electronics

Special Issue Reprint

Security and Privacy for Modern Wireless Communication Systems

2nd Edition

Edited by
Tao Huang, Shihao Yan, Guanglin Zhang, Tsz Hon Yuen, Yohan Park, Changhoon Lee and Jusak Jusak

mdpi.com/journal/electronics

MDPI

Security and Privacy for Modern Wireless Communication Systems, 2nd Edition

Security and Privacy for Modern Wireless Communication Systems, 2nd Edition

Guest Editors

Tao Huang
Shihao Yan
Guanglin Zhang
Tsz Hon Yuen
Yohan Park
Changhoon Lee
Jusak Jusak

MDPI

Basel • Beijing • Wuhan • Barcelona • Belgrade • Novi Sad • Cluj • Manchester

Guest Editors

Tao Huang
College of Science and Engineering
James Cook University
Smithfield
Australia

Shihao Yan
School of Science
Edith Cowan University
Joondalup
Australia

Guanglin Zhang
Department of Communication Engineering
Donghua University
Shanghai
China

Tsz Hon Yuen
Department of Software Systems & Cybersecurity
Monash University
Clayton
Australia

Yohan Park
College of Engineering
Keimyung University
Daegu
Republic of Korea

Changhoon Lee
Department of Computer Science & Engineering
Seoul National University of Science and Technology
Seoul
Republic of Korea

Jusak Jusak
School of Science and Technology
James Cook University
Singapore
Singapore

Editorial Office
MDPI AG
Grosspeteranlage 5
4052 Basel, Switzerland

This is a reprint of the Special Issue, published open access by the journal *Electronics* (ISSN 2079-9292), freely accessible at: https://www.mdpi.com/journal/electronics/special_issues/42RXIDLX57.

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

Lastname, A.A.; Lastname, B.B. Article Title. *Journal Name* **Year**, *Volume Number*, Page Range.

ISBN 978-3-7258-4757-0 (Hbk)
ISBN 978-3-7258-4758-7 (PDF)
https://doi.org/10.3390/books978-3-7258-4758-7

© 2025 by the authors. Articles in this book are Open Access and distributed under the Creative Commons Attribution (CC BY) license. The book as a whole is distributed by MDPI under the terms and conditions of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) license (https://creativecommons.org/licenses/by-nc-nd/4.0/).

Contents

About the Editors . . . . . **vii**

Tao Huang and Jusak Jusak
Security and Privacy for Modern Wireless Communication Systems, 2nd Edition
Reprinted from: *Electronics* **2025**, *14*, 2379, https://doi.org/10.3390/electronics14122379 . . . . . **1**

Minseong Yoon, Dongjun Seo, Soyoung Kim and Keecheon Kim
V2X Network-Based Enhanced Cooperative Autonomous Driving for Urban Clusters in Real Time: A Model for Control, Optimization and Security
Reprinted from: *Electronics* **2025**, *14*, 1629, https://doi.org/10.3390/electronics14081629 . . . . . **8**

Chathuranga Sampath Kalutharage, Saket Mohan, Xiaodong Liu and Christos Chrysoulas
Enhancing Automotive Intrusion Detection Systems with Capability Hardware Enhanced RISC Instructions-Based Memory Protection
Reprinted from: *Electronics* **2025**, *14*, 474, https://doi.org/10.3390/electronics14030474 . . . . . **34**

Geumtae Kim, Eunyoung Seo, Yongwoo Lee, Young-Sik Kim and Jong-Seon No
Lazy Modular Reduction for NTT
Reprinted from: *Electronics* **2024**, *13*, 4887, https://doi.org/10.3390/electronics13244887 . . . . . **55**

Paul Scalise, Robert Garcia, Matthew Boeding, Michael Hempel, Hamid Sharif
An Applied Analysis of Securing 5G/6G Core Networks with Post-Quantum Key Encapsulation Methods
Reprinted from: *Electronics* **2024**, *13*, 4258, https://doi.org/10.3390/electronics13214258 . . . . . **73**

Hyewon Park, Seunghwan Son, Youngho Park and Yohan Park
Provably Quantum Secure Three-Party Mutual Authentication and Key Exchange Protocol Based on Modular Learning with Error
Reprinted from: *Electronics* **2024**, *13*, 3930, https://doi.org/10.3390/electronics13193930 . . . . . **95**

IoT-GChain: Internet of Things-Assisted Secure and Tractable Grain Supply Chain Framework Leveraging Blockchain
IoT-GChain: Internet of Things-Assisted Secure and Tractable Grain Supply Chain Framework Leveraging Blockchain
Reprinted from: *Electronics* **2024**, , 3740, https://doi.org/10.3390/electronics13183740 . . . . . **120**

Eduard Zadobrischi and Ștefan Havriliuc
Enhancing Scalability of C-V2X and DSRC Vehicular Communication Protocols with LoRa 2.4 GHz in the Scenario of Urban Traffic Systems
Reprinted from: *Electronics* **2024**, *13*, 2845, https://doi.org/10.3390/electronics13142845 . . . . . **139**

Nut Aroon, Vicky Liu, Luke Kane, Yuefeng Li, Aklilu Daniel Tesfamicael and Matthew McKague
An Architecture of Enhanced Profiling Assurance for IoT Networks
Reprinted from: *Electronics* **2024**, *13*, 2832, https://doi.org/10.3390/electronics13142832 . . . . . **160**

Brianna Bace, Elisabeth Dubois and Unal Tatar
Resilience against Catastrophic Cyber Incidents: A Multistakeholder Analysis of Cyber Insurance
Reprinted from: *Electronics* **2024**, *13*, 2768, https://doi.org/10.3390/electronics13142768 . . . . . **181**

Giulio Tripi, Antonio Iacobelli, Lorenzo Rinieri, Marco Prandini
Security and Trust in the 6G Era: Risks and Mitigations
Reprinted from: *Electronics* **2024**, *13*, 2162, https://doi.org/10.3390/electronics13112162 . . . . . **204**

Gianmarco Baldini and Fausto Bonavitacola
LoRa Radio Frequency Fingerprinting with Residual of Variational Mode Decomposition and Hybrid Machine-Learning/Deep-Learning Optimization
Reprinted from: *Electronics* **2024**, *13*, 1925, https://doi.org/10.3390/electronics13101925 . . . . . **222**

Xabiel G. Pañeda, David Melendi, Víctor Corcoba, Alejandro G. Pañeda, Roberto García and Dan García
Forensic Analysis of File Exfiltrations Using AnyDesk, TeamViewer and Chrome Remote Desktop
Reprinted from: *Electronics* **2024**, *13*, 1429, https://doi.org/10.3390/electronics13081429 . . . . . **239**

About the Editors

Tao Huang

Tao Huang received his Ph.D. in Electrical Engineering from the University of New South Wales, Australia, in 2016; his M.Eng. degree in Sensor System Signal Processing from the University of Adelaide, Australia, in 2007; and his B.Eng. degree in Electronics and Information Engineering from Huazhong University of Science and Technology, China, in 2003. He is currently a Senior Lecturer at James Cook University (JCU), Cairns, Australia. Dr. Huang has held positions as an Endeavour Australia Cheung Kong Research Fellow, a visiting scholar at the Chinese University of Hong Kong, a research associate at the University of New South Wales, and a postdoctoral research fellow at JCU. Before entering academia, he held industry roles as a senior engineer, senior data scientist, project team lead, and technical lead. He is also a co-inventor of an international patent related to MIMO systems. His awards include the Australian Postgraduate Award, the Engineering Research Award from the University of New South Wales, the Best Paper Award at IEEE WCNC (2011), the IEEE Outstanding Leadership Award (2022), the IEEE Access Outstanding Associate Editor (2023 and 2024), and the JCU Citation for Outstanding Contribution to Student Learning (2022). Dr. Huang currently serves as Vice Chair of the IEEE Northern Australia Section and Chair of the local MTT-S/ComSoc Chapter. He has also held leadership roles in international conferences, including workshop co-chair, publication co-chair, technical program committee chair, program vice chair, and symposium chair. He is an Associate Editor for the *Scientific Reports*, *IEEE Open Journal of the Communications Society*, *IEEE Access*, and *IET Communications*. His research interests include deep learning, intelligent sensing, computer vision, pattern recognition, wireless communications, system optimization, electronic systems, and IoT security.

Shihao Yan

Shihao Yan received a Ph.D. degree in Electrical Engineering from the University of New South Wales (UNSW), Sydney, Australia, in 2015. He received a B.S. in Communication Engineering and an M.S. in Communication and Information Systems from Shandong University, Jinan, China, in 2009 and 2012, respectively. He was a Postdoctoral Research Fellow at the Australian National University, a University Research Fellow at Macquarie University, and a Senior Research Associate at the School of Electrical Engineering and Telecommunications, UNSW, Sydney, Australia. He is currently a Senior Lecturer in the School of Science, Edith Cowan University (ECU), Perth, Australia. He is also the Team Lead of Emerging Technologies for Cybersecurity in the Security Research Institute (SRI) at ECU. He was a Technical Co-Chair and Panel Member of a number of IEEE conferences and workshops, including the IEEE GlobeCOM 2018 Workshop on Trusted Communications with Physical Layer Security and IEEE VTC 2017 Spring Workshop on Positioning Solutions for Cooperative ITS. He was also awarded the Endeavour Research Fellowship by the Department of Education, Australia. His current research interests are in the areas of signal processing for wireless communication security and privacy, including covert communications, covert sensing, location spoofing detection, physical layer security, IRS-aided wireless communications, and UAV-aided communications.

Guanglin Zhang

Guanglin Zhang received a B.S. degree in Applied Mathematics from Shandong Normal University, Jinan, China, in 2003; an M.S. degree in Operational Research and Cybernetics from Shanghai University, Shanghai, China, in 2006; and a Ph.D. in Information and Communication Engineering from Shanghai Jiao Tong University, Shanghai, in 2012. From 2013 to 2014, he served as a Postdoctoral Research Associate at the Institute of Network Coding, The Chinese University of Hong Kong. He joined Donghua University as an Associate Professor in 2014 and was promoted to full Professor in 2017. From 2015 to 2021, he was the Department Chair of Communication Engineering. From 2020 to 2023, he acted as the Associate Dean of the College of Information Science and Technology, Donghua University. Currently, he is serving as the Special Appointment Eastern Scholar Professor, the Director of the Office of Talent Affairs, and the Director of the Department of Human Resources at Donghua University. He has previously served as a Technical Program Committee Member for IEEE Globecom 2016–2017; IEEE ICC 2014, 2015, and 2017; IEEE VTC2017-Fall; IEEE/CIC ICCC 2014; and WCSP 2014, APCC 2013, and WASA 2012. He has also served as the Local Arrangement Chair of ACM TURC 2017 and the Vice TPC Co-Chair of ACM TURC 2018. He is currently an Editor on the Editorial Board of *China Communications*. His research interests include online algorithms, the capacity scaling of wireless networks, vehicular networks, smart microgrids, and mobile edge computing.

Tsz Hon Yuen

Tsz Hon Yuen is an associate professor in the Department of Software Systems and Cybersecurity at Monash University. He currently serves as the Director of Research in the department, and as the Associate Director (Technology) of the Fintech Impact Lab of Monash University. He received his Ph.D. degree from the University of Wollongong in 2010. His research interests include cryptography (such as public key encryption, digital signatures, and identity-based encryption), privacy-preserving protocols (such as anonymous credentials and zero-knowledge proof systems), and blockchain (such as payment channels and confidential transactions). He has published more than 50 technical papers, including top journals and conferences such as Crypto, Eurocrypt, CCS, and IEEE S&P. He received the Best Paper Award in ESORICS 2014. He is also active in industrial fields, with more than 10 patents. He has served as the PC members and reviewers for many security and cryptography conferences, workshops, and journals.

Yohan Park

Yohan Park received his B.S., M.S., and Ph.D. in electronic engineering from Kyungpook National University, Daegu, South Korea in 2006, 2008, and 2013, respectively. Currently, he is an Assistant Professor in the Department of Computer Engineering at Keimyung University in Daegu. Prior to this, he held the same position at Nazarene University from 2017 to 2019, worked as a BK21 Postdoctoral Researcher at Kyungpook National University from 2016 to 2017, and lectured there from 2014 to 2016. In addition, he held a postdoctoral research position at Singapore National University from 2013 to 2014. Dr. Park's research interests include computer networks, mobile security, blockchain, and the Internet of Things (IoT).

Changhoon Lee

Changhoon Lee is a distinguished scholar and leading expert in the fields of information security and computer science. He earned his Ph.D. from the Graduate School of Information Management and Security (GSIMS) at Korea University, a premier institution in South Korea. Over the course of his career, Prof. Lee has held several prominent academic positions, including Research Professor at the Center for Information Security Technologies at Korea University and Professor in the School of Computer Engineering at Hanshin University. He is currently serving as a Professor in the Department of Computer Science and Engineering at Seoul National University of Science and Technology (SeoulTech). Prof. Lee plays an active role in the global academic and professional communities, contributing as a conference chair, program committee member, and Guest Editor for esteemed international journals. His research interests are wide-ranging and include cyber threat intelligence (CTI), information security, cryptography, digital forensics, IoT security, and blockchain technologies. He is a member of several leading academic and professional societies, including IEEE, IEEE Computer Society, IEEE Communications Society, ACM, IACR, ICRP, KIISC, KDFS, and ICTPS. Through his extensive expertise, prolific research, and academic service, Prof. Lee has earned recognition both nationally and internationally as a respected authority in the cybersecurity domain.

Jusak Jusak

Jusak Jusak received Ph.D. degree in electrical engineering from Royal Melbourne Institute of Technology (RMIT) University, Melbourne, Australia. He is now a Senior Lecturer in Internet of Things, School of Science and Technology, James Cook University (Singapore Campus). He was a Postdoctoral Research Fellow with Massey University, Palmerston North, New Zealand, working on a Next Generation Networks research project supported by the Telecom New Zealand. He has been a Senior Lecturer in Computer Engineering department, University of Dinamika, Surabaya, Indonesia, since 2011. Dr. Jusak is a member of IEEE and IEEE Communications Society with research interest in the Internet of Things applications for healthcare and the development of lightweight blockchain architecture to achieve the need for a scalable, fault-tolerant, secure, and traceable private blockchain network to match with the requirements for healthcare Internet of Things. He is also focusing on the creation and deployment of tiny machine learning, a lightweight machine learning method for devices with limited resources.

electronics

MDPI

Editorial

Security and Privacy for Modern Wireless Communication Systems, 2nd Edition

Tao Huang [1,*] and Jusak Jusak [2]

[1] College of Science and Engineering, James Cook University, Cairns, QLD 4878, Australia
[2] School of Science and Technology, James Cook University, Singapore 387380, Singapore; jusak.jusak@jcu.edu.au
* Correspondence: tao.huang1@jcu.edu.au

1. Introduction

Wireless communication systems are undergoing rapid transformation, driven by the advent of 6G technologies, the proliferation of Internet of Things (IoT) devices, and the growing reliance on intelligent and connected infrastructure. As wireless networks expand in complexity and capability, ensuring security and privacy becomes increasingly vital—not only to protect sensitive data but also to maintain trust, system integrity, and operational reliability in critical applications such as remote healthcare, autonomous vehicles, and smart manufacturing.

This Special Issue, titled "Security and Privacy for Modern Wireless Communication Systems, 2nd Edition", aims to address urgent challenges by highlighting the latest innovations in protocols, architectures, software, and hardware solutions designed to enhance the security and privacy of contemporary wireless networks. Unlike traditional systems, modern wireless environments must accommodate a diverse array of resource-constrained devices, ultra-low-latency applications, and emerging vulnerabilities introduced by technologies such as intelligent reflective surfaces, blockchain, edge/fog/cloud computing, and artificial intelligence (AI).

The 12 papers published in this Special Issue discuss a wide range of research topics. Several contributions focus on lightweight and quantum-resistant cryptography. These include discussions on the use of the CHERI architecture to create memory-safe intrusion detection systems and MLWE-based key exchange methods for post-quantum secure authentication. Other papers address network-level threats, proposing innovative solutions such as hybrid communication architectures, blockchain-based security for supply chains, and frameworks for cyber insurance to handle catastrophic cyber incidents. Additionally, this Special Issue highlights advancements in physical-layer security and edge security, including methods for physical-layer fingerprinting of LoRa devices, detection of covert communications, and the development of trust models within 6G edge–node–cloud ecosystems. Practical aspects of cybersecurity are also examined; several studies explore log-based forensic analysis, testbed validation, and real-time monitoring to detect remote access and file exfiltration.

Collectively, the papers cover foundational research and applied solutions across topics such as 5G/6G communications, physical-layer key generation, lightweight cryptography, blockchain networks, anonymity in data transmission, the use of machine learning/deep learning, and security and privacy designs tailored for Vehicle-to-Everything (V2X), supply chain, and LoRa networks. Furthermore, this Special Issue underscores the importance of prototyping, forensic tools, and testbed evaluations to ensure real-world viability.

By collating these contributions from leading researchers, this Special Issue provides a comprehensive and timely overview of the state of the art while also identifying emerging research directions that will shape the future of secure and private wireless communication systems.

2. Summary of the Contributions

The first contribution, provided by Yoon et al., proposes a cooperative autonomous driving system tailored for urban environments by leveraging a V2X network-based framework. Unlike previous studies that primarily focus on highway scenarios, this work addresses the unique challenges of urban driving, such as managing traffic lights, pedestrian interactions, and dynamic vehicle clusters. A key innovation is the introduction of the Throttle–Steer–Brake (TSB) driving technique, which enables following vehicles to replicate the control commands of a leading vehicle, significantly reducing computational overhead and improving real-time responsiveness. The system also incorporates an HSV-based traffic light perception method, which outperforms AI-based models in speed and accuracy, and a pedestrian detection mechanism using V2P communication to trigger emergency braking when necessary. The system employs pseudonymized vehicle identifiers, AES encryption, and EdDSA digital signatures to ensure secure and reliable communication. Validated in the CARLA simulator, the proposed approach demonstrated approximately 466 times faster computation using TSB driving compared to waypoint-based methods and achieved 99.58% accuracy in traffic light recognition. The results confirm the feasibility of real-time, safe, and secure cooperative autonomous driving in urban settings. Future work will focus on deploying the system in real-world conditions, accounting for network latency and failures, and enhancing TSB driving in complex traffic scenarios.

The second contribution, provided by Kalutharage et al., explores the enhancement of automotive intrusion detection systems (IDSs) by integrating CHERI (Capability Hardware Enhanced RISC Instructions), a hardware-enforced memory protection architecture. Motivated by the growing cybersecurity risks in connected vehicles, especially threats like IP spoofing and rule manipulation, the authors propose a CHERI-enabled IDS that isolates memory and safeguards detection rules from unauthorized access. The system enforces strict spatial and temporal memory safety by embedding capability-based access controls directly into the IDS design, effectively blocking advanced attacks. Experimental results show that the CHERI-enhanced IDS achieves 100% detection of spoofed packets and prevents all unauthorized rule modifications, even under real-time constraints. Implemented on the ARM Morello board, the solution operates within the latency thresholds required for automotive applications. Compared to traditional software-based IDS solutions, this hardware-centric approach ensures rule integrity, enhances detection robustness, and maintains system performance in resource-constrained environments. The authors conclude that CHERI-based IDS frameworks offer a scalable, secure foundation for future automotive cybersecurity systems, with potential for further optimization and integration with machine learning for anomaly detection.

The third contribution, authored by Kim et al., introduces LazyNTT, a novel approach to optimizing the Number Theoretic Transform (NTT), which is fundamental to lattice-based post-quantum cryptographic schemes. Traditional NTT relies heavily on Montgomery multiplication for modular reduction, which, while efficient, is still more costly than standard integer multiplication. To solve this problem, the authors propose two variants, SM-LazyNTT and SSM-LazyNTT, which strategically replace certain Montgomery multiplications with standard multiplications during intermediate stages of the NTT computation. These "lazy" replacements postpone modular reductions without affecting the final correctness, reducing computational overhead. SM-LazyNTT alternates between

standard and Montgomery multiplications, while SSM-LazyNTT allows two consecutive standard multiplications before performing a Montgomery multiplication. The method is validated using parameters from Falcon and Kyber cryptosystems, showing performance improvements of up to 9% and 28%, respectively. The approach is generalizable to more complex variants and incurs only minimal additional memory requirements. Experimental results confirm that LazyNTT significantly reduces cycle counts while preserving output correctness, making it a promising optimization for post-quantum cryptography and potentially homomorphic encryption.

The fourth contribution, produced by Scalise et al., presents an applied study on enhancing the cybersecurity of 5G and future 6G core networks by integrating post-quantum cryptography (PQC), specifically Key Encapsulation Methods (KEMs), into the free5GC open-source platform. Recognizing that the virtualized nature of modern 5G Core Networks increases the attack surface, the authors aim to address the threat posed by quantum computing, especially the risk of "Store Now, Decrypt Later" (SNDL) attacks, by replacing vulnerable key exchange methods in TLS with quantum-resistant alternatives like CRYSTALS-Kyber. The team implements these PQC KEMs into TLS 1.3 for inter-VNF communication within free5GC, leveraging the Open Quantum Safe (OQS) project and modifying the system using Go-OpenSSL bindings for compatibility. The evaluation demonstrates that PQC integration results in only a minimal increase in latency and bandwidth overhead during UE (User Equipment) connection establishment. Specifically, standalone Kyber KEMs outperform the widely used elliptic curve X25519 in latency performance. The authors conclude that PQC-enabled 5G cores are both feasible and beneficial, offering strong quantum-resilient security with negligible impact on real-time system performance. This work sets a precedent for integrating PQC in production-grade 5G and 6G systems to future-proof their cryptographic integrity.

The fifth contribution, authored by Park et al., introduces a provably quantum-secure three-party mutual authentication and key exchange protocol based on the Modular Learning with Error (MLWE) problem. Motivated by the limitations of existing work in the literature, which is vulnerable to several attacks, including replay attacks, impersonation attacks, insider attacks, and denial-of-service attacks, the authors develop a more robust and efficient protocol that integrates biometric data for three-factor authentication. The proposed scheme ensures that users and servers can mutually authenticate and securely exchange session keys even with quantum adversaries. Formal security validation is conducted using BAN logic, the Real-or-Random (RoR) model, and the AVISPA tool, confirming resistance against various attacks and verifying mutual authentication. Performance analysis demonstrates computational efficiency and improved security features compared to related protocols. By leveraging the strengths of MLWE in combination with biometric authentication and secure hash-based operations, the protocol offers a scalable and practical solution for post-quantum secure communications in multi-party settings.

The sixth contribution, provided by Thakur et al., presents IoT-GChain, a blockchain- and IoT-assisted grain supply chain framework aimed at enhancing transparency, traceability, and security in grain distribution. The system addresses long-standing issues in traditional supply chains, such as middleman exploitation, record tampering, and lack of real-time grain condition monitoring, by integrating non-fungible tokens (NFTs), smart contracts, and sensor data. Each grain lot is tokenized as an NFT containing critical information, including quality metrics, weight, and environmental conditions (e.g., temperature and humidity) collected via DHT22 sensors and GPS modules. These NFTs are dynamically updated as grain moves through a multi-tiered supply chain, from central authorities to local retailers. Smart contracts deployed on the Ethereum blockchain automate verification, traceability, and ownership transfer, with real-time updates stored in Firebase and metadata

hosted on IPFS. The framework includes a user interface with QR code scanning for grain status verification and integrates with MetaMask for secure interaction. Security is assessed using SolidityScan, with a reported security score of 84.56. Experimental validation confirms the system's capability to streamline operations, enhance user trust, and prevent fraud. This layered, decentralized design demonstrates a scalable and tamper-resistant model for modernizing grain supply chains in developing economies.

The seventh contribution, provided by Zadobrischi et al., explores a hybrid communication architecture that enhances the scalability and interoperability of vehicular communication protocols, specifically C-V2X, DSRC, and LoRa 2.4 GHz, within urban Intelligent Transportation Systems (ITS). The paper is motivated by the need to improve road safety, reduce traffic congestion, and support smart city development amidst increasing urbanization. The authors assess the strengths and limitations of each protocol: DSRC offers low latency for short-range safety-critical applications; C-V2X, built on LTE/5G, supports high-throughput and long-range communication; and LoRa 2.4 GHz provides energy-efficient, long-distance communication ideal for non-critical infrastructure monitoring. Through simulation and field experiments, the study demonstrates that LoRa 2.4 GHz can complement DSRC and C-V2X by filling communication gaps in high-mobility or obstructed urban areas. Using advanced signal modeling and ray tracing, the authors validate LoRa's effectiveness under urban interference and confirm its low bit error rate and acceptable latency. A proposed architecture integrates LoRa nodes across vehicles, infrastructure, and pedestrians, linking them to a gateway that communicates with cloud-based analytics platforms. The study concludes that combining these technologies offers a scalable, low-cost, and resilient communication framework for future ITS applications and recommends further research into interoperability, dynamic parameter tuning, and security to support autonomous vehicle networks and smart city infrastructure.

The eighth contribution, produced by Aroon et al., proposes the Enhanced Profiling Assurance (EPA) architecture to address the limitations of the Manufacturer Usage Description (MUD) standard in detecting malicious activity within IoT networks. While the MUD framework enforces stateless access control rules based on predefined profiles, it fails to inspect the state of ongoing communications, potentially allowing traffic amplification and other attacks to bypass detection. To overcome this, the EPA incorporates a two-layer Intrusion Detection and Prevention System (IDPS) that combines both stateless and stateful inspection. It leverages three-way decision theory—allow, deny, and uncommitted—to support ongoing analysis of uncertain network behaviors rather than forcing binary decisions. The architecture also integrates a Network Behavior Analysis (NBA) system that consults dynamically updated knowledge bases of normal, abnormal, and uncertain behaviors. This enables real-time packet inspection and adaptive rule updates for improved security enforcement. Experimental evaluations demonstrate that the EPA significantly reduces false negatives compared to the conventional MUD approach. The study provides a flexible and scalable implementation model for small-scale IoT networks, with potential for broader application in securing edge environments against evolving cyber threats.

The ninth contribution, provided by Bace et al., offers a comprehensive multistakeholder analysis on how the United States can enhance resilience against catastrophic cyber incidents (CCIs) through the strategic use of cyber insurance and potential federal intervention. Using content analysis of 56 public comments submitted in response to a U.S. Treasury Department Request for Information, the authors investigate four key areas: how CCIs are defined, how they can be mitigated, whether the current cyber insurance sector can manage such risks, and what role the government should play. The study finds that while stakeholders agree that CCIs involve large-scale financial loss, cross-sector impact, and critical infrastructure disruption, no consistent definition complicates risk modeling. Many

comments support the use of basic cybersecurity measures (e.g., MFA, patching, employee training) to mitigate CCIs and suggest that insurers should mandate such practices. Most respondents believe the private insurance sector lacks the capacity or willingness to manage CCIs alone, advocating for a federal backstop, modeled partly on the Terrorism Risk Insurance Program (TRIP), to stabilize the market. Concerns such as moral hazard, physical loss exclusions, and excessive regulation are acknowledged but seen as manageable through careful design. The study concludes with a proposed framework for a federal backstop and offers actionable insights for policymakers to strengthen national cybersecurity resilience.

The tenth contribution, authored by Tripi et al., provides a comprehensive analysis of security, privacy, and trust challenges in 6G networks, structured around a layered security architecture that spans the physical, connection, and application layers. The paper begins by highlighting the transformative potential of 6G, including its ultra-low-latency capabilities, massive connectivity, and support for emerging technologies like XR, CRAS, BCI, and blockchain, while emphasizing that these advances demand a proactive approach to security from the ground up. At the physical layer, the authors detail vulnerabilities like eavesdropping, jamming, and spoofing and propose mitigations including Reconfigurable Intelligent Surfaces (RISs), Friendly Jamming, NOMA, frequency hopping, and reinforcement learning. Key risks such as DoS, MitM, and replay attacks for the connection layer are countered through Quantum Key Distribution (QKD), network slicing, and deep learning-based intrusion detection systems (IDSs). The application layer faces threats like social engineering, for which biometric authentication, Quantum Homomorphic Cryptography (QHC), and Authentication and Key Agreement (AKA) protocols are proposed. The paper also explores trust management models involving blockchain, smart contracts, and trust anchors, advocating for decentralized, transparent mechanisms. The authors offer a forward-looking blueprint for building secure, resilient, and trustworthy 6G infrastructures by integrating Zero Trust Architecture- and AI-driven defense strategies.

The eleventh contribution, from Baldini et al., presents a novel hybrid machine learning and deep learning approach for Radio Frequency Fingerprinting (RFF) of LoRa devices, leveraging the residuals from Variational Mode Decomposition (VMD) instead of the commonly used modes. The authors argue that residual signals, which exclude common modulation components, better capture device-specific physical characteristics essential for accurate fingerprinting. The authors propose a feature-driven segmentation method to handle the computational burden of applying CNNs to spectrograms derived from these signals. This involves calculating Local Binary Pattern (LBP) and Shannon entropy to identify and extract the most discriminative regions of the spectrograms for input into a multi-head attention CNN. The method is evaluated on a public dataset of 10 LoRa devices, each contributing 500 signal bursts. Experimental results show that this residual-based and segmented spectrogram approach significantly improves classification accuracy (up to 91.9%) while reducing computational time compared to baseline methods using full spectrograms or original signals. The study concludes that combining residual analysis with targeted image reduction and deep learning leads to a more efficient and accurate RFF framework suitable for IoT environments with constrained resources.

The final contribution, from Pañeda et al., presents a forensic analysis of file exfiltration activities using three widely used remote desktop applications: TeamViewer, AnyDesk, and Chrome Remote Desktop. Motivated by the surge in remote work during the COVID-19 pandemic and the associated cybersecurity risks, the authors investigate whether digital forensics can effectively determine what files were exfiltrated, as well as when and by whom, using such tools. The study simulates both push (the user exfiltrates data to an external device) and pull (the external user accesses internal data) scenarios on virtual machines. Using Wireshark, forensic evidence is collected from application logs, system

event logs, browser histories, and network traffic. Key findings show that TeamViewer provides detailed logs, including file names, paths, and timestamps, while AnyDesk reveals connection details and remote IPs but lacks file names in logs. Chrome Remote Desktop, being browser-based, offers the least forensically useful logging, with evidence primarily found in browser histories and system events. Network behavior, such as traffic patterns and acknowledgement timings, can be analyzed to infer exfiltration events despite encryption. The study concludes that while some evidence is partial, combining multiple sources (logs, audits, traffic) can yield legally viable digital evidence and early warning indicators for unauthorized data transfers.

3. Conclusions

This Special Issue features a diverse collection of cutting-edge research that addresses the evolving challenges of security and privacy in modern wireless communication systems. The selected papers strike a balance between theoretical advancements and practical solutions, covering a wide range of topics, including post-quantum cryptography, physical-layer security, network intrusion detection, and forensic analysis. As wireless technologies advance toward 6G and beyond, and as the integration of AI, IoT, and edge computing deepens, ensuring secure and trustworthy communication infrastructures remains a critical priority. Through this collection, we not only hope to platform current research but also inspire further innovation in this vital field.

Author Contributions: Writing—original draft preparation, T.H.; writing—review and editing, J.J. and T.H. All authors have read and agreed to the published version of the manuscript.

Funding: This article received no external funding.

Acknowledgments: We would like to express our sincere gratitude to all the Guest Editors of this Special Issue for their invaluable contributions throughout the editorial process. Their expertise, dedication, and thoughtful oversight were instrumental in ensuring the high quality and thematic coherence of the published works. We also extend our appreciation to the reviewers for their time and constructive feedback and to the authors for submitting their outstanding research. Finally, we thank the editorial team at *Electronics* for their continuous support and professionalism in managing this Special Issue.

Conflicts of Interest: The authors declare no conflicts of interest.

List of Contributions

1. Yoon, M.; Seo, D.; Kim, S.; Kim, K. V2X Network-Based Enhanced Cooperative Autonomous Driving for Urban Clusters in Real Time: A Model for Control, Optimization and Security. *Electronics* **2025**, *14*, 1629. https://doi.org/10.3390/electronics14081629.
2. Kalutharage, C.S.; Mohan, S.; Liu, X.; Chrysoulas, C. Enhancing Automotive Intrusion Detection Systems with Capability Hardware Enhanced RISC Instructions-Based Memory Protection. *Electronics* **2025**, *14*, 474. https://doi.org/10.3390/electronics14030474.
3. Kim, G.; Seo, E.; Lee, Y.; Kim, Y.-S.; No, J.-S. Lazy Modular Reduction for NTT. *Electronics* **2024**, *13*, 4887. https://doi.org/10.3390/electronics13244887.
4. Scalise, P.; Garcia, R.; Boeding, M.; Hempel, M.; Sharif, H. An Applied Analysis of Securing 5G/6G Core Networks with Post-Quantum Key Encapsulation Methods. *Electronics* **2024**, *13*, 4258. https://doi.org/10.3390/electronics13214258.
5. Park, H.; Son, S.; Park, Y.; Park, Y. Provably Quantum Secure Three-Party Mutual Authentication and Key Exchange Protocol Based on Modular Learning with Error. *Electronics* **2024**, *13*, 3930. https://doi.org/10.3390/electronics13193930.

6. Thakur, K.S.; Ahuja, R.; Singh, R. IoT-GChain: Internet of Things-Assisted Secure and Tractable Grain Supply Chain Framework Leveraging Blockchain. *Electronics* **2024**, *13*, 3740. https://doi.org/10.3390/electronics13183740.
7. Zadobrischi, E.; Havriliuc, Ș. Enhancing Scalability of C-V2X and DSRC Vehicular Communication Protocols with LoRa 2.4 GHz in the Scenario of Urban Traffic Systems. *Electronics* **2024**, *13*, 2845. https://doi.org/10.3390/electronics13142845.
8. Aroon, N.; Liu, V.; Kane, L.; Li, Y.; Tesfamicael, A.D.; McKague, M. An Architecture of Enhanced Profiling Assurance for IoT Networks. *Electronics* **2024**, *13*, 2832. https://doi.org/10.3390/electronics13142832.
9. Bace, B.; Dubois, E.; Tatar, U. Resilience against Catastrophic Cyber Incidents: A Multistakeholder Analysis of Cyber Insurance. *Electronics* **2024**, *13*, 2768. https://doi.org/10.3390/electronics13142768.
10. Tripi, G.; Iacobelli, A.; Rinieri, L.; Prandini, M. Security and Trust in the 6G Era: Risks and Mitigations. *Electronics* **2024**, *13*, 2162. https://doi.org/10.3390/electronics13112162.
11. Baldini, G.; Bonavitacola, F. LoRa Radio Frequency Fingerprinting with Residual of Variational Mode Decomposition and Hybrid Machine-Learning/Deep-Learning Optimization. *Electronics* **2024**, *13*, 1925. https://doi.org/10.3390/electronics13101925.
12. Pañeda, X.G.; Melendi, D.; Corcoba, V.; G. Pañeda, A.; García, R.; García, D. Forensic Analysis of File Exfiltrations Using AnyDesk, TeamViewer and Chrome Remote Desktop. *Electronics* **2024**, *13*, 1429. https://doi.org/10.3390/electronics13081429.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

electronics

MDPI

Article

V2X Network-Based Enhanced Cooperative Autonomous Driving for Urban Clusters in Real Time: A Model for Control, Optimization and Security

Minseong Yoon [1], Dongjun Seo [1], Soyoung Kim [2] and Keecheon Kim [1,*]

[1] Department of Computer Science and Engineering, Konkuk University, Seoul 05029, Republic of Korea; tlssla0531@konkuk.ac.kr (M.Y.); ggaebi99@konkuk.ac.kr (D.S.)
[2] Department of IT Convergence and Information Security, Konkuk University, Seoul 05029, Republic of Korea; ypd07026@konkuk.ac.kr
* Correspondence: kckim@konkuk.ac.kr; Tel.: +82-2-450-3518

Abstract: For the commercialization of connected vehicles and smart cities, extensive research is carried out on autonomous driving, Vehicle-to-Everything (V2X) communication, and platooning. However, limitations remain, such as restrictions to highway environments, and studies are conducted separately due to challenges in ensuring reliability and real-time performance under external influences. This paper proposes a cooperative autonomous driving system based on V2X network implemented in the CARLA simulator, which simulates an urban environment to optimize vehicle-embedded systems and ensure safety and real-time performance. First, the proposed Throttle–Steer–Brake (TSB) driving technique reduces the computational overhead for following vehicles by utilizing the control commands of a leading vehicle. Second, a V2X network is designed to support object perception, cluster escape, and joining. Third, an urban perception system is developed and validated for safety. Finally, pseudonymized vehicle identifiers, Advanced Encryption Standard (AES), and the Edwards-curve Digital Signature Algorithm (EdDSA) are employed for data reliability and security. The system is validated in processing time and accuracy, confirming feasibility for real-world application. TSB driving demonstrates a computation speed approximately 466 times faster than conventional waypoints-based driving. Accurate urban perception and V2X communication enable safe cluster escape and joining, establishing a foundation for cooperative autonomous driving with improved safety and real-time capabilities.

Keywords: cooperative autonomous driving; V2X network; TSB driving; control optimization; vehicle privacy security

1. Introduction

In cooperative autonomous driving, related studies, such as autonomous driving control, urban environmental perception, and V2X networks, are often conducted separately. In the existing research on platooning and connected vehicles, the ultra-low latency and high reliability of V2X networks for real-time vehicle communication are essential requirements [1]. In addition, most research focuses primarily on highway environments. According to a 2020 survey by the International Road Federation, the motorway or highway network measures 105,948 km, whereas the main or national road network extends to 251,957 km [2]. The length of national roads is approximately 2.5 times longer than that of highways. Therefore, it is necessary to develop a cooperative autonomous driving system tailored to urban environments [3].

This paper integrates cooperative autonomous driving technologies and evaluates them in a CARLA simulator-based urban environments that resembles real-world conditions. The experiment prioritizes real-time performance and safety. If both real-time capability and safety are ensured, the system is expected to enable immediate deployment and commercialization in connected vehicles and smart cities in the near future.

The vehicle is controlled using a Proportional–Integral–Differential (PID) controller and the Pure Pursuit algorithm [4]. The CARLA server is configured, and cluster information is generated and managed by transmitting and receiving data such as vehicle ID and x and y coordinates. Based on the established V2X network, the throttle, steer, and brake values from the leading vehicle in cooperative autonomous driving are forwarded to the following vehicle to realize TSB driving. Urban environments perception compares traffic light recognition based on Artificial Intelligence (AI) with that based on Hue Saturation Value (HSV), and HSV-based perception is selected due to its superior performance [5]. Pedestrian recognition and stop procedures are triggered based on pedestrian location data exchanged with the edge server. If the distance between the vehicle and the pedestrian is less than a predefined threshold, an emergency stop is initiated. The most important cluster escape and joining algorithms in cooperative autonomous driving operate based on the vehicle's order on the server and its distance from the leading vehicle. Additionally, to ensure secure communication, a security framework is constructed using pseudonymized vehicle identifiers, AES encryption, and EdDSA digital signatures to protect privacy and maintain data integrity [6].

TSB driving demonstrates a time gain approximately 466 times greater than that of waypoints-based driving, indicating its suitability for real-time vehicle communication and control. HSV-based traffic light perception achieves an accuracy of 99.58% and operates nearly 100 times faster than AI-based perception. The urban environments, including pedestrians, are successfully recognized, and vehicles perform safe stops within lanes. In this study, vehicles navigate urban environments at safe speeds and successfully perform cluster escape and joining. In conclusion, the integrated safety of cooperative autonomous driving in urban environments is validated, and the security aspect of the V2X network is carefully considered. The proposed system is expected to contribute significantly to the foundational research in V2X Network and autonomous driving.

The structure of this paper is as follows: Section 2 presents a review of related work, Section 3 describes the methodology and models, Section 4 provides detailed experimental results, and Section 5 concludes the paper and discusses future research directions.

2. Related Works and Contribution

2.1. Platooning

Most preceding studies on platooning and cooperative autonomous driving focused on trucks or heavy-duty vehicles [7].

Platooning is typically based on a distributed architecture, where control operates within each vehicle. However, a method proposed in the existing research adopts a centralized control approach, in which a Multi-access Edge Computing (MEC) server collects vehicle data and manages the cluster [8]. In the experimental setup, the number of vehicles on the highway and in the cluster is set to 20, with an inter-vehicle distance of 10 m. This architecture offers greater scalability and safer management than a simple Vehicle-to-Vehicle (V2V) network. However, in urban environments, vehicle communication is often restricted, and the low-latency benefits of V2V communication are not fully exploited. In this study, we construct a server that manages information for all vehicles, integrating the strengths of both V2V and Vehicle-to-Infrastructure (V2I) communication. This enables safe cooperative autonomous driving in urban environments.

One of the key functions of platooning is perception, including vehicle detection. In the ARGO Project, a 3D vision system is constructed using two cameras to detect obstacles and vehicles. In addition, the system aims to track the preceding vehicle based on its position and speed. A target point is defined near the preceding vehicle, and in this study, the waypoints-based autonomous driving method for the latter vehicle uses the position of the preceding vehicle as a reference [9]. Since additional control strategies are required to enhance the performance of the cooperative autonomous driving system, this study introduces safe control method based on cluster escape and joining algorithms, as well as TSB driving.

Vehicular Ad hoc Networks (VANETs) are utilized to distribute the fundamental tasks of autonomous driving among agents. VANETs can be implemented using either traditional Internet Protocol (IP)-based networking or Information-Centric Networking (ICN). Constructing efficient VANETs involving thousands of vehicles in urban environments remains a major challenge [10]. In this study, a cluster of three vehicles is implemented in an urban setting, allowing vehicles to dynamically join and escape the cluster. The system's performance is evaluated in terms of time efficiency and computational overhead.

2.2. Autonomous Driving Control

Geometric and dynamic models are commonly employed in the design of autonomous vehicle control systems. The dynamic bicycle model provides a more accurate representation of vehicle motion, incorporating tire–road interactions. The Model Predictive Controller (MPC) is utilized to perform trajectory tracking by computing optimal control signals based on the vehicle's current state [11]. In this study, a predictive control model is designed based on the dynamic bicycle model and the MPC structure.

Autonomous driving can be modeled as a two-layer control architecture [12]. The first layer determines the target point of the vehicle using fuzzy logic, while the second layer employs conventional control methods to actuate the vehicle toward the target point defined by the first layer. This approach demonstrates strong performance in real-world autonomous vehicle control. In this study, a similar principle to waypoints-based driving is adopted, where the target position of the vehicle is defined accordingly. TSB driving controls the vehicle using the throttle, steer, brake values to reach the target position, and it is significant, in that the target position is updated in real time through the V2X network.

2.3. V2X Network

The V2X network technology is a next-generation communication framework that includes V2V, V2I, Vehicle-to-Pedestrian (V2P), and Vehicle-to-Network (V2N) communication. V2X communication technologies include Dedicated Short-Range Communication (DSRC) and Long-Term Evolution (LTE)-V2X, which are currently deployed, and 5th Generation (5G)-V2X, which is under development. V2X network offers advantages such as enhanced traffic safety, traffic flow optimization, and environmental sustainability. However, its commercialization remains premature, and the verification of its reliability and security is still insufficient. To improve this, prior research has proposed a comprehensive testing methodology for V2X network [13]. Security threats and real-world application scenarios have been systematically analyzed across multiple areas, including protocol conformance testing, functional testing, and performance evaluation. However, the results remain hypothetical due to a lack of real-world experimental validation. In this study, we conduct a test under real-world-like conditions and verify the practical reliability of the V2X network through empirical evaluation.

Numerous studies have conducted comprehensive reviews of the Resource Allocation (RA) mechanisms of Cellular Vehicle-to-Everything (C-V2X). Compared to conventional DSRC-based V2X network, C-V2X offers advantages such as broader coverage, higher

reliability, and Quality of Service (QoS) guarantees. When comparing LTE-V2X with 5th Generation New Radio (5G-NR) V2X, LTE-V2X provides strong reliability guarantees, while 5G-NR V2X enhances QoS and supports AI-based optimization. The advancements in 5G-NR V2X have been summarized, and future research directions have been suggested [14]. In this study, a V2X network is implemented in a simulated real-world environment with consideration for security.

A comprehensive tutorial on 5G-NR V2X communication technology and the corresponding standard, introduced in the 3rd Generation Partnership Project (3GPP) Release 16, is available [15]. 5G-NR V2X complements LTE-V2X by addressing its limitations and supporting V2X networks optimized for Intelligent Transport Systems (ITS). While LTE-V2X lacks a mechanism to verify packet reception, 5G-NR V2X introduces Hybrid Automatic Repeat Request (HARQ) and the Physical Sidelink Feedback Channel (PSFCH) to confirm transmission. In addition, 5G-NR V2X incorporates a QoS management framework that utilizes various V2X applications. This related research proposes using LTE-V2X for basic safety message transmission and 5G-NR V2X for enhanced functionality. The specifications of 5G-NR V2X are systematically organized, and its core technologies are explained in detail to suggest future research directions.

To implement cooperative autonomous driving, a V2X network is a critical component. Table 1 compares the characteristics of the V2X network, the V2V network, and the network in CARLA simulator utilized in this study [16,17]. The V2X network encompasses various communication entities, including Road Side Units (RSUs), pedestrians, network infrastructure, and government agencies, in addition to vehicles. It provides wide-area communication capabilities (extending several kilometers) by leveraging technologies such as 5G, LTE, Wi-Fi, DSRC, and C-V2X. In contrast, the V2V network is based on direct communication between vehicles and utilizes DSRC and C-V2X technologies, supporting communication within a range of approximately 800 m. In the case of the network in CARLA simulator, communication occurs between the server and clients, as well as between vehicles, and theoretically enables communication over an unlimited range. In terms of communication standards, the V2X network supports both Institute of Electrical and Electronics Engineers (IEEE) 802.11 p and IEEE 802.11 bd, whereas the V2V network is limited to IEEE 802.11 p [18,19]. Additionally, the V2X network serves comprehensive purposes, including traffic congestion mitigation, accident prevention, the enhancement of autonomous driving performance, and environmental impact reduction. Meanwhile, the V2V network primarily focuses on traffic flow optimization and accident reduction. In addition to these functions, the network in CARLA simulator offers a virtual environment for evaluating and improving algorithmic processing speed.

Table 1. Comparison between V2X network, V2V network, and network in CARLA simulator.

	V2X Network	**V2V Network**	**Network in CARLA Simulator**
Communication Agents	Vehicles, RSU, Pedestrian, Network, Government, etc.	Vehicles	Vehicles, Server, Clients
Communication Range	5G, LTE, Wi-Fi, DSRC, C-V2X (about km)	DSRC, C-V2X (about 800 m)	Unlimited
Communication Standard	IEEE 802.11 p, IEEE 802.11 bd	IEEE 802.11 p	
Purpose	Reduced Congestion and Incidents, Improved Completeness of Autonomous Driving, Environmental Pollution Mitigation	Traffic Flow Optimization, Reduced Congestion and Incidents	Traffic Flow Optimization, Reduced Congestion and Incidents, Improved Algorithm Processing Speed

2.4. Urban Environment Perception

A related study aims to improve traffic light perception performance in Advanced Driver Assistance Systems (ADASs) by enhancing You Only Look Once version 8 (YOLOv8) with meta-learning techniques. Existing YOLO and Convolutional Neural Network (CNN) models exhibit low accuracy under varying illumination conditions and are significantly affected by the size of the traffic lights and surrounding elements such as dataset characteristics. By applying the Model-Agnostic Meta-Learning (MAML) algorithm, an AI model that is robust to new environments is developed, maximizing performance in color differentiation. Unlike conventional methods that detect the entire traffic light as a bounding box, this approach detects only the illuminated section, thereby optimizing the labeling process. As a result, both accuracy is improved and computational overhead is reduced. Through feature weight optimization, the model further enhances color recognition by assigning higher weights to color features, resulting in improved accuracy and reduced computational overhead. This method demonstrates higher accuracy than the baseline YOLOv8 model under various conditions such as daytime, nighttime, and fog, while maintaining real-time performance with 30 Frames Per Second (FPS). However, it remains susceptible to errors when colors are similar, and the model size is relatively large at 25 million parameters, requiring substantial computational resources. AI-based perception methods still face limitations related to color confusion and latency [20]. Therefore, we aims to compare the proposed approach with a simpler HSV value-based method, and to identify a more accurate and faster object perception technique for vehicular environments through cross-validation.

Related research proposes a coarse-to-fine deep learning framework to improve the reliability of existing YOLO-based traffic light perception methods [21]. These existing methods struggle with perception due to the small size and diverse shapes of traffic lights, and they have difficulty distinguishing directional features, such as arrow indicators. To address these challenges, traffic light perception is divided into two stages. In the first stage, YOLOv8-nano is enhanced to detect small traffic lights and to group them into super classes for an initial search. In the second stage, the system accurately classifies traffic light types. It applies depthwise separable convolutions and expansion layers to improve both weight efficiency and accuracy. Super class detection demonstrates better performance than conventional YOLOv5 and YOLOv7, while the second stage achieves higher accuracy with fewer parameters compared to traditional CNN models. As a result, the YOLO-based detection performance is maximized, and the method maintains real-time operation with an average detection speed of 9.9 ms. However, performance still degrades in low-light conditions, and validation remains limited in real-world autonomous driving environments. In this study, a new dataset tailored to the experimental environment is constructed and evaluated in simulated autonomous driving environments.

2.5. Contribution in Integrated System

This study integrates different technological component (autonomous vehicle control, urban environments perception, and V2X network), distinguishing itself from prior research limited to single domains. By comprehensively addressing the feasibility of cooperative autonomous driving systems within the real world, the study provides meaningful contributions to the field. Vehicle control is implemented using both TSB driving and waypoints-based driving. The proposed TSB driving approach allows the latter vehicle to utilize the throttle, steering, and brake values of the preceding vehicle, thereby reducing the computational overhead and overcoming hardware limitations to satisfy the real-time requirements.

Unlike previous studies focused on highway environments, this method maintains real-time performance and safety even under increased computational overhead due to traffic light and pedestrian perception, as well as data transmission via the V2X network

in urban settings. This highlights the advantage of the proposed control technique in ensuring lightweight and scalable cooperative autonomous driving. In terms of urban environments perception, this study develops a stop algorithm by recognizing traffic lights and pedestrians. For traffic light perception, the HSV-based perception method is adopted and compared against a CNN-based YOLO model. By analyzing both approaches in terms of accuracy and computational efficiency, the study selects the method better suited to the urban environments, contrasting with prior research that relies on a single perception model.

Furthermore, vehicles in the cluster are defined as clients, and a V2X network is designed to enable the real-time exchange of location data. This allows vehicles to dynamically escape or join the cluster, ensuring stable cooperative autonomous driving even in complex urban environments. Notably, this study proposes an algorithm that enables vehicles to flexibly escape or join the cluster depending on traffic conditions, particularly the stop situations frequently encountered in urban environments, which is an aspect rarely addressed in earlier studies.

3. Methodology and Control Model Analysis

This paper employs the TSB driving method through a V2X network to realize cooperative autonomous driving in an urban environments. For vehicle control, the Pure Pursuit algorithm and PID control technique are applied. In the urban environments setting, traffic lights are detected using the Open Source Computer Vision Library (OpenCV, 4.9.0) with HSV-based color filtering, and pedestrians are recognized through communication with the CARLA server. A central communication server is established within the V2X network, which interacts with the CARLA server. The CARLA server communicates with vehicles via socket programming. Additionally, a cluster escape and joining algorithm is designed, enabling vehicles to dynamically change cooperative driving states in real time. There are five reference terms used for vehicle roles in this study: leading vehicle, following vehicle, preceding vehicle, latter vehicle, and tail vehicle. The leading vehicle is positioned at the front of the cluster, while the tail vehicle is located at the rear. The preceding vehicle refers to the one ahead during an information exchange, transmitting data to the vehicle behind. The latter vehicle is the one behind, receiving data from the preceding vehicle. Vehicles that follow the preceding vehicle and rely on its data are referred to as following vehicles, which include all vehicles in the cluster except for the leading vehicle. The overall architecture of this study is illustrated in Figure 1, and each method is described in detail in the subsequent sections. The red arrows indicate the TSB driving method, the blue arrows represent the waypoints-based driving method, and the black arrows denote the process applied to all vehicles. The colors of the boxes are used solely for visual distinction.

3.1. TSB Driving and Wapoints Driving

Waypoints-based driving employs the Pure Pursuit algorithm to control the steering angle and uses PID control to regulate the vehicle's speed [22]. This control strategy is applied to the leading vehicle and to following vehicles when entering curved sections.

$$R = \frac{L}{\tan(\delta)}, \quad \delta = \arctan\left(\frac{2L\sin(\alpha)}{l_d}\right) \tag{1}$$

The look ahead path following with the target point and actual vehicle path is illustrated in Figure 2, and the Pure Pursuit algorithm, as expressed in Equation (1), is applied to compute the steering angle [23]. The orange line indicates the look ahead distance between the vehicle and the target point, while the blue line represents the actual driving path of the vehicle. Dashed lines represent reference trajectories based on waypoints, whereas solid lines indicate the actual path traveled by the vehicle. The steering angle is calculated based

on the distance to the target point, the wheelbase, and the angle formed with respect to the center of the rear axle. As the vehicle moves, the target point continuously updates, and the steering angle is adjusted accordingly. In this study, the distance between waypoints is set to 1 m, and to ensure stable trajectory tracking, the next waypoint is selected when the distance between the vehicle and the waypoint is less than 2 m. Therefore, the look ahead distance l_d is set to 2 m. The Tesla Model 3, used in the CARLA simulator, has a wheelbase L = 2.85 m. R represents the turning radius of the vehicle, which is the radius of the circular path that the vehicle follows based on its current steering angle. The angle α is computed by subtracting the current vehicle yaw from the angle between the vehicle's current position and the coordinates of the target point. Finally, the derived steering angle δ is converted into radians and transmitted as the control command *cmd_steer*.

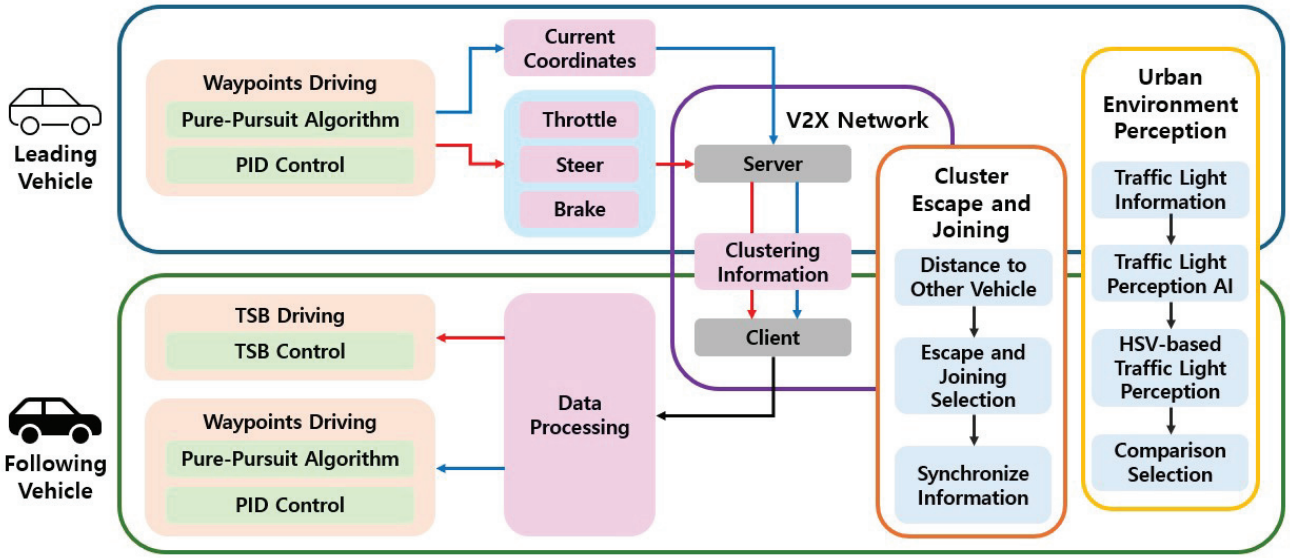


Figure 1. Structure of cooperative autonomous driving using TSB driving and waypoints driving.

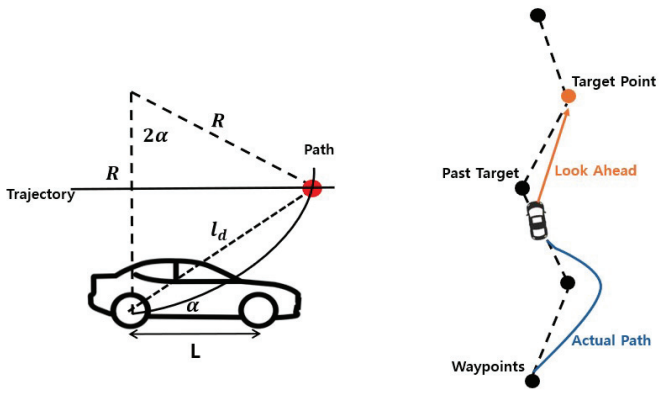


Figure 2. Look ahead path following with target point and actual vehicle path.

$$u(t) = K_p e(t) + K_i \int_0^t e(t)dt + K_d \frac{de(t)}{dt} \tag{2}$$

For speed control, the PID control technique is employed, and a formula such as Equation (2) is applied [24]. Here, $u(t)$ represents the calculated output and serves as the vehicle's final throttle value. The error term $e(t)$ is defined as the difference between the target throttle value of 0.3, which corresponds to a target speed of 30 km/h, and the current

throttle value. This control process utilizes the proportional, integral, and differential constants K_p, K_i, and K_d to avoid sudden acceleration or deceleration. In the CARLA simulator, the throttle value ranges from 0.0 to 1.0, with 1.0 representing full throttle. Based on experimental tuning, the constants are set as K_p = 1.0, K_i = 0.05, and K_d = 0.01. The calculated throttle value is clipped between 0.0 and 1.0 using the np.clip() function and transmitted as the command value.

TSB driving is a method in which the latter vehicle receives and applies the throttle, steer, and brake values generated by the preceding vehicle during driving. While conventional waypoints-based driving computes the control commands based on the target and current x, y coordinates, TSB driving directly utilizes the TSB values produced by the preceding vehicle during straight line motion as control inputs for the latter vehicle. This approach allows following vehicles to perform optimized driving without the need for separate calculations.

When the preceding vehicle uses a threshold-based method to distinguish between straight and curved driving, the accuracy degrades in S-shaped or complex curved road segments. To address this limitation, the preceding vehicle employs an AI approach based on the yaw_diff value to classify driving states. The yaw_diff represents the rate of change in the vehicle's heading direction and is modeled as time-series data using a Recurrent Neural Network (RNN). Based on this classification, the control system switches to transmitting TSB values to the following vehicles during straight line driving, while applying waypoints-based driving in curved segments. Comparative analysis confirms that this hybrid approach enables stable driving while significantly reducing computational overhead compared to conventional waypoints-based methods.

3.2. Communication and Security Between Vehicles in Cooperative Autonomous Driving

In cooperative autonomous driving system, the accurate management of cluster information is essential due to continuous inter-vehicle communication and the handling of complex requests [25,26]. In this research, the CARLA simulator is considered an edge server operating within a single wireless network cell, reflecting the coverage area of a typical 5G base station. This assumption does not imply the need to deploy a server in every cell but rather serves to model communication delays and cluster control. V2V communication is based on the direct communication method between vehicles and other devices (PC5) interface, while V2I communication conceptually corresponds to the network-based communication method between vehicles and the infrastructure via the cellular base station (Uu) interface. Without implementing the physical layer, the communication is abstractly modeled, and message transmission is secured through Transmission Control Protocol (TCP)-based socket communication. Since the focus of this study lies in evaluating the effectiveness of control strategies rather than precise network implementation, factors such as latency and channel effects are not simulated. The edge server manages the cluster by validating vehicle data and enhances security through pseudonymization, encryption, and digital signatures [27]. The simulation environment is configured under ideal conditions to compare the performance of TSB-based driving and waypoints-based driving. The proposed network architecture is illustrated in Figure 3. The red and blue groups indicate separate vehicle clusters.

The network structure primarily consists of 5G-NR communication among the V2X server, the CARLA server, and the preceding and latter vehicles. The V2X server stores and utilizes information received from the CARLA server to support efficient cluster management and data processing. To verify the digital signature of a pseudonymized vehicle, the system is linked to a database that maps pseudonymized IDs to actual vehicle IDs. Upon receiving an authentication request from the CARLA server, the public key of the

real ID that is mapped from the pseudonym is retrieved and returned for verification. The CARLA server functions as an edge server and provides the V2X server with various data, including actor information, map data, and traffic flow information such as pedestrian locations. It also supports the creation and management of RSUs. Each vehicle transmits pseudonymized information such as its vehicle ID and current location to the CARLA server. The server records these data to determine whether clustering should be performed. Additionally, the preceding vehicle uses the PC5 interface to communicate with the latter vehicle and determine its control behavior [28].

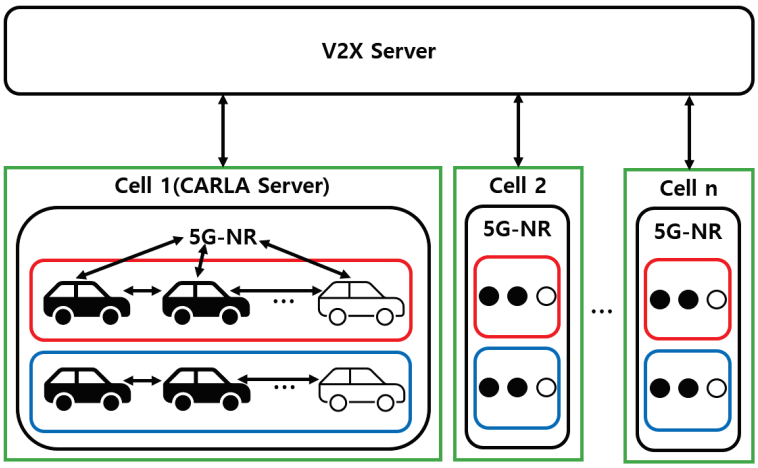


Figure 3. Network topology for the proposed cooperative autonomous driving system.

Table 2 presents the message format used for inter-vehicle communication, which is designed to securely transmit vehicle identification and driving information. The vehicle ID, front ID, and rear ID are pseudonymized for vehicle distinction and are managed per session. A session ID is assigned to maintain and track the communication link between each vehicle and the CARLA server.

Table 2. Message packet details for the V2X network.

Field Name	Size	Description
Vehicle ID	2 Bytes	Unique identifier for the vehicle (pseudonymized ID))
Front ID	2 Bytes	Identifier of the leading vehicle (pseudonymized ID)
Rear ID	2 Bytes	Identifier of the following vehicle (pseudonymized ID)
Session ID	4 Bytes	Identifier for tracking the connection state during a communication session
Message Type	1 Byte	Type of V2X message (e.g., maneuver coordination, Cluster Joining, Cluster Escape)
Status	1 Byte	Status of the message (e.g., pending, accepted, rejected)
Timestamp	4 Bytes	Transmission timestamp (Unix Timestamp)
Encrypted Data	Variable	Encrypted vehicular mobility data (e.g., Coordinates, Speed, Steering angle)
Checksum	2 Bytes	Error detection checksum for message integrity
Hash (SHA224)	28 Bytes	Cryptographic hash value of the message (SHA-224)
Signature	64 Bytes	Security signature (EdDSA, optimized for low overhead compared to RSA)

When the message type is defined as driving data exchange, the Status field is used to distinguish between straight and curved sections. In the case of escape or joining messages,

this field indicates the vehicle's state, such as standby, acceptance, or rejection. Encrypted data include driving information such as the vehicle's location, speed, and steering angle, and are protected using symmetric key encryption methods such as AES. For escape and joining messages, this data field is padded with zeros. The Checksum field is responsible for detecting message errors and verifying data integrity through the Secure Hash Algorithm-224 (SHA-224). The Signature field ensures the authenticity of the message source, incorporating an EdDSA-based digital signature. This EdDSA-based signature is used only during the "Init" stage to authenticate the pseudonymized ID.

The vehicle information exchange and cluster formation process defines and implements initialization, cluster escape, and joining procedures according to the flow illustrated in Figure 4.

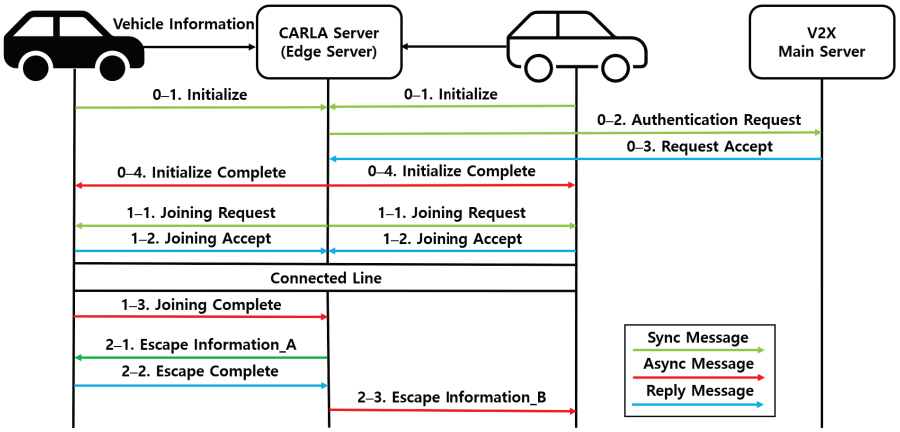


Figure 4. Flow for vehicle information exchange and cluster management.

3.2.1. Init

When a vehicle enters a new region managed by a different CARLA server and first connects to the corresponding edge server, its vehicle ID and initial location information are encrypted and transmitted. At this stage, the pseudonymized vehicle ID is inserted into the vehicle ID, front ID, and rear ID fields. The message type is set to "Init", and the driving information field is padded with zeros. The CARLA server updates and manages vehicle information based on the received packet. Simultaneously, the V2X server obtains the public key corresponding to the actual vehicle ID that is mapped from the pseudonymized ID. Using this key, the vehicle is authenticated through a digital signature. Once authentication is completed, the CARLA server sends an initialization completion message to the vehicle along with waypoints information. The vehicle then begins autonomous driving based on the assigned waypoints and its destination.

3.2.2. Cluster Joining

When a vehicle approaches within 15 m of another, the CARLA server identifies the preceding and latter vehicles and generates a "Cluster Joining Request" message, which initiates the cluster formation procedure. Upon receiving this request, each vehicle updates its joining status and responds to the CARLA server with a message indicating standby, acceptance, or rejection. If a vehicle is successfully joined to the cluster, its own ID is recorded in the vehicle ID field. The ID of the preceding vehicle is recorded in the front ID field, while that of the latter vehicle is recorded in the rear ID field. During this stage, the encrypted driving information field is padded with zeros, and the message type is set to "Cluster Joining". Subsequently, the preceding and latter vehicles are connected via the PC5 interface. The preceding vehicle determines the necessary driving information and transmits it to the latter vehicle. Finally, when the latter vehicle sends a confirmation

packet indicating successful joining, the CARLA server updates the vehicle information accordingly.

3.2.3. Cluster Escape

When a vehicle moves more than 30 m away, the server sends a "Cluster Escape" message to the latter vehicle. The ID of the corresponding vehicle is recorded in both the vehicle ID and front ID fields. If a latter vehicle exists, its ID is recorded in the rear ID field, and if not, the ID of the corresponding vehicle is used instead. The encrypted driving information field is padded with zeros, and the message type is set to "Cluster Escape". The vehicle then updates its status and responds to the server. The server subsequently updates the vehicle information and processes the escape request.

3.3. Urban Environment Perception

Unlike highways, urban environments require the careful consideration of traffic lights and pedestrians, particularly at intersections. At a red traffic light, the vehicle must come to a complete stop. When the leading vehicle stops, multiple following vehicles may join the cluster. Conversely, if only the following vehicles stop, they may escape the cluster. As the proposed cooperative autonomous driving system operates in urban environments, traffic light and pedestrian perception are regarded as critical factors. Pedestrians transmit their real-time location information to the edge server, which identifies the pedestrian closest to the vehicle. The server then determines a safe distance based on the road width. If a pedestrian is detected within this distance, the vehicle stops immediately [29].

3.3.1. Traffic Light Perception

The traffic light perception process is illustrated in Figure 5. This study compares an AI-based method, which utilizes a trained traffic light detection model, with an HSV value-based method and selects the more suitable approach for the urban environments.

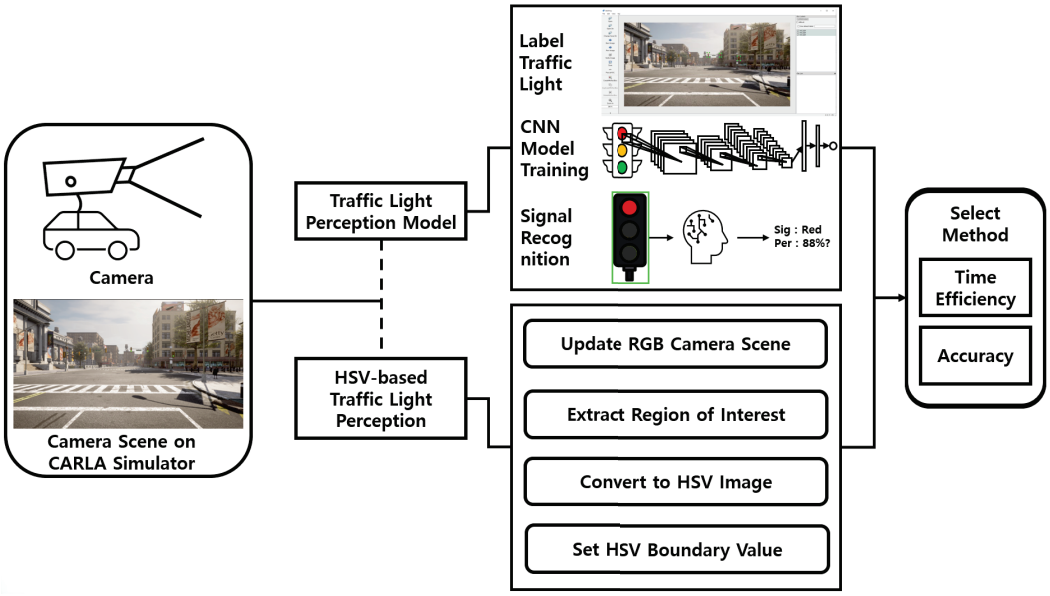


Figure 5. Structure of traffic light perception in urban environment.

The AI-based perception method uses a Red–Green–Blue (RGB) camera in the CARLA simulator to capture traffic light scenes at intersections. Red, yellow, and green signals are collected at equal frequencies. The entire traffic light is labeled to preprocess the data for training a CNN model. Based on the training results, a traffic light perception model is generated. This study employs YOLOv12nano, the lightest version of YOLOv12, which

is the latest iteration in the YOLO series. YOLO is an object detection model originally proposed by Joseph Redmon in 2015. While earlier object detection models require multiple passes over an image and require substantial computational resources, YOLO improves efficiency by detecting objects in a single pass. Because traffic light detection must operate continuously during driving and the vehicle's embedded system must manage multiple simultaneous tasks, the computational resources allocated for perception are limited. Therefore, YOLOv12nano is selected for its higher accuracy and faster performance compared to YOLOv11, while requiring fewer computational resources [30].

The HSV value-based traffic light perception method is implemented using OpenCV. Traffic light locations at various intersections are defined as Regions of Interests (ROIs) and extracted, enabling focused analysis on the traffic light area. The input image is converted into the HSV color space, and color-based filtering is applied by setting HSV thresholds corresponding to traffic light HSV colors. If the number of pixels within the threshold in the ROI is eight or more, the traffic light is considered successfully detected. As HSV values may vary due to weather or environmental conditions, repeated experiments are conducted to validate robustness. The results confirm that setting the pixel count threshold to eight yields accurate and reliable perception.

The two methods are compared in terms of time efficiency and accuracy. As this study is conducted in an urban environment, real-time performance is prioritized. Accordingly, if both methods satisfy a predefined accuracy threshold, the method with superior time efficiency is selected. The accuracy threshold is experimentally determined to ensure safe and reliable traffic light perception.

3.3.2. Pedestrian Perception

Pedestrian detection is based on V2P communication within the V2X network [31]. If a pedestrian's location received via the V2X network is within the vehicle's predefined safety distance, an emergency response is triggered and the vehicle comes to a stop. The Tesla Model 3, used in this research, has a total length of 4.72 m and a width of 2.09 m. The distance from the vehicle center is approximately 2.36 m in the longitudinal direction and 1.05 m laterally. Pedestrian positions are continuously received through the V2X network. If a pedestrian is detected within 6.5 m of the client vehicle, emergency braking is executed.

In this study, the average driving speed is set to 30 km/h, requiring a minimum braking distance of 1.5 m. The predefined braking distance of 6.5 m satisfies this requirement [32]. On actual roads, assuming an average urban speed of 50 km/h in South Korea, the required braking distance is approximately 20 m. Therefore, applying this system to real vehicles would require communication constraints to limit signal exchange within the roadway. If the vehicle speed exceeds 40 km/h, the braking distance threshold is adjusted to 20 m within the algorithm.

3.4. Cluster Escape and Joining

In urban environments, the cluster escape and joining algorithm plays a crucial role [33,34]. Unlike highways, where vehicles maintain continuous movement, vehicles in urban environments frequently stop. Therefore, safe cooperative autonomous driving is feasible only when this algorithm is properly implemented. In this study, the algorithm is designed using each vehicle's ID and its distance from the preceding vehicle within the V2X network.

The pseudo-code of the cluster escape and joining algorithm is presented in Algorithm 1. The input includes the vehicle ID and its distance from the preceding vehicle. The output provides the cluster state, represented by $data[\text{“}Message\ Type\text{”}]$, as well as the IDs of vehicles that have escaped or joined the cluster.

Algorithm 1 Algorithm of the cluster escape and joining.

Input: Distance of preceding vehicle (Distance), ID of preceding vehicle ($self.front$), ID of player ($self.player.id$)
Output: Situation of cluster ($data[\text{“}Message\ Type\text{”}]$), ID of escaped and joined vehicle

if $self.front$!= $self.player.id$ **then**
    if Distance > 30 **then**
        $data[\text{“}Message\ Type\text{”}]$ = “escape”
    else if 30 >= Distance > 15 **then**
        throttle = 0.25
    else if 15 >= Distance > 8 **then**
        throttle = -0.25
    else if Distance <= 8 **then**
        brake = 1

if $self.front == self.player.id$ **then**
    if distance to rear vehicle of another cluster < 10 **then**
        $data[\text{“}Message\ Type\text{”}]$ = “join”

return Complete Cluster Escape and Joining – using $data[\text{“}Message\ Type\text{”}]$, ID of escaped and joined vehicle

If the ID of the leading vehicle received through the V2X network does not match the ID of the current vehicle, it is classified as a following vehicle. The cluster escape and joining algorithm operates based on the distance to the preceding vehicle. In cooperative autonomous driving, the maximum allowable distance between vehicles in a cluster is set to 30 m. If the distance exceeds this threshold, the vehicle is considered to have escaped from the cluster. In such a case, the system sets $data[\text{“}Message\ Type\text{”}]$ to “escape” and designates the vehicle as a new leading vehicle, transitioning it to waypoints-based driving mode. The optimal inter-vehicle distance within a cluster is set to 15 m. If the distance to the preceding vehicle exceeds 15 m but remains below 30 m, the throttle value is increased to close the gap to 15 m. Conversely, if the distance is between 8 m and 15 m, the throttle is reduced to maintain the target spacing. If the preceding vehicle comes to a stop, adjusting the throttle alone may result in a collision. Therefore, if the distance to the preceding vehicle falls below 8 m, the brake is applied to bring the vehicle to a complete stop. As the following vehicles also stop sequentially, this mechanism prevents collisions even when the current vehicle remains stationary.

A joining event occurs when the ID of the current vehicle matches that of the leading vehicle. If the leading vehicle approaches within 10 m of the tail vehicle in another cluster, a joining is triggered. This event is indicated by setting $data[\text{“}Message\ Type\text{”}]$ to “join”, and the current vehicle is reclassified as the tail vehicle of the target cluster. Since the tail vehicle functions as a following vehicle within the group, the TSB driving mode may be applied depending on the situation.

4. Results

In this research, cooperative autonomous driving in an urban environments is implemented using the CARLA simulator. CARLA simulator provides open-source digital assets for autonomous driving system development and urban layout simulation, making it suitable for this research. Moreover, the CARLA simulator offers an effective experimental environments by supporting environmental perception through various sensors, such as cameras and Light Detection and Ranging (LiDAR), enabling vehicles to perform autonomous driving accordingly.

This study is conducted based on TOWN 10 in the CARLA simulator, which features a grid layout, including numerous different junctions including a 4-way yellow-box junction, dedicated turning lanes and central reservations. Traffic lights are installed at all junctions where two or more roads meet, and the driving route is configured to pass through a total of four traffic lights. When the number of pedestrians is set to 10, interactions between vehicles and pedestrians at intersections rarely occur. Conversely, when set to 50, pedestrians continuously cross intersections, causing vehicles to remain stopped. Therefore, the number of pedestrians is set to 25 to allow the vehicle to detect pedestrians at intersections, perform a stop, and resume driving. The entire route, including the departure point, arrival point, and traffic light sections, is illustrated in Figure 6.

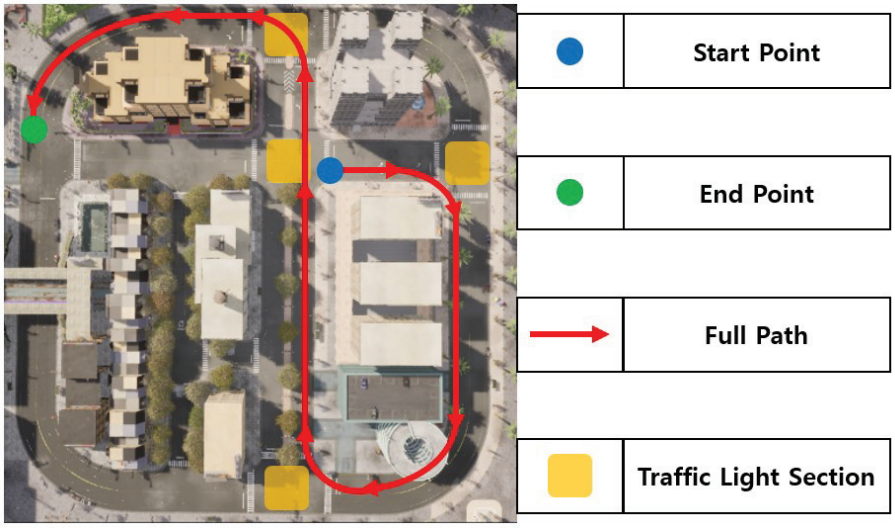


Figure 6. Full path and location information in urban environment.

To evaluate the performance of the proposed cooperative autonomous driving system, experiments are conducted using the CARLA simulator version 0.9.12. The experimental environment is configured on a high-performance system equipped with an NVIDIA RTX 3080 Graphics Processing Unit (GPU, emTek, Jakarta, Indonesia) and 64 gigabytes (GB) of RAM.

4.1. TSB Driving and Wapoints Driving

Autonomous driving in this study utilizes both TSB driving and waypoints-based driving. In the TSB driving mode, the following vehicle measures the total travel distance using the Euclidean distance method and divides it by the total execution time to calculate the execution time per meter [35]. The corresponding results are summarized in Table 3.

Table 3. Performance indicators for TSB driving and waypoints-based driving.

	ALL Meter (m)	ALL Time	Meter per Time
TSB Driving	189.0302	0.1114	0.0005
Waypoints-based Driving	166.5100	38.8167	0.2331

The total travel distance for TSB driving is approximately 189 m, whereas waypoints-based driving covers around 166 m, resulting in a difference of about 23 m. This discrepancy arises from variations in route selection and is considered acceptable for comparative analysis. Despite the longer travel distance, the total execution time for TSB driving is measured at approximately 0.1114 s, which is about 388 times shorter than that of waypoints-based driving, 38.8167 s. Consequently, the execution time per meter for TSB driving is approxi-

mately 466 times lower, indicating that TSB driving is significantly more efficient in terms of computational performance and practical applicability.

4.2. Communication and Security Between Vehicles in Cooperative Autonomous Driving

Efficient communication and security between vehicles are critical for optimizing clustering and ensuring safety in cooperative autonomous driving systems [36]. This research designs a system that enables clustering through real-time V2X communication, allowing vehicles to cooperate for safe and efficient operation. The key considerations for communication and security in this process are outlined as follows.

4.2.1. Assessment of Interoperability

The cooperative autonomous driving system proposed in this study optimizes clustering through real-time information exchange between vehicles and employs V2X communication. Interoperability is a key factor in ensuring seamless data exchange across different vehicle manufacturers and communication technologies. Designed based on 5G-NR V2X, the system supports low latency, high-speed transmission, and high reliability, thereby enabling real-time, delay-free communication among heterogeneous vehicles. In particular, the PC5 interface facilitates direct communication between vehicles, ensuring efficient data exchange regardless of the vehicle brands or protocol differences. The CARLA simulator and the V2X server operate in conjunction to manage vehicle location, speed, and behavior in real time, enhancing both clustering and driving performance. Additionally, vehicle authentication is achieved by mapping pseudonymized IDs to real IDs, and data integrity is maintained through public key-based authentication. Throughout this process, authenticated data exchange ensures secure and interoperable communication between vehicles and infrastructures.

4.2.2. Assessment of Personal Information Protection

Privacy protection is a critical aspect of real-time V2X communication. In this research, various security mechanisms are designed to ensure data confidentiality and identity protection. Specifically, the use of pseudonymized vehicle IDs, AES encryption, and EdDSA digital signatures are designed as core security measures to achieve this objective.

The pseudonymized vehicle ID prevents the external exposure of vehicle information, thereby protecting the driver's privacy. The vehicle ID, front ID, and rear ID are all pseudonymized and managed in a way that prevents association to the vehicle's actual identity. These IDs are maintained on a session basis and securely mapped to the real ID through interaction with the V2X server. This approach ensures that the vehicle's real-time location and operational data cannot be tracked, playing a vital role in minimizing privacy breaches. Moreover, the pseudonymized vehicle ID is updated either per session or periodically when a vehicle connects to a cluster network, preventing the continuous tracking of the vehicle's location or behavior. Such periodic updates are essential for minimizing vehicle traceability and ensuring both privacy protection and anonymity.

Critical vehicle communication data, such as location, speed, and steering angle, can be encrypted and transmitted using AES, thereby preventing external attackers from eavesdropping or tampering with the messages. EdDSA-based digital signatures enable the verification of message authenticity, ensuring that transmitted data originates from authenticated vehicles. This mechanism prevents man-in-the-middle attacks and guarantees both message origin and integrity. EdDSA is particularly effective in V2X network, as it provides faster processing and smaller signature sizes compared to Rivest–Shamir–Adleman (RSA), enhancing both performance and security. However, since signature generation still incurs latency, EdDSA is applied only during the "Init" stage to authenticate pseudonymized vehicle IDs. During subsequent V2V communication, digital signatures are omitted to

allow fast and efficient packet exchange. This selective use of signatures supports real-time clustering and minimizes potential privacy breaches during continuous V2X communications. Standards such as IEEE 1609.2 define digital signature and encryption methods to secure V2X communications [37]. In this study, EdDSA and AES are adopted to implement these cryptographic mechanisms, along with additional security features such as pseudonymized ID management and session-based identity handling. Consequently, the proposed method inherits the strengths of existing standards while enhancing real-time performance and communication efficiency, making a significant contribution to the secure design of cooperative autonomous driving systems.

4.2.3. Countermeasures Against Attack Simulations

V2V communication systems are vulnerable to various attacks, including man-in-the-middle attacks, replay attacks, and denial-of-service (DoS) attacks. The security mechanisms implemented in this study demonstrate strong resistance against such simulated threats.

First, EdDSA digital signatures and AES encryption are employed to ensure data integrity and authenticity. These techniques prevent attackers from intercepting or tampering with transmitted messages. Since each message is cryptographically signed, the system can verify whether it originated from an authenticated vehicle, effectively mitigating man-in-the-middle attacks. Additionally, timestamps and checksums are embedded in each message to distinguish between newly transmitted data and previously sent packets, thereby protecting against replay attacks. Furthermore, pseudonymized vehicle IDs and location-based encryption are utilized to prevent external tracking of vehicle positions. This approach safeguards driver privacy and significantly reduces the risk of traceable data leakage.

4.3. Urban Environment Perception

4.3.1. Traffic Light Perception

Traffic light perception in the urban environments is performed using two methods. In Figures 7 and 8, and Table 4, the HSV value-based perception method is referred to as HSV perception, while the AI-based method, which employs a trained traffic light perception model, is referred to as AI perception. The AI-based model is trained on a balanced dataset of 225 images, with an equal distribution of red, yellow, and green traffic light states. Since the purpose is not to optimize model performance but to establish a baseline for comparison with the HSV-based method, data augmentation is not applied. This design choice is necessary to evaluate the two approaches under controlled experimental conditions. The perception results are compared in terms of accuracy and time efficiency. If no traffic light is detected, the system retains the previous control command to maintain safe autonomous driving in the urban environments. Figure 7 shows the results of the HSV-based perception, while Figure 8 presents the AI-based perception results. A quantitative comparison of accuracy and time efficiency is summarized in Table 4.

The AI-based perception method demonstrates relatively lower performance in the experimental evaluation. In contrast, the HSV-based perception method successfully detects all traffic signals after 800 iterations, achieves an accuracy of 99.5833%, and requires only 3.851 ms to recognize a single traffic light. As discussed in Section 3.3.1, real-time performance is prioritized due to the urban driving environments. Given that the HSV-based method is approximately 101.8 times faster, it is adopted for urban environments perception. The high perception accuracy of 99.5833% confirms its reliability. Since it is efficient in both time and accuracy, no explicit accuracy threshold is defined. In most scenarios, traffic lights are absent. These cases are classified as no_signal. When the signal is categorized as no_signal, the vehicle retains its previous control state, allowing it to stop at

red lights and proceed through green lights. Through repeated experiments, it is confirmed that HSV-based perception enables reliable urban environments perception and supports safe autonomous driving.

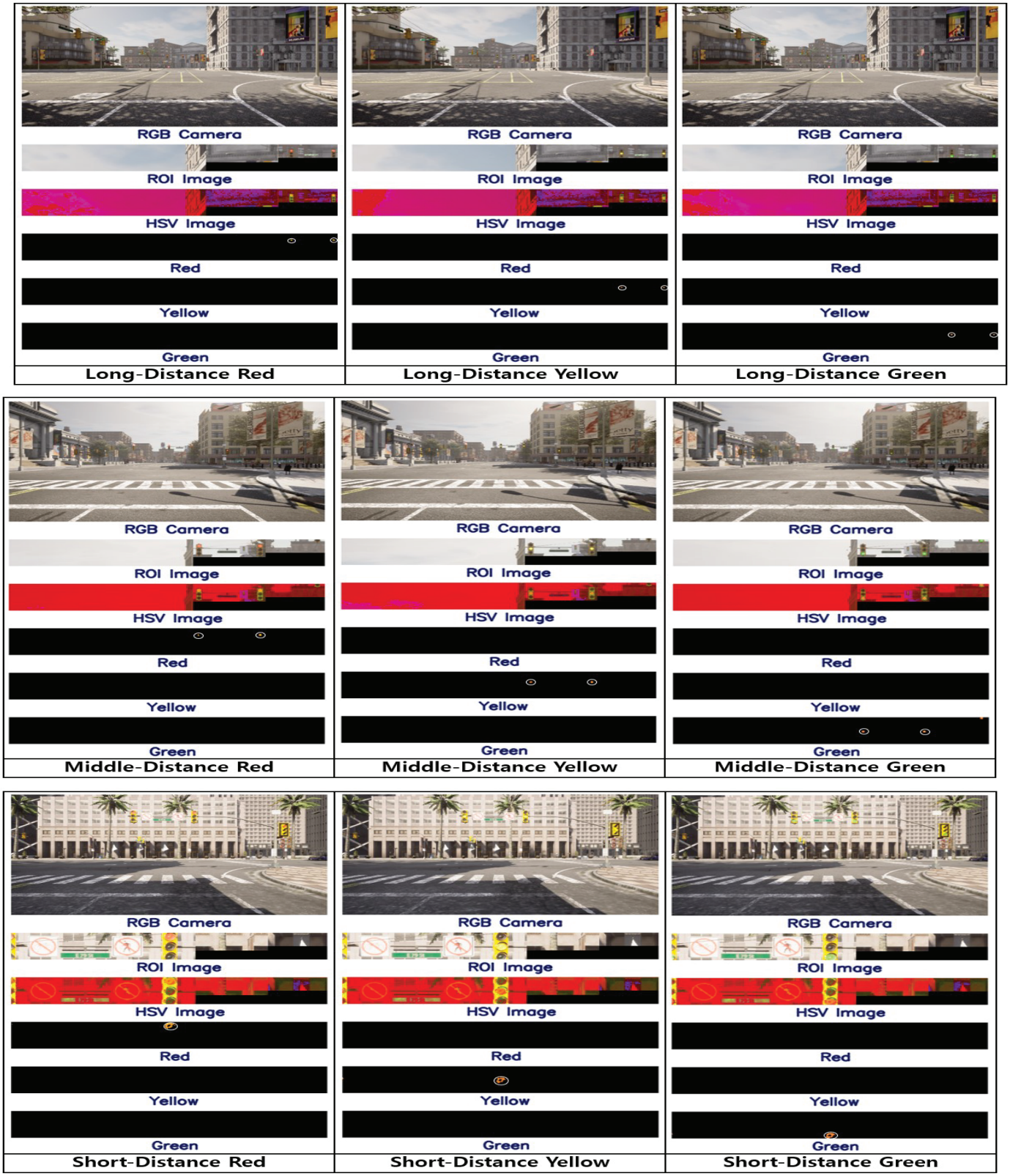


Figure 7. Result and process images of HSV perception's traffic light perception experiment.

Figure 8. Result images of AI perception's traffic light perception experiment.

Table 4. Performance indicators for AI perception and HSV perception.

	Perception Accuracy (%)	Time Efficiency (ms)
AI Perception	70.7727	392.074
HSV Perception	99.5833	3.851

4.3.2. Pedestrian Perception

The presence of pedestrians within the maximum distance from the center of the vehicle to the curb is continuously monitored via the CARLA server. As illustrated in Figure 9, the vehicle initiates braking to prevent collisions when a pedestrian is detected. The blue circle indicates the pedestrian detection range of the vehicle, while the yellow circle visualizes the pedestrian's location. Given the average vehicle speed of 30 km/h, the vehicle stops if a pedestrian is located within 6.5 m of its center. Once the pedestrian moves beyond the safe threshold, the vehicle resumes driving. If the vehicle speed exceeds 40 km/h, the threshold distance is adjusted to 20 m. This behavior is essential for ensuring safety in urban environments. For static obstacles, the vehicle must generate an avoidance path. However, due to the dynamic nature of pedestrian movement, a full stop is required to guarantee safety. With future advancements in V2X networks, it is anticipated that

vehicles will not only stop for pedestrians but also dynamically generate alternative routes through areas with lower pedestrian density, thereby improving traffic flow and driver convenience.

Figure 9. Pedestrian perception in urban environment.

4.4. Cluster Escape and Joining

In Figure 10, the vehicle states during a cluster escape and joining scenario, along with the server's communication status, are illustrated. The red square indicates the initial cluster, while the yellow square shows a new cluster formed by a vehicle that escaped from the original group. The server facilitates sequential communication between the preceding and latter vehicles. The vehicles involved in this scenario have the vehicle IDs "49", "52", and "59". Cluster escape and joining typically occur when a vehicle stops at an intersection. In this case, the tail vehicle remains at the intersection and subsequently escapes the cluster. The server notifies that client 3, with vehicle ID "59", has escaped. The vehicle with ID "59" receives this notification, records the information, and responds to the server. The server then assigns vehicle ID "59" as the leading vehicle in a newly formed cluster and resumes waypoints-based driving. Additionally, the server updates client 2, with vehicle ID "52", that its role has changed to the new tail vehicle. Client 2 acknowledges this change and records the update accordingly.

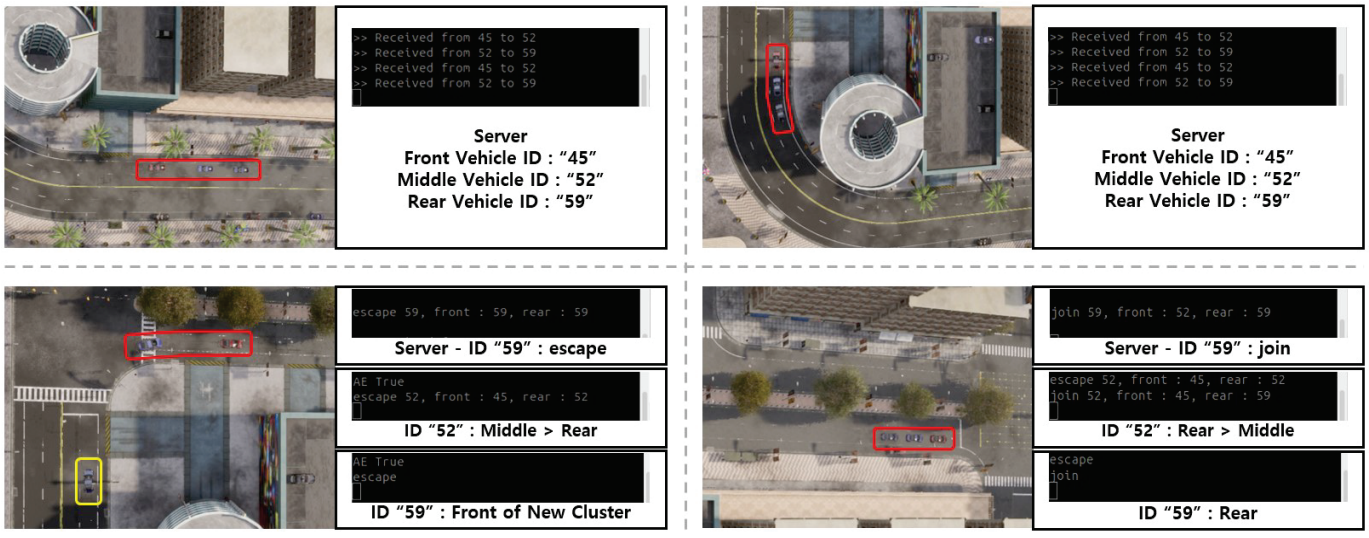


Figure 10. Experimental results of cluster escape and joining algorithm.

4.5. Integrated System of Cooperative Autonomous Driving

To implement the cooperative autonomous driving system, the experiment is conducted in an integrated environment composed of three independent components: the CARLA server, the V2X server, and the vehicles. Each component is executed in a separate Ubuntu terminal to ensure modularity and process isolation.

4.5.1. CARLA Server

When the CARLA simulator is launched, the CARLA server is initialized to manage various elements such as actors, traffic, and simulation time through the World class. The server continuously monitors the real-time locations of all actors, assuming the use of Global Positioning System (GPS) sensors, and utilizes this information to perform tasks such as accident prevention. To ensure secure communication, the CARLA server exchanges data with the V2X server, receiving information such as vehicle ID, front ID, and rear ID from each vehicle. Based on this information, the server analyzes inter-vehicle relationships and, if the predefined conditions are met, manages them as a single cluster. Upon receiving vehicle data, the server first checks the message type and then executes the corresponding operation. In addition, the cluster and cell information maintained by the V2X server is periodically transmitted to the CARLA server, where it is stored and used for synchronized management.

4.5.2. Vehicles

Vehicles are registered as vehicle objects among the Actor instances in the World class of the CARLA server. To reflect the real-world driving conditions, each vehicle undergoes an initialization process via the "INIT" message, and the experiment is conducted under the assumption that all vehicles enter a defined cell. Communication between the vehicle and the server is established through the Uu interface, while the PC5 interface is assumed for V2V communication. After completing initialization, each vehicle receives waypoints from the server based on its assigned destination and proceeds accordingly. Safety, particularly accident prevention during driving, is treated as the highest priority. When the CARLA server detects a potential collision with a pedestrian or another vehicle, it issues an emergency signal, and the vehicle immediately sets the brake value to 1 to come to a stop. Similarly, when a red signal is detected during traffic light perception, a control_red signal is generated, and the brake is also set to 1. Each vehicle transmits its x and y coordinates to the server in real time. If the vehicle belongs to a cluster, this positional information is also shared with both the preceding and latter vehicles. A vehicle designated as a latter vehicle receives the throttle, steer, brake or coordinates information from the preceding vehicle and adjusts its movement accordingly. It then transmits its own x and y coordinates to the latter vehicle via V2V communication, depending on the status value contained in the most recently received message.

The vehicle control process is structured into three stages as illustrated in Algorithm 2. First, to prioritize pedestrian safety, the system determines whether the vehicle should stop based on pedestrian detection. Next, traffic light information is used to control vehicle behavior in accordance with traffic signals. These two steps help prevent potential collisions in advance. In the final stage, the vehicle identifies its position through communication with the server and selects either TSB-based or waypoints-based driving, depending on the presence of a preceding vehicle and the assigned driving mode. The proposed method contributes to both safety and driving efficiency in cooperative autonomous driving within urban environments.

Algorithm 2 Integrated algorithm for perception and autonomous driving control.

Input: Camera image, CARLA Server message (Cluster_information), ID of preceding vehicle ($self.front$), ID of player ($self.player.id$)
Output: Driving command (`throttle`, `steer`, `brake`)

▷ **Step 1: Pedestrian Detection**
if pedestrian is detected in camera image **then**
    issue `Stop` command
    return

▷ **Step 2: Traffic Light Perception**
($self.world.camera_manager.state$) ← recognize traffic light (via HSV-based method)
if `light_state` == "RED" **or** `light_state` == "YELLOW" **then**
    issue `Stop` command
    return

▷ **Step 3: Autonomous Driving Control**
if $self.front == self.player.id$ **then** ▷ Vehicle is the leader
    Perform `waypoint`-based driving control
else ▷ Vehicle is a follower
    Receive driving information from `CARLA` server
    if message mode == ''`waypoints`'' **then**
        Follow received `waypoints` using tracking controller
    else if message mode == ''`TSB`'' **then**
        Apply received `throttle`, `steer`, `brake` values directly

4.5.3. V2X Server

The V2X server is anticipated to be expanded in future work to serve as a central communication server in real-world environments. At the current stage, each CARLA server is assumed to represent a single cell, where all relevant information is collected, recorded, and used for tasks such as ID verification via a connected database. The collected data are stored in JavaScript Object Notation (JSON) format, and specific methods for data storage and processing will be developed in future work.

Figure 11 illustrates the cooperative autonomous driving route between the preceding and latter vehicles, as a result of the integrated system proposed in this study. The preceding vehicle follows a waypoints-based driving method, while the latter vehicle alternates between TSB-based and waypoints-based driving. In the region where the x-coordinate is less than –100, the path of the latter vehicle does not exactly align with that of the preceding vehicle. However, since both trajectories remain within the lane boundaries, the situation is considered safe. A detailed analysis reveals a slight delay in the transition between curved and straight driving modes, indicating that the AI model takes some time to switch from curved to straight section control. This issue is attributed to the boundary conditions set by the AI model and does not compromise the safety of the TSB driving approach. As a result, both vehicles successfully demonstrate safe driving behavior, confirming the practical applicability of the proposed system.

In this research, both the transmission time of the V2X server and the time required for a vehicle to join or leave a cluster are measured to be within 1 ms. The V2X server's transmission time is verified 1781 times, yielding an average of 0.152 ms. The time for a vehicle to complete cluster joining or escape is verified 1015 times, with an average time of 0.23 ms. Although the simulation is conducted under ideal conditions without external forces, the proposed system is expected to maintain its safety, accuracy, and efficiency, even when applied in real-world environments where various external influences exist.

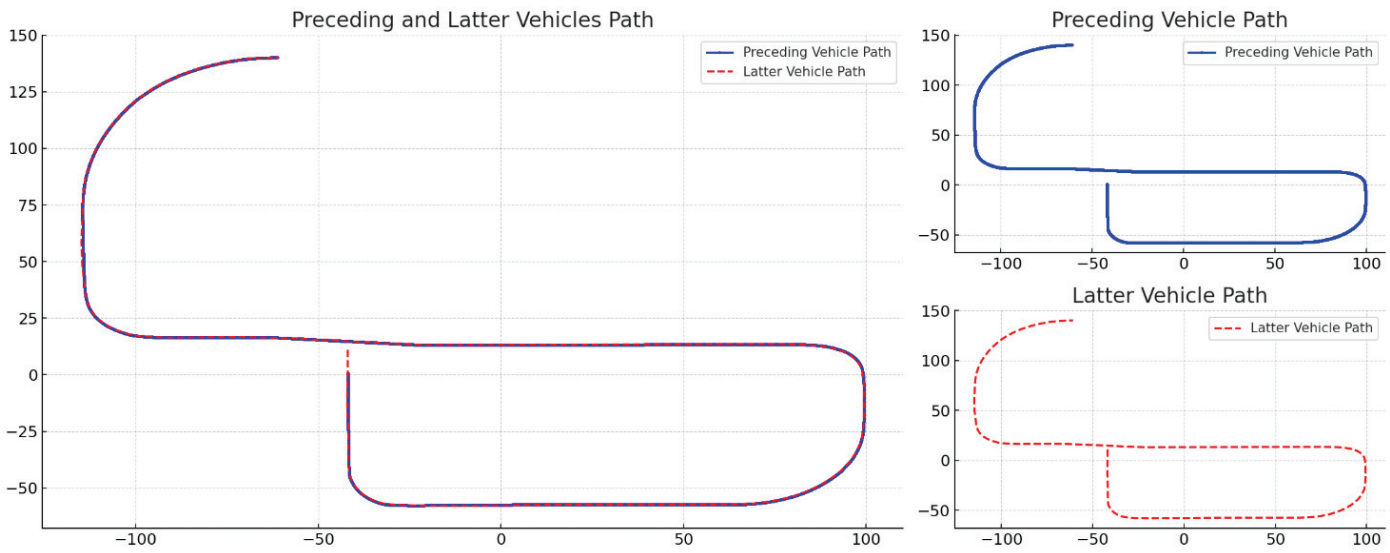


Figure 11. Integrated system result of cooperative autonomous driving.

5. Conclusions and Future Works

In this paper, cooperative autonomous driving is implemented and verified in an urban environments. A new driving technique, called TSB driving, is proposed to ensure real-time performance in V2X communication by minimizing the computational overhead of the following vehicle in cluster scenarios. Autonomous driving, V2X network, urban environments perception, and security components are integrated to enable seamless information exchange. Based on this shared information, vehicles are able to perform cooperative autonomous driving safely within the urban environments.

When cooperative autonomous driving is conducted in urban environments, this study verifies whether safe and efficient driving is achievable by evaluating the computational overhead, speed, and accuracy across each domain. In urban environments, pedestrians, infrastructure, and vehicles are interconnected through a V2X network, enabling real-time communication and coordination across multiple domains. This real-time interoperability is a key factor for realizing cooperative autonomous driving in urban environments. By integrating multiple domains, including autonomous driving, V2X communication, urban environments perception, and security, the proposed system demonstrates improvements in accuracy and time efficiency compared to related studies. Furthermore, a new driving method, TSB driving, is compared with conventional waypoints-based driving to validate its practicality. The adoption of TSB driving contributes to research and development by reducing hardware dependency and enhancing system efficiency.

In future work, the proposed functionality will be expanded by implementing cooperative autonomous driving verified in the CARLA simulator on real roads. For example, by utilizing experimental testbed vehicles, the system can generate traffic information and obtain real-time location data of target vehicles through communication with various RSUs. In addition, autonomous vehicle safety can be further enhanced by diversifying driving routes. Rather than performing complete stops, avoidance routes can be dynamically generated by refining pedestrian detection. The braking distance is adjusted based on the vehicle's speed, and the pedestrian detection range is modified accordingly. Since pedestrians may occasionally be detected outside of the road area, the communication region is planned to be restricted to valid road boundaries using GPS information. As this study is conducted in a simulated environment, no communication errors have occurred between V2X devices. However, future research will incorporate scenarios involving network latency and connection loss to enhance the system's robustness for real-world deployment.

From a technical perspective, various studies are planned to optimize autonomous driving control and enhance safety and real-time performance within the V2X network. Although TSB driving offers advantages in real-time responsiveness, it remains limited in terms of safety. To address this, scenarios will be further subdivided, or corrective parameters will be applied to ensure that TSB-based driving follows a trajectory as closely as possible to that of waypoints-based driving. Furthermore, the experimental environment is planned to be extended to include multiple CARLA server cells, which conceptually represent base stations deployed in different regions of a real-world network. The findings of this study are expected to provide a foundational framework for the development of connected vehicles and smart cities by enabling direct application in real-world scenarios.

In conclusion, this paper introduces TSB driving to optimize the computational overhead on vehicles and proposes a cooperative autonomous driving system that integrates urban environments perception and information exchange via the V2X network. A method for enhancing reliability with consideration for network security is also presented, and its immediate applicability is demonstrated through experiments conducted in the CARLA simulator, which closely replicates urban environments. Future work aims to contribute to the advancement of connected vehicles and smart cities by conducting experiments in real-world environments, deploying multiple RSUs, and further expanding and refining the scope of information exchange.

Author Contributions: Conceptualization, M.Y. and D.S.; methodology, M.Y., D.S. and S.K.; software, M.Y. and D.S.; validation, M.Y., D.S. and S.K.; writing—original draft preparation, M.Y., D.S. and S.K.; writing—review and editing, M.Y. and D.S.; visualization, M.Y. and D.S.; supervision, K.K.; project administration, K.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: https://github.com/ggaebi99/Cooperative-autonomous-driving-in-CARLA-Simulator (accessed on 9 April 2025).

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

V2X	Vehicle-to-Everything
TSB	Throttle–Steer–Brake
AES	Advanced Encryption Standard
EdDSA	Edwards-curve Digital Signature Algorithm
PID	Proportional–Integral–Differential
AI	Artificial Intelligence
HSV	Hue Saturation Value
MEC	Multi-access Edge Computing
V2V	Vehicle-to-Vehicle
V2I	Vehicle-to-Infrastructure
VANETs	Vehicular Ad hoc NETworks
ICN	Information-Centric Networking
IP	Internet Protocol
MPC	Model Predictive Controller
V2P	Vehicle-to-Pedestrian
V2N	Vehicle-to-Network
DSRC	Dedicated Short-Range Communication
LTE	Long-Term Evolution

5G	5th Generation
RA	Resource Allocation
C-V2X	Cellular Vehicle-to-Everything
QoS	Quality of Service
5G-NR	5th Generation New Radio
3GPP	3rd Generation Partnership Project
ITS	Intelligent Transport Systems
HARQ	Hybrid Automatic Repeat and Request
PSFCH	Physical Sidelink Feedback Channel
RSU	Road Side Unit
IEEE	Institute of Electrical and Electronics Engineers
ADAS	Advanced Driver Assistance System
YOLOv8	You Only Look Once version 8
CNN	Convolutional Neural Network
MAML	Model-Agnostic Meta-Learning
FPS	Frames Per Second
OpenCV	Open Source Computer Vision Library
RNN	Recurrent Neural Network
PC5	The direct communication method between vehicles and other devices
Uu	The network-based communication method between vehicles and the infrastructure via the cellular base station
TCP	Transmission Control Protocol
SHA-224	Secure Hash Algorithm-224
RGB	Red–Green–Blue
ROI	Region of Interest
LiDAR	Light Detection and Ranging
GPU	Graphic Processing Unit
GB	Gigabytes
RSA	Rivest–Shamir–Adleman
Dos	Denial-of-Service
GPS	Global Positioning System
JSON	JavaScript Object Notation

References

1. Mande, S.; Ramachandran, N. A Comprehensive Survey on Challenges and Issues in V2X and V2V Communication in 6G Future Generation Communication Models. *Ing. Des Syst. D'Information* **2024**, *6*, 951–960. [CrossRef]
2. IRF World Road Statistics. Total Road Network (by Road Type)—United States. Available online: http://18.134.46.19/dashboard/checksumgraph (accessed on 16 April 2025).
3. Louati, A.; Louati, H.; Kariri, E.; Neifar, W.; Hassan, M.K.; Khairi, M.H.H.; Farahat, M.A.; El-Hoseny, H.M. Sustainable Smart Cities through Multi-Agent Reinforcement Learning-Based Cooperative Autonomous Vehicles. *Sustainability* **2024**, *2*, 1779. [CrossRef]
4. Jain, H.; Babel, P. A Comprehensive Survey of PID and Pure Pursuit Control Algorithms for Autonomous Vehicle Navigation. *arXiv* **2024**, arXiv:2409.09848.
5. Hema, D.; Kannan, D.S. Interactive color image segmentation using HSV color space. *Sci. Technol.* **2020**, *6*, 37–41. [CrossRef]
6. Muslam, M.M.A. Enhancing security in vehicle-to-vehicle communication: A comprehensive review of protocols and techniques. *Vehicles* **2024**, *2*, 450–467. [CrossRef]
7. Malik, S.; Khan, M.A.; El-Sayed, H.; Khan, M.J. Should Autonomous Vehicles Collaborate in a Complex Urban Environment or Not? *Smartcities* **2023**, *7*, 2447–2483. [CrossRef]
8. Quadri, C.; Mancuso, V.; Ajmone Marsan, M.; Rossi, G.P. Platooning on the edge. In Proceedings of the 23rd International ACM Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems, Alicante, Spain, 16–20 November 2020; Volume 11, pp. 1–10.
9. Broggi, A.; Bertozzi, M.; Fascioli, A.; Bianco, C.G.L.; Piazzi, A. Visual perception of obstacles and vehicles for platooning. *IEEE Trans. Intell. Transp. Syst.* **2000**, *9*, 164–176. [CrossRef]
10. Yurtsever, E.; Lambert, J.; Carballo, A.; Takeda, K. A survey of autonomous driving: Common practices and emerging technologies. *IEEE Access* **2020**, *3*, 58443–58469. [CrossRef]

11. Kong, J.; Pfeiffer, M.; Schildbach, G.; Borrelli, F. Kinematic and dynamic vehicle models for autonomous driving control design. In Proceedings of the 2015 IEEE Intelligent Vehicles Symposium (IV), Seoul, Republic of Korea, 28 June–1 July 2015; Volume 8, pp. 1094–1099.
12. Naranjo, J.E.; González, C.; García, R.; de Pedro, T.; Haber, R.E. Power-steering control architecture for automatic driving. *IEEE Trans. Intell. Transp. Syst.* **2005**, *12*, 406–415. [CrossRef]
13. Wang, J.; Shao, Y.; Ge, Y.; Yu, R. A survey of vehicle to everything (V2X) testing. *Sensors* **2019**, *1*, 334. [CrossRef]
14. Sehla, K.; Nguyen, T.M.T.; Pujolle, G.; Velloso, P.B. Resource Allocation Modes in C-V2X: From LTE-V2X to 5G-V2X. *IEEE Internet Things J.* **2022**, *3*, 8291–8314. [CrossRef]
15. Garcia, M.H.C.; Molina-Galan, A.; Boban, M.; Gozalvez, J.; Coll-Perales, B.; Şahin, T.; Kousaridas, A. A Tutorial on 5G NR V2X Communications. *IEEE Commun. Surv. Tutor.* **2021**, *2*, 1972–2026. [CrossRef]
16. Mannoni, V.; Berg, V.; Sesia, S.; Perraud, E. A comparison of the V2X communication systems: ITS-G5 and C-V2X. In Proceedings of the 2019 IEEE 89th Vehicular Technology Conference (VTC2019-Spring), Kuala Lumpur, Malaysia, 28 April–1 May 2019; Volume 4, pp. 1–5.
17. CARLA Unreal Engine 4 Documentation. Available online: https://carla.readthedocs.io/en/latest (accessed on 3 August 2021).
18. IEEE Standard for Information Technology—Local and Metropolitan Area Networks—Specific Requirements—Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 6: Wireless Access in Vehicular Environments. Available online: https://standards.ieee.org/ieee/802.11p/3953/ (accessed on 17 June 2010).
19. IEEE Standard for Information Technology–Telecommunications and Information Exchange between Systems—Local and Metropolitan Area Networks–Specific Requirements—Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment: Enhancements for Ambient Power Communication (AMP). Available online: https://standards.ieee.org/ieee/802.11bp/11592/ (accessed on 21 March 2024).
20. Tammisetti, V.; Stettinger, G.; Cuellar, M.P.; Molina-Solana, M. Meta-YOLOv8: Meta-Learning-Enhanced YOLOv8 for Precise Traffic Light Color Detection in ADAS. *Electronics* **2025**, *14*, 468. [CrossRef]
21. Yao, Z.; Liu, Q.; Fu, J.; Xie, Q.; Li, B.; Ye, Q.; Li, Q. A Coarse-to-Fine Deep Learning Based Framework for Traffic Light Recognition. *IEEE Trans. Intell. Transp. Syst.* **2024**, *4*, 13887–13899. [CrossRef]
22. Huang, Y.; Tian, Z.; Jiang, Q.; Xu, J. Path tracking based on improved pure pursuit model and pid. In Proceedings of the 2020 IEEE 2nd International Conference on Civil Aviation Safety and Information Technology, Weihai, China, 14–16 October 2020; Volume 10, pp. 359–364.
23. Coulter, R.C. *Implementation of the Pure Pursuit Path Tracking Algorithm*; Carnegie Mellon University: Pittsburg, PA, USA, 1992; Volume 1, pp. 1–15.
24. Ang, K.H.; Chong, G.; Li, Y. PID control system analysis, design, and technology. *IEEE Trans. Control Syst. Technol.* **2005**, *7*, 559–576.
25. Khalid, I.; Maglogiannis, V.; Naudts, D.; Shahid, A.; Moerman, I. Optimizing hybrid v2x communication: An intelligent technology selection algorithm using 5g, c-v2x PC5 and dsrc. *Future Internet* **2024**, *3*, 107. [CrossRef]
26. Hur, D.; Lee, D.; Oh, J.; Won, D.; Song, C.; Cho, S. Survey on challenges and solutions of C-V2X: LTE-V2X communication technology. In Proceedings of the 2023 Fourteenth International Conference on Ubiquitous and Future Networks (ICUFN), Paris, France, 4–7 July 2023; Volume 7, pp. 639–641.
27. Sutradhar, K.; Pillai, B.G.; Amin, R.; Narayan, D.L. A survey on privacy-preserving authentication protocols for secure vehicular communication. *Comput. Commun.* **2024**, *3*, 1–18. [CrossRef]
28. Bazzi, A.; Berthet, A.O.; Campolo, C.; Masini, B.M.; Molinaro, A.; Zanella, A. On the design of sidelink for cellular V2X: A literature review and outlook for future. *IEEE Access* **2021**, *7*, 97953–97980. [CrossRef]
29. Verstraete, T.; Muhammad, N. Pedestrian collision avoidance in autonomous vehicles: A review. *Computers* **2024**, *3*, 78. [CrossRef]
30. YOLO12: Attention-Centric Object Detection. Available online: https://docs.ultralytics.com/models/yolo12 (accessed on 22 February 2025).
31. Merdrignac, P.; Shagdar, O.; Nashashibi, F. Fusion of perception and V2P communication systems for the safety of vulnerable road users. *IEEE Trans. Intell. Transp. Syst.* **2016**, *12*, 1740–1751. [CrossRef]
32. Fei, R.; Ma, M.; Hei, X.; Yang, L.; Li, A.; Qiu, Y. A new safety distance model based on braking process considering adhesion coefficient. *Electronics* **2024**, *1*, 421. [CrossRef]
33. Kasi, K.; Karuppanan, G. Framework to Identify Vehicle Platoons under Heterogeneous Traffic Conditions on Urban Roads. *Sustainability* **2024**, *1*, 78. [CrossRef]
34. Segata, M.; Bloessl, B.; Joerer, S.; Dressler, F.; Cigno, R.L. Supporting platooning maneuvers through IVC: An initial protocol analysis for the JOIN maneuver. In Proceedings of the 2014 11th Annual Conference on Wireless On-Demand Network Systems and Services (WONS), Obergurgl, Austria, 2–4 April 2014; Volume 4, pp. 130–137.
35. Tian, D.; Wang, Y.; Lu, G.; Yu, G. A VANETs routing algorithm based on Euclidean distance clustering. In Proceedings of the 2010 2nd International Conference on Future Computer and Communication, Wuhan, China, 21–24 May 2010; Volume 5, pp. 1–183.

36. Nagy, R.; Török, Á.; Pethő, Z. Evaluating V2X-Based Vehicle Control under Unreliable Network Conditions, Focusing on Safety Risk. *Appl. Sci.* **2024**, *6*, 5661. [CrossRef]
37. IEEE Standard for Wireless Access in Vehicular Environments–Security Services for Application and Management Messages. Available online: https://standards.ieee.org/ieee/1609.2/10258/ (accessed on 3 December 2022).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

electronics

MDPI

Article

Enhancing Automotive Intrusion Detection Systems with Capability Hardware Enhanced RISC Instructions-Based Memory Protection

Chathuranga Sampath Kalutharage [1,2,*], Saket Mohan [1], Xiaodong Liu [2] and Christos Chrysoulas [3]

[1] Secure Elements, Coventry CV1 2NT, UK; saket.mohan@secureelements.co.uk
[2] School of Computing, Engineering & the Build Environment, Edinburgh Napier University, Edinburgh EH10 5DT, UK; x.liu@napier.ac.uk
[3] School of Mathematical and Computer Sciences, Heriot-Watt University, Edinburgh EH14 4AS, UK; c.chrysoulas@hw.ac.uk
* Correspondence: sampath.kalutharage@secureelements.co.uk

Abstract: The rapid integration of connected technologies in modern vehicles has introduced significant cybersecurity challenges, particularly in securing critical systems against advanced threats such as IP spoofing and rule manipulation. This study investigates the application of CHERI (Capability Hardware Enhanced RISC Instructions) to enhance the security of Intrusion Detection Systems (IDSs) in automotive networks. By leveraging CHERI's fine-grained memory protection and capability-based access control, the IDS ensures the robust protection of rule configurations against unauthorized access and manipulation. Experimental results demonstrate a 100% detection rate for spoofed IP packets and unauthorized rule modification attempts. The CHERI-enabled IDS framework achieves latency well within the acceptable limits defined by automotive standards for real-time applications, ensuring it remains suitable for safety-critical operations. The implementation on the ARM Morello board highlights CHERI's practical applicability and low-latency performance in real-world automotive scenarios. This research underscores the potential of hardware-enforced memory safety in mitigating complex cyber threats and provides a scalable solution for securing increasingly connected and autonomous vehicles. Future work will focus on optimizing CHERI for resource-constrained environments and expanding its applications to broader automotive security use cases.

Keywords: automotive cybersecurity; IP spoofing; memory protection

1. Introduction

The automotive industry is undergoing transformative innovation, driven by rapid technological advancements and the advent of autonomous vehicles. Modern vehicles are no longer purely mechanical systems; they have evolved into complex mechatronic platforms, integrating numerous Electronic Control Units (ECUs) interconnected through sophisticated communication networks [1–3]. With the integration of advanced software systems and electronic hardware components, vehicles now support a wide array of functionalities. Notably, the rise of autonomous vehicles has driven a substantial increase in the deployment of ECUs. These ECUs manage critical vehicle functions, such as engine control, braking, and driver assistance, as well as secondary systems like lighting, infotainment, and connectivity services. Operating within interconnected networks, they utilize various communication protocols to ensure seamless and efficient performance [4].

However, this increased interconnectivity introduces a broad attack surface for cyber threats, posing significant risks to vehicle security, operational reliability, and passenger safety [5].

One of the critical components in ensuring vehicle cybersecurity is the Intrusion Detection System (IDS) [6]. IDSs play a pivotal role in detecting malicious activities at an early stage, thereby minimizing the potential impact of cyberattacks. They offer both known and unknown threats, and even zero-day vulnerabilities are derived based on pattern analysis within network traffic to deviate from usual behavior [7]. IDSs monitor network traffic to identify and reduce threats, including unauthorized resource access, data manipulation, and violation of protocol rules. However, traditional IDS solutions have failed to ensure the safety and security of automotive systems against complex attack vectors. Key concerns include, but are not limited to, IP spoofing—the attacker forges the source IP address to appear as a legitimate device–and manipulation of IDS rules, enabling them to bypass detection. Furthermore, Hamada et al. [8] present the development of an IDS at the central gateway of automotive networks to detect spoofing messages. This underscores the prevalence of message spoofing attacks and highlights the urgent need for robust detection mechanisms within automotive systems. These studies put together prove that traditional IDSs, while indispensable for automotive cybersecurity, often cannot handle such advanced and evolving attack vectors as IP spoofing and rule manipulation. This underscores the need for innovative approaches to strengthen IDS capabilities in modern automotive environments.

Other notable attack vectors include denial-of-service (DoS) attacks, which flood networks with excessive traffic to overwhelm IDS processing capabilities. These attacks are particularly concerning in resource-constrained environments, such as automotive networks, where even a single malicious source can disrupt operations. While distributed denial-of-service (DDoS) attacks are more complex and involve traffic from multiple sources, DoS attacks are easier to execute and can cause significant disruption with limited resources. This underscores the critical need for IDS solutions designed to withstand such targeted attacks. Additionally, message replay attacks disrupt operations by resending previously captured valid messages. Payload tampering exploits vulnerabilities through data packet injection or modification, while man-in-the-middle (MITM) attacks enable malicious actors to intercept, alter, or inject messages between ECUs. Timing attacks exploit delays in message processing, and protocol-specific exploits target communication standards like CAN, CAN-FD, or SOME/IP. These vulnerabilities highlight the limitations of traditional IDS solutions and the urgent need for robust, hardware-based security measures to safeguard IDS functionality.

As vehicles integrate more advanced connectivity and autonomous features, the attack surface for cyberattacks targeting Electronic Control Units (ECUs) expands significantly. Research highlights that modern vehicles contain an increasing number of ECUs, with high-end models surpassing 100 units, compared to around 25 in typical vehicles from a decade ago [9]. This proliferation of ECUs, coupled with their connectivity to external networks, correlates with a rising number of reported vulnerabilities. For example, studies have demonstrated that cyberattacks can remotely exploit these systems to disable critical functions, such as braking and steering, or even drain a vehicle's battery [10]. The number of connected vehicles globally is projected to reach over 400 million by 2025, further emphasizing the urgent need for robust cybersecurity measures [11]. Attackers can exploit network vulnerabilities to inject malicious packets or spoof legitimate IP addresses, bypassing conventional IDS mechanisms. Furthermore, unauthorized access to IDS configurations can enable attackers to modify detection rules, undermining the system's ability to flag

suspicious activity. Without strong protections, these threats compromise both vehicle safety and security.

However, traditional IDS solutions face significant challenges in addressing complex attack vectors such as IP spoofing, rule manipulation, and memory exploitation. These vulnerabilities highlight a critical gap in the current state of IDS development: the lack of robust hardware-enforced mechanisms for ensuring memory safety and rule integrity. Current IDS implementations rely heavily on software protections, which are insufficient against advanced attacks targeting ECUs in resource-constrained automotive environments. This study bridges this gap by leveraging the CHERI (Capability Hardware Enhanced RISC Instructions) architecture to introduce fine-grained memory protection for IDS systems. This novel approach ensures spatial and temporal memory safety, isolating IDS rules from unauthorized access and manipulation while maintaining compatibility with automotive standards.

This study contributes by introducing a novel CHERI-enhanced IDS framework tailored for automotive networks. Key contributions include the following:

- **Integration of CHERI Capabilities:** Demonstrating the application of CHERI to enforce memory protection, preventing unauthorized access to IDS configurations.
- **The 100% Prevention of IDS Rule Manipulation**: Ensuring rule integrity under simulated attack scenarios, including IP spoofing and rule tampering.
- **Real-Time Detection of Spoofing Attacks**: Accurate anomaly detection with minimal performance overhead, ensuring suitability for real-time automotive environments.
- **Evaluation on Realistic Automotive Scenarios**: Simulation of ECUs and attack models to validate the effectiveness of the proposed system, adhering to AUTOSAR standards.

The remainder of this paper is organized as follows: Section 2 provides the Background, outlining key challenges in automotive cybersecurity, IDS vulnerabilities, and an introduction to CHERI technologies. Section 3 presents the Literature Review, discussing related work on IDS frameworks, CHERI applications, and automotive network security. Section 4 explains the Methodology, detailing the experimental setup, CHERI-enhanced IDS framework, and attack simulation scenarios. Section 5 covers the Results and Analysis, highlighting the effectiveness of CHERI in mitigating IP spoofing and rule manipulation attacks. Finally, Section 6 concludes the paper, offering insights into CHERI's potential in real-world automotive applications and recommendations for future research directions.

2. Background

2.1. CHERI Project Overview

CHERI (Capability Hardware Enhanced RISC Instructions) is a novel hardware-based architectural framework designed to provide fine-grained memory safety and compartmentalization [12]. It replaces traditional pointers with capabilities—enhanced constructs carrying metadata, including base, bounds, and permissions, enabling explicit control over memory access. This approach addresses vulnerabilities such as buffer overflows, unauthorized memory access, and code injection attacks.

The CHERI project is a research initiative originally developed through a collaboration between the University of Cambridge, SRI International, and ARM Research, with funding from the DARPA-sponsored CRASH and MRC programs [13]. At its core, CHERI introduces a capability-based model that extends traditional architectures, offering hardware-enforced spatial memory safety by preventing out-of-bounds memory access. This design significantly reduces vulnerabilities like buffer overflows and code injection [14].

Compared to traditional software-only approaches, CHERI provides hardware-enforced isolation with minimal performance overhead. While software solutions rely on runtime checks, CHERI integrates security at the hardware level, ensuring strict enforcement of the principles of least privilege and intentional use. Furthermore, its compatibility with widely used languages like C and C++ allows seamless integration with existing systems without requiring extensive code rewrites. This compatibility is especially critical for automotive systems, where C/C++ dominate due to their low-level control and performance efficiency [15].

In our study, we demonstrated that incorporating CHERI into Intrusion Detection Systems (IDSs) significantly enhances their resilience against advanced threats, including IP spoofing and IDS rule manipulation attacks. Unlike software-only IDS frameworks, CHERI protects rule configurations by enforcing memory bounds, preventing attackers from altering detection logic. This hardware-enforced security guarantees rule integrity, even against attackers using legitimate-appearing spoofed IP addresses.

The ARM Morello board serves as a prominent implementation of the CHERI architecture, showcasing its practical applicability in real-world scenarios [16]. By adopting CHERI-based solutions, the automotive industry can address network and software-level vulnerabilities, achieving robust security by design. Furthermore, CHERI aligns with emerging standards like ISO/SAE 21434 [17,18], making it a viable option for protecting critical systems from sophisticated attacks such as rule manipulation and IP spoofing.

Our research highlights the practicality of CHERI by implementing it on the ARM Morello board, showcasing its low-latency performance and alignment with real-time automotive environments. By adopting CHERI-based solutions, the automotive industry can effectively address network- and software-level vulnerabilities, ensuring a robust defense for connected and autonomous vehicles.

2.2. CHERI and ARM Morello Technologies

The CHERI (Capability Hardware Enhanced RISC Instructions) architecture represents a transformative approach to addressing memory safety vulnerabilities, one of the most critical challenges in modern computing systems. By integrating fine-grained memory protection and scalable software compartmentalization directly into hardware, CHERI provides a robust mechanism for securing both traditional and modern computing environments. The novelty of this research lies in its application of CHERI's memory protection capabilities to enhance IDS systems in automotive networks. Unlike traditional IDS solutions, the proposed framework integrates hardware-enforced security mechanisms to protect IDS configurations and detect advanced threats, such as IP spoofing and rule manipulation, in real time. By embedding CHERI's capabilities into automotive-friendly programming environments like C/C++, this approach ensures robust cybersecurity measures without compromising system performance.

2.2.1. CHERI Architecture Overview

CHERI is built upon two foundational principles [19]:

1. **Principle of Least Privilege**: This principle ensures that each component operates with only those permissions required to perform its function, thus reducing the attack surface area and limiting the damage in case of compromise. In automotive cybersecurity, this becomes even more important to ensure that sensitive subsystems, such as IDS rule configurations and ECU communications, are protected. By keeping each process strictly within its predefined boundaries, CHERI hardware prevents unauthorized accesses to critical memory regions. For example, the capability model

of CHERI prevents an attempt by a compromised ECU to access the IDS rules or manipulate configurations and thus isolates and contains the threat.

2. **Principle of Intentional Use**: A principle that enforces the explicit declaration of permissions and memory accesses so that actions actually carried out by a program are intentional and well authorized. This rules out ambiguities that result in misusing or exploiting privileges, such as the confused deputy problem. Scientific studies have shown that intentional use policies can greatly enhance system robustness since it ensures every action follows some predetermined constraints set in security [19,20].

Taken together, these principles form the bedrock of how CHERI is able to offer fine-grained memory safety and software compartmentalization. In the scope of this work, they guarantee IDS rules are isolated from unauthorized processes, and their usage is only possible through explicit and intentional permissions, which in turn mitigate critical vulnerabilities such as the manipulation of rules and IP spoofing. CHERI capabilities are hardware-enforced tokens that combine memory addresses with metadata, such as bounds and permissions, ensuring that memory access is strictly controlled. Each memory region is protected by a unique capability, and these capabilities are immutable in ways that increase their security. For instance, CHERI employs tagged memory to distinguish valid capabilities from arbitrary data, preventing their corruption or unauthorized modification.

2.2.2. CHERI Enhancements and Implementation

CHERI enhances conventional architectures like MIPS, RISC-V, and ARM by introducing a capability-based memory model that coexists with existing software [21]. This hybrid approach facilitates the gradual adoption of CHERI technologies, enabling legacy code to operate alongside CHERI-enabled applications. The architecture extends pointers to include additional metadata, such as base, bounds, and permissions, ensuring that only authorized memory regions are accessed.

Key features of CHERI include the following:

- Memory Protection: By attaching bounds and permission bits to pointers, CHERI ensures spatial memory safety, preventing unauthorized access beyond designated memory regions.
- Software Compartmentalization: CHERI supports the isolation of software components, enabling secure interaction between mutually untrusting programs or processes.
- Tagged Memory: This feature ensures that capabilities cannot be forged or corrupted, as the hardware tracks their validity independently of their location in memory.

2.2.3. ARM Morello: A Prototype for Commercial Application

ARM Morello, a prototype implementation of the CHERI architecture, extends the ARMv8-A architecture with CHERI capabilities to evaluate its feasibility in commercial processors. Developed as part of a collaboration with the UK government, Morello explores the integration of CHERI's fine-grained memory protection mechanisms into mass-market hardware. By leveraging CHERI's capability-based security model, Morello showcases the practical implementation of hardware-enforced memory safety and software compartmentalization, which are critical for mitigating modern cybersecurity threats and addressing evolving challenges [22].

The Morello board incorporates several CHERI advancements:

- Compatibility with Existing Software: Morello provides a hybrid execution environment, enabling unmodified ARMv8-A applications to coexist with CHERI-enhanced programs. This compatibility reduces adoption barriers for developers and organizations.

- Enhanced Security Properties: By integrating CHERI capabilities, Morello strengthens security guarantees for memory safety, compartmentalization, and access control at the hardware level.
- Performance Considerations: ARM Morello demonstrates that CHERI's additional security features can be implemented with minimal performance overhead, making it viable for real-time systems like those in automotive networks.

2.3. IP Spoofing and IDS Rule Manipulation

Spoofing is one of the most general techniques in cyberattacks, where an attacker forges the source IP address in packets to impersonate a legitimate ECU within an automotive network. This tactic can allow unauthorized access to critical systems, disrupt operations, or inject malicious commands within automotive systems. For instance, an attacker may impersonate a trusted ECU in order to manipulate either the braking or steering functions, which will be very dangerous [23,24]. The distributed and resource-constrained nature of automotive networks makes the spoofed packet detection even more challenging. Most of the automotive IDSs depend on IP-based filtering that is generally too weak to identify forged packets when the attacker uses a legitimate-appearing IP address. For example, some works indicate that spoofing attacks of this type could easily evade traditional IDS systems because of a lack of strong memory protection combined with dynamic rule adaptation mechanisms [25,26]. Once attackers establish a foothold within the network, they may attempt to alter IDS rules to avoid detection in future attacks. This manipulation could involve disabling specific detection rules or creating exceptions for their activity [26]. Without robust memory protection for IDS configurations, rule manipulation remains a persistent vulnerability, as rule-based detection alone cannot prevent access to the IDS's internal logic if an attacker gains sufficient privileges.

The following challenges must be addressed with hardware-enforced security mechanisms, such as CHERI capabilities, for advanced IDS solutions, ensuring strict memory isolation for IDS rule configurations. In particular, storing the IDS rules in protected memory, CHERI efficiently prevents attackers from tampering with the detection logics even when spoofed packets are received. Such hardware-enforced protection strengthens IDS resiliency in accurately detecting anomalies without significant performance trade-offs [25,27]. Furthermore, most of the IP spoofing attacks are part of larger campaigns that involve other attack vectors, such as rule manipulation and denial-of-service attacks. These combined threats indicate the need for comprehensive security frameworks, hardware-enhanced for automotive security. This will help in mitigation not only against IP spoofing but also provide long-term scalability and adaptability against upcoming security challenges in connected and autonomous vehicles by integrating CHERI capabilities within IDS systems.

3. Literature Review

Automotive networks have evolved from isolated, closed systems to highly interconnected networks due to the inclusion of wireless interfaces such as Wi-Fi, Bluetooth, and cellular connectivity [1]. These interfaces, while providing benefits like over-the-air (OTA) updates and telematics services, expose vehicles to numerous cyber risks [23,28]. One of the primary concerns is the complexity and distributed nature of automotive software, with modern vehicles containing over 100 million lines of code distributed across ECUs. These ECUs interconnect through the Controller Area Network (CAN) and Ethernet networks, using protocols such as Scalable service-Oriented Middleware over IP (SOME/IP) to enable efficient communication.

To enhance communication security in automotive electrical and electronic architectures, various approaches have been developed. One notable example is SecOC, standard-

ized by AUTOSAR, which employs symmetric cryptography to secure communication over the Controller Area Network (CAN) [29]. Lee et al. in [30] suggest a hybrid framework leveraging blockchain technology to improve network security within vehicles. Nevertheless, the framework does not include a comprehensive explanation of the protocol's interaction process or demonstrate the practicality of the proposed solution. Islam et al. in [31] present a CAN-based scheme designed to mitigate multiple attack scenarios. They assert that the method ensures security without requiring changes to the underlying CAN protocol. Additionally, validation experiments are conducted, demonstrating the effectiveness of the proposed approach in preventing and defending against a range of attacks. Hafeez et al. in [24,32] employ a neural network approach to authenticate messages within vehicle networks, including the conventional CAN protocol. They assert that this method significantly enhances the accuracy of Electronic Control Unit (ECU) communication. While the study includes analysis and validation of the approach, it does not provide a detailed assessment of its security. Woo et al. in [33] propose a security architecture leveraging the CAN with Flexible Data Rate (CAN-FD) network to establish a secure transmission environment. They introduce a hierarchical encryption transmission technique for the vehicle network, which has been validated but lacks a formal verification process to confirm its feasibility. Lodge et al. [34] analyze vulnerabilities in CAN and CAN-FD protocols, such as replay, injection, and DoS attacks, emphasizing the need for robust automotive security frameworks. They propose solutions like lightweight cryptography, blockchain-based key provisioning, and multi-layer architectures to enhance communication security. While promising, these methods face challenges in compatibility with legacy systems and increased complexity, highlighting the importance of advanced IDS for Ethernet-based networks.

Kreissl et al. in [35] explored methods for securing SOME/IP-based vehicular communication networks. The authors suggest a central entity to manage key material distribution through (D)TLS during the event group subscription process. However, the "service offer" and "find offer" steps remain unprotected. To support broadcast communication, SOME/IP is enhanced with the TESLA protocol. Despite its advantages, TESLA's delayed authentication for broadcast messages requires modifications to the SOME/IP process flow, such as caching messages until verification data is received. This introduces latency, making it unsuitable for time-sensitive data transmission. Iorio et al. [36] propose a new framework to enhance the security of SOME/IP. While the framework includes security protocol modeling and verification, offering strong assurance of the desired security properties, it imposes significant network overhead and lacks a clear and intuitive formal verification proof. Herold et al. [37] propose an Intrusion Detection System (IDS) for SOME/IP using Complex Event Processing (CEP) to detect attacks. The system identifies threats such as malformed packets, protocol violations, and timing issues. However, like other IDS solutions, it faces challenges with false positives and false negatives. Zorman et al. [27] propose a firewall implementation for embedded automotive systems, utilizing rule-based security and deep packet inspection (DPI) to validate SOME/IP payloads. The solution demonstrates feasibility in resource-constrained environments, highlighting its potential for enhancing security in automotive Ethernet communication. Qi Liu et al. [38] proposed a method that combines deep learning models with residual self-attention to analyze Ethernet traffic, aiming to detect cyberattacks like DDoS and MITM, as well as random hardware failures. This framework demonstrates high accuracy and real-time detection capabilities, highlighting the importance of robust cybersecurity solutions for modern Ethernet-based automotive communication systems.

Recent advances in hardware-based security frameworks have introduced novel approaches to mitigating vulnerabilities in modern automotive systems. Among these, Trusted Execution Environments (TEEs) and Hardware Security Modules (HSMs)have

gained prominence. TEEs, such as Intel SGX and ARM TrustZone, provide isolated secure execution environments for sensitive operations, protecting them from unauthorized access by the main operating system. TEEs are widely applied in automotive systems to safeguard critical processes such as firmware updates, cryptographic computations, and secure communications. For example, studies demonstrate that TEEs can effectively isolate critical vehicle control logic, ensuring protection against malware or unauthorized modifications [39,40]. However, TEEs focus primarily on secure execution and do not directly address memory safety violations, such as buffer overflows or memory corruption, which are significant concerns in automotive ECUs [41].

Hardware Security Modules (HSMs) are dedicated hardware devices designed to securely store cryptographic keys and perform cryptographic operations. In automotive applications, HSMs are integrated into Electronic Control Units (ECUs) to enable secure boot, message authentication, and key management. The AUTOSAR Secure Onboard Communication (SecOC) module relies heavily on HSMs to ensure message integrity and authenticity in Controller Area Networks (CANs) [29,41]. Despite their utility in cryptographic tasks, HSMs do not inherently protect against broader security challenges, such as memory exploitation or unauthorized rule manipulation.

In contrast to TEEs and HSMs, the CHERI (Capability Hardware Enhanced RISC Instructions) architecture provides fine-grained memory safety and compartmentalization by implementing a capability-based hardware model. CHERI prevents memory corruption and unauthorized access to critical IDS configurations, offering hardware-enforced isolation that mitigates advanced attack vectors, such as rule manipulation and spoofing. While TEEs ensure the secure execution of processes and HSMs protect cryptographic assets, CHERI focuses on memory safety at the hardware level, making it uniquely suited for Intrusion Detection Systems in automotive environments [42].

Intrusion Detection Systems (IDS) in automotive environments monitor network traffic to identify suspicious activities that may indicate an attack. Common threats include IP spoofing, where attackers modify IP headers to disguise their identity. While IDS can flag such anomalies, conventional approaches are often limited in countering spoofing when attackers effectively mask their actions. Additionally, if attackers compromise IDS configurations, they could manipulate detection rules to allow undetected future attacks. This research addresses a critical gap in existing solutions, which do not tackle IP spoofing and rule manipulation attacks with a security-by-hardware design approach. To the best of our knowledge, this is the first study to leverage CHERI's hardware-enforced memory safety features to mitigate IDS rule manipulation attacks and detect IP spoofing in automotive networks.

4. Methodology

This section describes the methodology and experimental setup, including the network and components simulated, the configuration of the IDS, and the process of executing spoofing and rule manipulation attacks. The experiment evaluates the CHERI capabilities in restricting unauthorized access and protecting the IDS configuration in an automotive network.

4.1. CHERI in Automotive Security Applications

The integration of CHERI (Capability Hardware Enhanced RISC Instructions) into automotive security offers a transformative approach to mitigating critical vulnerabilities in modern vehicles. CHERI introduces fine-grained memory protection by replacing traditional pointers with capabilities, which include metadata such as base, bounds, and permissions. This architecture prevents unauthorized access and memory safety violations,

addressing issues such as buffer overflows and code injection—common attack vectors in automotive systems.

In automotive applications, CHERI can enhance the security of Ethernet-based protocols like SOME/IP by safeguarding IDS configurations and protecting critical ECUs from malicious tampering. By embedding CHERI capabilities, automotive systems can enforce hardware-level isolation for IDS rules, preventing rule manipulation and ensuring secure anomaly detection. Additionally, CHERI mitigates risks associated with IP spoofing by ensuring that memory access is explicitly authorized, eliminating the possibility of attackers altering network configurations.

This approach aligns with the principles of security by design, enabling robust protections against advanced cyberattacks in the dynamic and interconnected environment of modern vehicles. CHERI's capability-based model ensures that automotive security solutions remain both effective and resilient, providing a foundational framework for safeguarding next-generation vehicle networks.

4.2. CHERI Capability Usage

CHERI (Capability Hardware Enhanced RISC Instructions) enhances security by replacing traditional pointers with capabilities. Capabilities are pointers with extra information, such as memory bounds and access permissions, that strictly control how memory can be accessed. This ensures that a program cannot inadvertently or maliciously access memory outside its designated boundaries. For instance, in standard C programming, a pointer can move freely across memory regions, which might allow it to read or overwrite sensitive data unintentionally. CHERI enforces strict boundaries for each pointer. If a pointer tries to access memory outside its permitted range, the hardware immediately stops the action and raises an alert. Memory safety violations often result from coding errors in unsafe languages like C and C++, making software susceptible to security vulnerabilities and unauthorized access. CHERI introduces capabilities that enforce spatial memory safety, effectively addressing these issues. Figure 1 illustrates a comparison between the behavior of ISO C and CHERI C in handling pointer arithmetic and memory access.

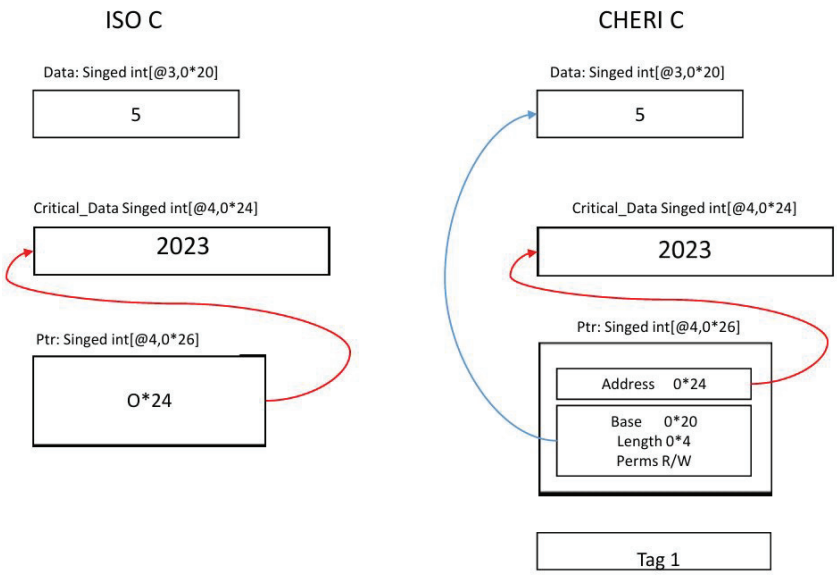


Figure 1. Fine-grained memory protection: comparing ISO C with CHERI C. In ISO C, the flawed program allows undefined behavior, leading to the potential leakage of critical_data due to unrestricted pointer arithmetic. In contrast, CHERI C enforces strict memory bounds through capabilities, ensuring that any out-of-bounds memory access triggers a hardware exception, thereby preventing unauthorized access to critical_data.

In ISO C, two variables, data and critical_data, are declared:

- data is allocated at address 0x20 with a value of 5.
- critical_data is allocated at address 0x24 with a value of 2023.

A pointer, ptr, is initialized to point to data. In ISO C, pointer arithmetic allows ptr to be incremented to point to 0x24, accessing and leaking the value of critical_data unintentionally. This behavior is described mathematically in Equation (1).

$$\begin{aligned} &\text{Variable } data \text{ : signed int, allocated at address } 0x20, \text{ value } 5. \\ &\text{Variable } critical_data \text{ : signed int, allocated at address } 0x24, \text{ value } 2023. \\ \textbf{ISO C:}\ &\text{Pointer } ptr \text{ : signed int*, initialized to address of } data. \\ &\text{After increment: } ptr = ptr + 1 \implies ptr = 0x24. \\ &\text{Dereferencing: } y = *ptr \implies y = critical_data = 2023. \end{aligned} \tag{1}$$

In **CHERI C**, pointers are replaced with capabilities that include metadata such as `base`, `length`, and `permissions`. The capability for `ptr` is confined to the memory bounds of `data` (`0x20` to `0x24`). Any attempt to dereference `ptr` outside these bounds (e.g., at `0x24`) triggers a hardware trap, preventing unauthorized access to `critical_data`. This behavior is defined in Equation (2). CHERI enforces the **principle of least privilege** by ensuring `ptr` cannot access memory beyond its defined bounds, maintaining memory safety by design.

$$\begin{aligned} &\text{Variable } data \text{ : signed int, allocated at address } 0x20, \text{ value } 5. \\ &\text{Variable } critical_data \text{ : signed int, allocated at address } 0x24, \text{ value } 2023. \\ \textbf{CHERI C:}\ &\text{CHERI C Pointer } ptr \text{ : signed int* with metadata as a CHERI capability :} \\ &\quad \text{Address: } 0x20, \text{ Base: } 0x20, \text{ Length: } 0x4, \text{ Permissions: R/W, Tag: Valid.} \\ &\text{Dereferencing: } y = *ptr \implies \text{Bounds check by CHERI: Access Denied if} \\ &\quad ptr > 0x20 + 0x4. \end{aligned} \tag{2}$$

By integrating CHERI capabilities into the program, the hardware ensures that memory safety violations, such as accessing `critical_data` through `ptr`, are impossible. This approach not only protects against common vulnerabilities like buffer overflows and out-of-bounds memory access but also mitigates intentional attacks aimed at exploiting memory corruption.

This demonstrates how CHERI implements **memory safety by design** on our IDS, transforming insecure C code into a secure execution model without significant performance overhead or extensive code changes. As C remains a dominant and automotive-friendly language due to its low-level hardware access, efficiency, and real-time capabilities, CHERI's compatibility with C ensures that security can be integrated into existing automotive systems without compromising these critical advantages.

Implementation: CHERI Usage on IDS

The integration of CHERI (Capability Hardware Enhanced RISC Instructions) within the Intrusion Detection System (IDS) framework was designed to utilize its inherent capability-based memory model to ensure fine-grained memory protection and enhance the security of automotive networks. Specifically, IDS rules are formulated based on SOME/IP packet features as per AUTOSAR standards, ensuring adherence to industry protocols. To maintain robust security, each ECU's rule set is stored in a dedicated and isolated memory region, safeguarded by CHERI's hardware-enforced capabilities. This isolation

prevents unauthorized access and ensures that packets are validated strictly against the corresponding ECU's rules, eliminating cross-rule interference.

CHERI's unique memory safety features, such as embedding metadata with bounds and permissions into pointers, enable strict compartmentalization. These capabilities guarantee that any access attempt beyond defined memory bounds triggers hardware traps, effectively preventing rule manipulation or unauthorized modification. During operation, incoming SOME/IP packets are analyzed in real time, with their attributes compared against the CHERI-protected rule sets. Packets failing to meet the defined criteria are flagged as anomalous. This implementation ensures that even sophisticated attacks, such as IP spoofing and rule manipulation, are mitigated with 100% accuracy. By leveraging CHERI's hardware-enforced memory protection, the IDS framework enhances resilience while maintaining compliance with real-time performance requirements in automotive environments.

4.3. Memory-Protected IDS

The proposed methodology integrates CHERI (Capability Hardware Enhanced RISC Instructions) to fortify automotive Ethernet-based Intrusion Detection Systems (IDSs) against advanced attack scenarios, such as IP spoofing and rule manipulation. By leveraging CHERI's capability-based memory protection, the framework ensures that IDS rules and configurations are isolated from unauthorized access, even in the event of system compromise. This section explains the approach with examples, attack scenarios, and code-level defenses.

In IP spoofing attacks, an attacker forges the source IP address in network packets to mimic a legitimate ECU, enabling unauthorized access or manipulation of in-vehicle communication. For example, a spoofed packet may claim to originate from 192.168.1.11 (a trusted ECU) but carry malicious payloads targeting safety-critical systems. This can be defended against. Using CHERI, IDS rules are stored with strict memory bounds, ensuring only authenticated packets can access or modify ECU configurations. Spoofed packets are detected by comparing their source IP and payload against predefined rules confined to CHERI-protected memory as in Listing 1.

Listing 1. Function to check ECU rules using CHERI capabilities.

```
int check_ecu_rule(const char *src_ip, const struct someip_packet *
    someip_pkt) {
    for (int i = 0; i < 5; i++) {
        const char *ecu_ip = (const char *)ecu_memory[i]->ecu_ip; //
             Access protected by~CHERI

        // Match source IP and packet attributes
        if (strcmp(src_ip, ecu_ip) == 0 &&
            ntohs(someip_pkt->method_id) == ecu_memory[i]->method_id
                 &&
            ntohs(someip_pkt->client_id) == ecu_memory[i]->client_id
                 &&
            ntohs(someip_pkt->session_id) == ecu_memory[i]->
                session_id) {
            return i;  // Rule matched
        }
    }
    return -1;  // No match found (potential spoofed packet)
}
```

Rule manipulation attacks occur when an attacker gains unauthorized access to the IDS configuration to modify detection thresholds or disable specific rules, enabling malicious packets to bypass detection. **CHERI** isolates IDS rules within protected memory regions, ensuring that only authorized processes can modify these configurations. For example, the `initialize_ecu_memory` function demonstrates how CHERI enforces tight memory bounds to prevent unauthorized access or modification, with any violations triggering hardware traps as in Listing 2.

Listing 2. Function to initialize ECU memory with CHERI capabilities.

```
void initialize_ecu_memory() {
    for (int i = 0; i < 5; i++) {
        // Allocate memory with CHERI capabilities
        ecu_memory[i] = (struct ecu_rule *__capability)malloc(sizeof
            (struct ecu_rule));
        if (!ecu_memory[i]) {
            perror("Failed to allocate memory for ECU rule");
            exit(EXIT_FAILURE);
        }
    }

    // Assign rules and enforce bounds
    ecu_memory[0]->ecu_ip = (char *__capability)cheri_setbounds("
        192.168.1.11", sizeof("192.168.1.11"));
    ecu_memory[0]->method_id = 0x1234;
    ecu_memory[0]->client_id = 0x5678;
    ecu_memory[0]->session_id = 0x9abc;
}
```

The **cheri_setbounds** function ensures that IDS rules are stored in tightly bounded memory regions. This prevents unauthorized processes from accessing or altering these rules. Any attempt to modify the rules outside authorized contexts results in a deterministic hardware trap.

Furthermore, the IDS must monitor and analyze network traffic for potential anomalies in real time, including spoofed packets and unauthorized rule manipulations. During live packet capture, the IDS compares incoming packets against the CHERI-protected rule set. CHERI ensures that only validated rules are accessed during this process, flagging spoofed or tampered packets. Packets are captured and validated against the CHERI-protected rule set, as shown in Listing 3. Any anomaly, such as a spoofed IP or a rule mismatch, triggers a warning, enabling the IDS to respond in real time. The **libpcap** library is used to capture network traffic in real time. Additionally, the libraries **arpa/inet.h, netinet/ip.h, and netinet/udp.h** are utilized to parse and analyze specific fields in the captured packets, such as IP and UDP headers.

Listing 3. Function to process packets and validate against ECU rules.

```
void process_packet(u_char *args, const struct pcap_pkthdr *header,
    const u_char *packet) {
    struct ip *ip_hdr = (struct ip *)(packet + 14);  // Parse IP
        header
    char src_ip[INET_ADDRSTRLEN];

    inet_ntop(AF_INET, &(ip_hdr->ip_src), src_ip, INET_ADDRSTRLEN);

    if (ip_hdr->ip_p == IPPROTO_UDP) {  // Check for UDP packets
```

```
        const struct someip_packet *someip_pkt = (struct
            someip_packet *)(packet + 14 + ip_hdr->ip_hl * 4);
        int rule_index = check_ecu_rule(src_ip, someip_pkt);

        if (rule_index >= 0) {
            printf("Valid packet matched with ECU%d rules\n",
                rule_index + 1);
        } else {
            printf("Unmatched packet from %s! Possible attack
                detected.\n", src_ip);
        }
    }
}
```

4.4. Experimental Setup

The experimental setup includes several simulated components representing elements of an automotive network, such as ECUs, an IDS, a packet sniffer, and an attacker model, as shown in Figure 2. The network layout mimics a typical vehicular communication structure, where the IDS monitors network traffic between ECUs and detects potential anomalies.

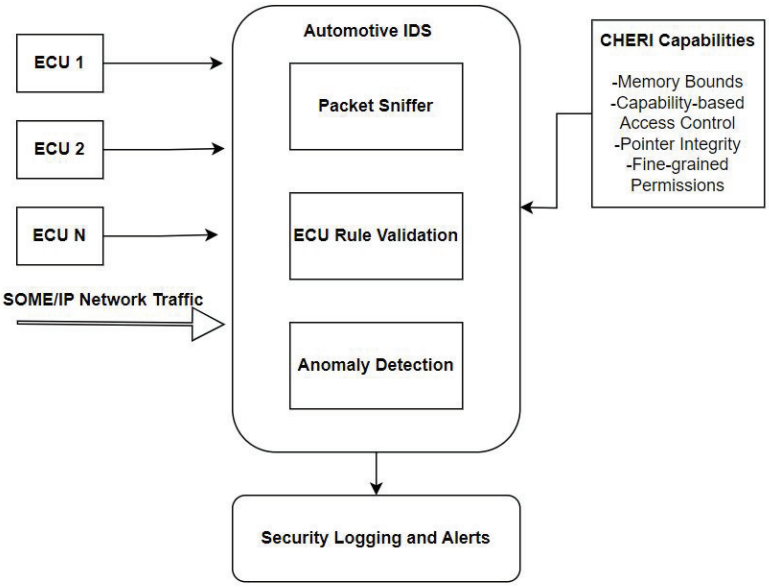


Figure 2. Automotive Intrusion Detection System (IDS) with CHERI capabilities experimental setup: The diagram illustrates the experimental setup of an IDS architecture for automotive networks. SOME/IP traffic from multiple ECUs, along with potential attacker scripts masquerading as random ECUs, is processed through a pipeline comprising packet sniffing, ECU rule validation, and anomaly detection. CHERI capabilities enhance the IDS by enforcing memory bounds, capability-based access control, pointer integrity, and fine-grained permissions. These features ensure robust protection against threats such as spoofed packets, rule manipulation, and unauthorized ECU impersonation. Detected anomalies are logged and trigger alerts, maintaining the integrity and security of the network.

- **ECU Components (ECU1, ECU2, ECU3, ECU4, and ECU5)**: These scripts simulate legitimate ECUs within the vehicle's network, transmitting standard, authorized packets. `ECU1` and `ECU2` generate and send routine communication packets across the network to emulate regular vehicular functions.
- **Intrusion Detection System (IDS)**: This IDS script monitors the network for suspicious packets or anomalies. With CHERI capabilities enabled, the IDS memory is

compartmentalized, ensuring that only authorized processes can access or modify its rule set. The IDS is configured to detect IP spoofing attempts and log unauthorized access to the IDS memory or rule configurations. Separate memory locations are defined on the IDS for each ECU's rules, meaning that each ECU has its own set of rules stored in an isolated memory location.

- **Attacker Model (Attacker)**: The attacker model is designed to send spoofed IP packets, using legitimate IP addresses to disguise malicious activities. The attacker's objective is to bypass the IDS by impersonating a legitimate ECU. If successful, the attacker aims to modify IDS rules to prevent future detections and alter the rules to allow malicious data to be sent to critical ECUs.
- **Packet Sniffer (Sniff)**: This component captures network traffic for analysis and logs all packets, enabling verification of whether spoofed packets are correctly identified by the IDS.

4.5. Execution of Attack

The Attacker: The attacker injects packets with spoofed IP addresses, mimicking the IP addresses of `ECU1` and `ECU2` to evade detection. The IDS monitors network traffic, comparing expected and actual packet source data to detect anomalies. CHERI capabilities within the IDS memory ensure that unauthorized packets cannot influence IDS functionality beyond their initial reception. Subsequently, the attacker attempts to access and modify the IDS rule set. By spoofing legitimate ECU IP addresses, the attacker seeks to gain access to IDS rule configurations and alter them to disable detection for future spoofing activities. However, the CHERI-enhanced IDS enforces strict memory protection, preventing unauthorized access or modification of its rules. Any access violation attempt is flagged by CHERI's memory bounds checks, and the IDS logs the incident while ensuring the attacker's script is terminated.

4.6. CHERI-Enhanced Memory Protections

CHERI capabilities introduce fine-grained memory protections to the IDS. In this experiment, the CHERI memory model is applied to compartmentalize IDS configurations, ensuring the following:

- **Access Permissions**: Only authorized memory regions can be accessed by the IDS. Unauthorized attempts to modify IDS rules or configurations trigger CHERI alerts.
- **Memory Bounds**: The memory boundaries around IDS rules are strictly enforced, preventing any modification or access from processes outside of predefined bounds.
- **Logging Violations**: Each unauthorized memory access attempt is logged, providing detailed records of any access violations. This functionality allows for verification of CHERI's effectiveness in blocking unauthorized access to IDS configurations.

By structuring the IDS with CHERI-enabled protections, the experiment tests how effectively CHERI prevents spoofed packets from bypassing IDS detection and protects against unauthorized rule manipulation.

5. Results and Analysis

The results section presents the findings from the attack simulation, demonstrating the impact of CHERI capabilities on the effectiveness of the IDS in detecting spoofed IP packets and blocking unauthorized access attempts. Table 1 outlines the simulation setup, which adheres to the AUTOSAR standard and is implemented on the Arm Morello board. Powered by a Cortex-A72 processor with up to eight cores, the board operates at typical frequencies around 1.5 GHz and integrates CHERI capabilities for fine-grained memory safety, including capability-based addressing, pointer integrity, and bound checking. These

features provide robust protection against vulnerabilities such as buffer overflows and use-after-free errors. The Morello board is equipped with 1 GB of DDR4 RAM, enabling efficient memory usage through CHERI's tagged memory system, which adds metadata for access control and memory isolation. Storage is supported via onboard eMMC and expansion options, while networking capabilities include integrated Gigabit Ethernet for high-speed data transfer. Additional I/O interfaces include USB 3.0 and USB 2.0 ports, HDMI for graphical output, and UART and JTAG for debugging. The platform also offers PCIe slots for peripheral expansions and GPIO pins for hardware interfacing, making it highly versatile for development and experimentation.

Table 1. Packet sending rates for ECU attack scenarios over 30 min.

ECU Type	Packet Rate (Packets/s)	Number of ECUs	Duration (s)	Total Packets Sent
Telemetry ECU	12.5	1	1800	22,500
Control Signal ECU	30	2	1800	108,000
Sensor Data ECU	75	2	1800	270,000

The IDS successfully flagged 100% of anomalous packets originating from undefined ECUs as shown in Figure 3, demonstrating a high detection rate for known attacks. In our tests, 15 undefined IPs were used to repeatedly send multiple packets over a 30-min timeframe, all of which were accurately detected. Furthermore, packets from known IPs were sent with altered features that deviated from the predefined rules assigned to each ECU and IP. In each scenario, the IDS accurately flagged the anomalies. The CHERI capabilities enabled the IDS to maintain rule integrity, preventing unauthorized packets from influencing detection mechanisms. This demonstrates that CHERI-enhanced memory isolation significantly enhances detection robustness, even against attackers using legitimate IPs as a disguise. The latency of the IDS response to anomalous packets was also measured to evaluate the performance impact of CHERI protections. Observed latency increases were minimal, indicating that the CHERI enhancements do not significantly affect IDS performance. This minimal overhead makes CHERI a viable solution for real-time applications in automotive IDS contexts. Furthermore, no existing literature could be found that use CHERI-based IDS for automotive applications, limiting direct comparisons.

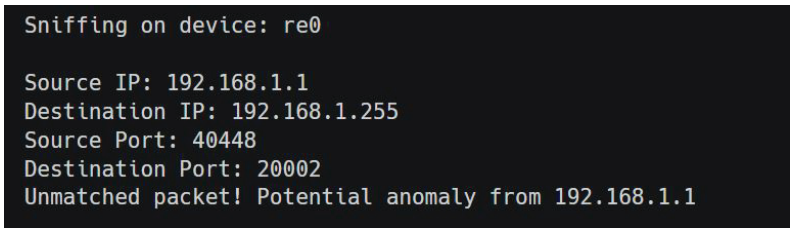


Figure 3. Detection of a potential anomaly in network traffic by the CHERI-enhanced IDS, flagging an unmatched packet from source IP 192.168.1.1 as a potential spoofing attempt.

5.1. IDS Rule Integrity and Unauthorized Access Blocking

During the IDS rule manipulation attempt, the attacker sought to alter the IDS configuration to disable specific detection rules.

- **Memory Access Violation Logs**: The CHERI-enabled IDS logged multiple unauthorized access attempts by the attacker as they tried to modify the IDS rules. Each access violation was flagged by CHERI's capability checks, which block any process outside the authorized scope from modifying memory. This illustrates CHERI's effectiveness in maintaining the integrity of the IDS rule set, as all modification attempts were successfully prevented as shown in Figure 4.

- **Rule Integrity Preservation**: The IDS maintained its rule configurations intact throughout the attack scenario, as evidenced by the lack of any successful modifications to the IDS rules. CHERI's memory protection mechanisms effectively isolated the IDS configuration, preventing the attacker from altering detection logic. This result demonstrates the value of CHERI capabilities in protecting critical security configurations, ensuring consistent IDS performance even under attack.

```
root@cheribsd:/home/sampath/IDS # ./Attacker
In-address space security exception (core dumped)
root@cheribsd:/home/sampath/IDS # ./Attacker
In-address space security exception (core dumped)
root@cheribsd:/home/sampath/IDS #
```

Figure 4. Execution of the Attacker program resulting in in-address space security exceptions, demonstrating the enforcement of memory safety features by CheriBSD (core dump generated).

5.2. Analysis of CHERI's Impact on IDS Security

The experimental results indicate that CHERI's fine-grained memory protections significantly enhance IDS security in the following ways:

- **Increased Resilience to Spoofing Attacks**: CHERI's hardware-enforced memory compartmentalization makes the IDS resilient to IP-based spoofing attacks. Although the attacker could mimic legitimate IPs, CHERI's memory protection mechanisms prevented unauthorized influence on the IDS logic.
- **Enhanced Rule Protection Against Manipulation Attempts**: The compartmentalized memory model prevented the attacker from modifying IDS rules, effectively securing the IDS from rule manipulation attacks. The unauthorized access logs demonstrate that CHERI successfully restricted all access attempts to the IDS configuration, ensuring rule integrity.
- **Implications for Real-Time Security Applications**: The minimal performance overhead observed in this experiment suggests that CHERI is suitable for real-time security applications in automotive contexts, where high-speed detection and low latency are critical.

The analysis of these results supports the conclusion that CHERI-based memory protection is an effective method for bolstering IDS security against IP spoofing and rule manipulation in automotive networks. To address sophisticated manipulation methods, such as real-time modifications of network traffic, the proposed system leverages several mechanisms:

1. **Hardware-Enforced Memory Isolation**: CHERI's capability-based architecture ensures that IDS rules and configurations are stored in strictly isolated memory regions. This isolation prevents unauthorized access, even if attackers employ advanced methods like spoofing or real-time packet alterations. CHERI's memory bounds enforce deterministic hardware traps when unauthorized attempts are made to access or modify IDS rule sets.
2. **Dynamic Anomaly Detection**: The IDS processes incoming SOME/IP packets in real time, validating attributes like source IP, payload structure, and session identifiers against CHERI-protected rule sets. CHERI ensures that, even if attackers manipulate live network traffic to mimic legitimate patterns, any deviation from the CHERI-enforced rules triggers alerts, ensuring the integrity of the detection mechanism.
3. **Real-Time Logging and Response**: CHERI capabilities enable the real-time logging of unauthorized memory access attempts. These logs allow the IDS to adapt dynamically

by flagging patterns of sophisticated manipulation, such as timing inconsistencies or malformed packets, enhancing its ability to counter evolving attack vectors.

To evaluate this, we simulated three types of ECU packet data based on the AUTOSAR standard: telemetry ECU, control signal ECU, and sensor data ECU. These simulations involved injecting spoofed IP packets and attempting to access and modify IDS rules. Each IDS rule was defined within distinct CHERI-protected memory regions corresponding to the specific ECU type. Despite multiple attempts by the attacker to exploit IP spoofing and alter IDS configurations, the CHERI-enforced memory boundaries successfully isolated these memory regions. As a result, all unauthorized access attempts triggered deterministic traps, and the attacker's script terminated with a core dump, indicating that the attack was completely mitigated. By leveraging CHERI's hardware-enforced protections in tandem with dynamic detection capabilities, the proposed IDS effectively counters real-time manipulation techniques, ensuring robust defense for automotive networks.

By enforcing strict memory access boundaries, CHERI not only prevents attackers from bypassing IDS detection mechanisms but also ensures the integrity of IDS rules. These results demonstrate CHERI's robust capability to safeguard against advanced threats, highlighting its significant potential for broader applications in automotive cybersecurity.

5.3. Simulated Scenarios and Scalability Analysis

In this study, we evaluated the performance and scalability of the CHERI-enhanced IDS by simulating three types of ECU packet data based on the AUTOSAR standard: **telemetry ECU**, **control signal ECU**, and **sensor data ECU**. Each ECU type transmitted packets at rates conforming to standard automotive communication protocols. Initially, the simulation included 15 ECUs, with both benign and malicious packets. The attackers employed IP spoofing to impersonate legitimate ECUs and attempted to manipulate IDS rule configurations stored in protected memory.

To assess scalability, the number of simulated ECUs was increased tenfold, from 15 to 150. Each ECU transmitted packets at rates aligned with AUTOSAR specifications:

- **Telemetry ECU**: Transmitted data packets at a rate of 12.5 packets per second.
- **Control Signal ECU**: Maintained a rate of 10 packets per second for command updates.
- **Sensor Data ECU**: Generated sensor readings at 25 packets per second.

Despite the substantial increase in packet volume, the CHERI-enhanced IDS maintained its performance metrics:

- **Detection Accuracy**: Consistently achieved 100% detection of spoofed packets across all scenarios.
- **Latency**: Average response time remained at 12 ms, only a marginal increase compared to the baseline IDS without CHERI, which exhibited a latency of 10 ms.

The implementation of CHERI in IDS systems does introduce certain performance trade-offs, but they are minimal and manageable, as demonstrated by our experimental results. These trade-offs are critical to balancing security with real-time operational requirements in automotive environments. Key findings include the following:

- **Latency Considerations**: Our study observed an average latency increase of 2 ms (from 10 ms to 12 ms) compared to a baseline IDS without CHERI. This marginal overhead, caused by the additional memory bound checks enforced by CHERI, remains well within the acceptable limits for real-time automotive systems. For instance, in automatic braking systems, where the latency threshold is typically below 50 ms, the additional 2 ms does not compromise responsiveness. This ensures that the enhanced security provided by CHERI does not hinder the performance of critical

safety mechanisms. Future work will aim to optimize CHERI's integration into more resource-constrained environments, further minimizing latency to meet the stringent requirements of ultra-fast vehicular control systems.

- **Memory Overheads**: The fine-grained memory protection mechanisms of CHERI require additional memory for storing metadata (capabilities). While this increases memory consumption, the impact was minimized by the efficient memory management features of the ARM Morello board. This trade-off is justified by the significant security benefits gained, including robust protection against memory corruption and unauthorized access.
- **Scalability in High-Throughput Scenarios**: During simulations with increased ECU counts and packet volumes, the CHERI-enhanced IDS maintained consistent performance. Even under high packet rates, the system demonstrated 100% detection accuracy with negligible degradation in processing times, proving its scalability for large-scale automotive networks.
- **Real-Time Packet Processing**: CHERI's capability-based protections allow the IDS to process SOME/IP packets dynamically, ensuring strict adherence to rule configurations. This ensures that security is maintained without significant delays, which is vital for time-sensitive operations like braking and collision avoidance.
- **Hardware Costs and Automotive Integration**: While adopting CHERI requires CHERI-enabled hardware, such as the ARM Morello board, these costs are offset by the long-term benefits of hardware-enforced security. However, it is important to note that current experiments are conducted on research-grade hardware, and further work is needed to deploy CHERI capabilities on automotive-grade devices, such as the SONATA board or equivalent platforms. This step will ensure real-world feasibility, compliance with industry standards, and suitability for production environments in the automotive sector.

In summary, the trade-offs associated with CHERI adoption are outweighed by its advantages, such as enhanced rule integrity, increased resilience against spoofing and manipulation, and real-time security assurance. By aligning with automotive standards like ISO/SAE 21434, CHERI-based solutions present a viable path toward secure, scalable, and high-performance IDS systems for the connected vehicle ecosystem.

6. Conclusions

The results highlight the effectiveness of CHERI capabilities in securing automotive networks against advanced threats such as IP spoofing and rule manipulation. CHERI's fine-grained memory protection mechanisms, particularly suited for systems developed in C/C++, enhance resilience while preserving performance, a critical requirement in automotive environments. C/C++ remain the dominant programming languages in the automotive domain due to their efficiency and close-to-hardware capabilities, making CHERI's compatibility with these languages a significant advantage for adoption in vehicle systems.

Although the results are promising, implementing CHERI in automotive systems requires careful consideration of performance trade-offs, hardware costs, and integration complexity. Future research should focus on optimizing CHERI configurations for resource-constrained environments and extending its applications to other critical automotive components beyond IDS. This study provides a practical framework for implementing CHERI-based IDS in real-world automotive environments. The results confirm the feasibility of deploying CHERI on industry-standard platforms, such as the ARM Morello board, ensuring compatibility with existing automotive communication protocols like SOME/IP. Testing under diverse cyber–physical attack scenarios and aligning CHERI-based solutions

with automotive standards, such as ISO/SAE 21434, would further validate its role in secure automotive design. The proposed IDS demonstrates scalability by efficiently handling increased network traffic volumes and ECU configurations without significant latency or resource overhead, making it viable for large-scale automotive systems. Additionally, deploying CHERI on lightweight, automotive-friendly platforms like the SONATA board could demonstrate its feasibility in real-world systems. Future work should also explore integrating machine learning-based anomaly detection into CHERI-enabled IDS to address advanced threats, including massive DoS attacks and compromised ECUs. Combining machine learning models with CHERI's memory safety can enhance detection accuracy, isolate compromised components, and improve resilience against evolving cyber threats in automotive networks.

This study underscores that CHERI-based memory protection significantly enhances IDS resilience against IP spoofing and rule manipulation attacks. By leveraging CHERI's fine-grained memory control, this work showcases a hardware-based security solution that maintains IDS rule integrity under attack. By addressing critical vulnerabilities, such as spoofed packets and unauthorized configuration changes, CHERI enhances the real-time security of connected and autonomous vehicles. The findings further highlight CHERI's alignment with industry standards, providing a robust foundation for secure automotive design. The findings suggest that adopting CHERI could strengthen the cybersecurity of critical vehicle control systems while seamlessly integrating with the C/C++-centric automotive software ecosystem.

Author Contributions: Conceptualization, C.S.K. and S.M.; methodology, C.S.K.; software, C.S.K.; validation, C.S.K., S.M., X.L. and C.C.; formal analysis, C.S.K.; investigation, C.S.K.; resources, C.S.K. and S.M.; data curation, C.S.K.; writing—original draft preparation, C.S.K writing—review and editing, S.M., X.L. and C.C.; visualization, C.S.K.; supervision, S.M., X.L. and C.C.; project administration, S.M.; funding acquisition, C.S.K. and S.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by DSDB cohort 5.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original data presented in this study are openly available on the first PoC's GitHub repository (https://github.com/sampathkcs/DSDB, accessed on 1 October 2024). The extended version of the experimental data is available upon request from the corresponding author, as the extended project is being developed in collaboration with an industry partner to increase TR levels.

Conflicts of Interest: Authors Chathuranga Sampath Kalutharage and Saket Mohan were employed by the company Secure Elements. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AUTOSAR	Automotive Open System Architecture
CAN	Controller Area Network
CAN-FD	Controller Area Network with Flexible Data Rate
CEP	Complex Event Processing
CHERI	Capability Hardware Enhanced RISC Instructions
DDoS	Distributed Denial of Service
DPI	Deep Packet Inspection

ECU	Electronic Control Unit
IDS	Intrusion Detection System
ISO	International Organization for Standardization
MITM	Man-in-the-Middle
OTA	Over-the-Air
PCAP	Packet Capture
SOME/IP	Scalable service-Oriented Middleware over IP
TESLA	Timed Efficient Stream Loss-tolerant Authentication

References

1. Luo, F.; Zhang, X.; Hou, S. Research on cybersecurity testing for in-vehicle network. In Proceedings of the 2021 International Conference on Intelligent Technology and Embedded Systems (ICITES), Chengdu, China, 31 October–2 November 2021; pp. 144–150.
2. Abdelkader, G.; Elgazzar, K.; Khamis, A. Connected vehicles: Technology review, state of the art, challenges and opportunities. *Sensors* **2021**, *21*, 7712. [CrossRef] [PubMed]
3. Kabilan, N.; Ravi, V.; Sowmya, V. Unsupervised intrusion detection system for in-vehicle communication networks. *J. Saf. Sci. Resil.* **2024**, *5*, 119–129.
4. Almehdhar, M.; Albaseer, A.; Khan, M.A.; Abdallah, M.; Menouar, H.; Al-Kuwari, S.; Al-Fuqaha, A. Deep learning in the fast lane: A survey on advanced intrusion detection systems for intelligent vehicle networks. *IEEE Open J. Veh. Technol.* **2024**, *5*, 869–906. [CrossRef]
5. Meneghello, F.; Calore, M.; Zucchetto, D.; Polese, M.; Zanella, A. IoT: Internet of threats? A survey of practical security vulnerabilities in real IoT devices. *IEEE Internet Things J.* **2019**, *6*, 8182–8201. [CrossRef]
6. Al-Jarrah, O.Y.; Maple, C.; Dianati, M.; Oxtoby, D.; Mouzakitis, A. Intrusion detection systems for intra-vehicle networks: A review. *IEEE Access* **2019**, *7*, 21266–21289. [CrossRef]
7. Zuech, R.; Khoshgoftaar, T.M.; Wald, R. Intrusion detection and big heterogeneous data: A survey. *J. Big Data* **2015**, *2*, 3. [CrossRef]
8. Hamada, Y.; Inoue, M.; Adachi, N.; Ueda, H.; Miyashita, Y.; Hata, Y. Intrusion detection system for in-vehicle networks. *SEI Tech. Rev.* **2019**, *88*, 76–81.
9. Samaila, M.G.; Neto, M.; Fernandes, D.A.B.; Freire, M.M.; Inácio, P.R.M. Challenges of Securing Internet of Things Devices: A Survey. *Secur. Priv.* **2018**, *1*, e20. [CrossRef]
10. Greenberg, A. *Millions of Vehicles Could Be Hacked and Tracked Thanks to a Simple Website Bug*; Wired: London, UK, 2021.
11. Connected Cars: On The Edge of a Breakthrough. Available online: https://aecc.org/resources/publications/ (accessed on 2 December 2024).
12. Woodruff, J.D. *CHERI: A RISC Capability Machine for Practical Memory Safety*; Technical Report UCAM-CL-TR-858; University of Cambridge, Computer Laboratory: Cambridge, UK, 2014.
13. University of Cambridge. The CHERI Project. Available online: https://www.cl.cam.ac.uk/research/security/ctsrd/cheri/ (accessed on 2 December 2024).
14. Davis, B.; Watson, R.N.; Richardson, A.; Neumann, P.G.; Moore, S.W.; Baldwin, J.; Chisnall, D.; Clarke, J.; Filardo, N.W.; Gudka, K.; et al. CheriABI: Enforcing valid pointer provenance and minimizing pointer privilege in the POSIX C run-time environment. In Proceedings of the Twenty-Fourth International Conference on Architectural Support for Programming Languages and Operating Systems, Providence, RI, USA, 13–17 April 2019; pp. 379–393.
15. Embedded.com. Why C/C++ Dominate in Automotive Systems. Available online: https://www.embedded.com/ (accessed on 2 December 2024).
16. ARM. Introducing the Morello Prototype Board. Available online: https://www.arm.com/morello (accessed on 2 December 2024).
17. *ISO/SAE 21434*; Road Vehicles—Cybersecurity Engineering. International Standards Organization: Geneva, Switzerland, 2020. https://www.iso.org/standard/70918.html (accessed on 2 December 2024).
18. Ibrahim, M.E.; Abbas, Q.; Daadaa, Y.; Ahmed, A.E. A Novel PPG-Based Biometric Authentication System Using a Hybrid CVT-ConvMixer Architecture with Dense and Self-Attention Layers. *Sensors* **2024**, *24*, 15. [CrossRef] [PubMed]
19. Saltzer, J.H. Protection and the control of information sharing in Multics. *Commun. ACM* **1974**, *17*, 388–402. [CrossRef]
20. Woodruff, J.; Watson, R.N.M.; Chisnall, D.; Moore, S.W.; Anderson, J.; Davis, B.; Laurie, B.; Neumann, P.G.; Norton, R.; Roe, M. The CHERI Capability Model: Revisiting RISC in an Age of Risk. In Proceedings of the 41st Annual International Symposium on Computer Architecture (ISCA), Minneapolis, MN, USA, 14–18 June 2014; pp. 457–468. [CrossRef]

21. Kho, N.M.D.; Uy, R.L. MIPSers: MIPS extension release 6 simulator. In Proceedings of the 2017IEEE 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM), Manila, Philippines, 1–3 December 2017; pp. 1–6.
22. Grisenthwaite, R.; Barnes, G.; Watson, R.N.; Moore, S.W.; Sewell, P.; Woodruff, J. The Arm Morello Evaluation Platform—Validating CHERI-based Security in a High-performance System. *IEEE Micro* **2023**, *43*, 50–57. Available online: https://www.cl.cam.ac.uk/research/security/ctsrd/pdfs/202305ieeemicro-morello-platform.pdf (accessed on 2 December 2024). [CrossRef]
23. Aliwa, E.; Rana, O.; Perera, C.; Burnap, P. Cyberattacks and countermeasures for in-vehicle networks. *ACM Comput. Surv. (CSUR)* **2021**, *54*, 21. [CrossRef]
24. Hafeez, A.; Topolovec, K.; Awad, S. ECU fingerprinting through parametric signal modeling and artificial neural networks for in-vehicle security against spoofing attacks. In Proceedings of the 2019 15th International Computer Engineering Conference (ICENCO), Cairo, Egypt, 29–30 December 2019; pp. 29–38.
25. Song, S.; Manikopoulos, C.N. IP Spoofing Detection Approach (ISDA) for Network Intrusion Detection System. In Proceedings of the 2006 IEEE Sarnoff Symposium, Princeton, NJ, USA, 27–28 March 2006; pp. 1–4. [CrossRef]
26. Parameshwarappa, P.; Chen, Z.; Gangopadhyay, A. Analyzing attack strategies against rule-based intrusion detection systems. In Proceedings of the Workshop Program of the 19th International Conference on Distributed Computing and Networking, Varanasi, India, 4–7 January 2018; pp. 1–4.
27. Zorman, E.H.; Isoaho, J.; Mohammad, T. Implementation of a SOME/IP Firewall with Deep Packet Inspection for Automotive Use-Cases. Ph.D. Thesis, University of Turku, Turku, Finland, 2024.
28. Ghadi, M.; Sali, Á.; Szalay, Z.; Török, Á. A new methodology for analyzing vehicle network topologies for critical hacking. *J. Ambient Intell. Humaniz. Comput.* **2021**, *12*, 7923–7934. [CrossRef]
29. AUTOSAR. *Specification of Module Secure Onboard Communication—CP*; Technical Report, Release 4.2.2; AUTOSAR: Hörgertshausen, Germany, 2017.
30. Lee, T.Y.; Lin, I.A.; Liao, R.H. Design of a FlexRay/Ethernet gateway and security mechanism for in-vehicle networks. *Sensors* **2020**, *20*, 641. [CrossRef]
31. Islam, R.; Refat, R.U.D. Improving CAN bus security by assigning dynamic arbitration IDs. *J. Transp. Secur.* **2020**, *13*, 19–31. [CrossRef]
32. Hafeez, A.; Malik, H.; Irtaza, A.; Uddin, M.Z.; Noori, F.M. Enhancing ECU identification security in CAN networks using distortion modeling and neural networks. *Front. Comput. Sci.* **2024**, *6*, 1392119. [CrossRef]
33. Woo, S.; Jo, H.J.; Kim, I.S.; Lee, D.H. A practical security architecture for in-vehicle CAN-FD. *IEEE Trans. Intell. Transp. Syst.* **2016**, *17*, 2248–2261. [CrossRef]
34. Lodge, N.; Tambe, N.; Saqib, F. Addressing Vulnerabilities in CAN-FD: An Exploration and Security Enhancement Approach. *IoT* **2024**, *5*, 290–310. [CrossRef]
35. Kreissl, J. Absicherung der Some/IP Kommunikation bei Adaptive Autosar. Master's Thesis, Universität Stuttgart, Stuttgart, Germany, 2017.
36. Iorio, M.; Reineri, M.; Risso, F.; Sisto, R.; Valenza, F. Securing SOME/IP for in-vehicle service protection. *IEEE Trans. Veh. Technol.* **2020**, *69*, 13450–13466. [CrossRef]
37. Herold, N.; Posselt, S.A.; Hanka, O.; Carle, G. Anomaly detection for SOME/IP using complex event processing. In Proceedings of the NOMS 2016—2016 IEEE/IFIP Network Operations and Management Symposium, Istanbul, Turkey, 25–29 April 2016; pp. 1221–1226.
38. Liu, Q.; Li, X.; Sun, K.; Li, Y.; Liu, Y. SISSA: Real-time Monitoring of Hardware Functional Safety and Cybersecurity with In-vehicle SOME/IP Ethernet Traffic. *IEEE Internet Things J.* **2024**, *11*, 27322–27339. [CrossRef]
39. McKeen, F.; Alexandrovich, I.; Berenzon, A.; Rozas, C.; Shafi, H.; Shanbhogue, V.; Savagaonkar, U. Innovative instructions and software model for isolated execution. In Proceedings of the 2nd International Workshop on Hardware and Architectural Support for Security and Privacy, Tel-Aviv, Israel, 23–24 June 2013; pp. 1–8.
40. ARM. Building a Secure System Using TrustZone Technology. 2019. Available online: https://developer.arm.com/documentation (accessed on 2 December 2024).
41. Kumar, A.; Gholve, A.; Kotalwar, K. *Automotive Security Solution Using Hardware Security Module (HSM)*; Technical Report, SAE Technical Paper; SAE International: Warrendale PA, USA, 2024.
42. Saltzer, J.H.; Schroeder, M.D. The protection of information in computer systems. *Proc. IEEE* **1975**, *63*, 1278–1308. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

electronics

MDPI

Article

Lazy Modular Reduction for NTT

Geumtae Kim [1], Eunyoung Seo [2,*], Yongwoo Lee [3,*], Young-Sik Kim [2] and Jong-Seon No [1]

[1] The Department of Electrical and Computer Engineering, Institute of New Media and Communications (INMC), Seoul National University, Seoul 08826, Republic of Korea; kkt1513@snu.ac.kr (G.K.); jsno@snu.ac.kr (J.-S.N.)
[2] The Department of Electrical Engineering and Computer Science, Daegu Gyeongbuk Institute of Science and Technology, Daegu 42988, Republic of Korea; ysk@dgist.ac.kr
[3] The Department of Electrical and Computer Engineering, Inha University, Incheon 22212, Republic of Korea
* Correspondence: eyseo_dgist@dgist.ac.kr (E.S.); yongwoo@inha.ac.kr (Y.L.)

Abstract: The number theoretic transform (NTT) is a fundamental operation in cryptography, especially for lattice-based cryptographic schemes. This paper introduces *LazyNTT*, a novel method that reduces the number of Montgomery multiplications required in the NTT computation by replacing some of them with standard multiplication without modular reduction. This approach enhances the performance of the NTT computation and modular polynomial multiplication in lattice-based cryptographic schemes. The proposed *LazyNTT* can be generalized by increasing the number of standard multiplications. The experimental results show that the proposed *LazyNTT* improves the cycle counts of the NTT by up to 28% and 9%, respectively, by allowing two and one standard multiplications.

Keywords: number theoretic transform (NTT); Montgomery multiplication; modular reduction; post-quantum cryptography (PQC); lattice-based cryptography

1. Introduction

The advent of quantum computers breaks classical public key cryptosystems such as RSA [1] using Shor's algorithm [2]. As a result, the development of post-quantum cryptography (PQC) has become crucial. The National Institute of Standards and Technology (NIST) has initiated a standardization process for PQC [3], with many candidates based on lattice cryptography.

Lattice-based cryptography heavily relies on modular arithmetic, particularly modular polynomial multiplication, which constitutes a significant computational bottleneck. The number theoretic transform (NTT) reduces the complexity of modular polynomial multiplication from $O(n^2)$ to $O(n \log n)$ by leveraging component-wise multiplication. Despite this, modular reduction within the NTT remains computationally expensive.

This paper proposes *LazyNTT*, a method designed to reduce the number of Montgomery multiplications [4] in the NTT computation, which is an efficient workaround for naive modular reduction but is still more expensive than the standard integer multiplication. By replacing some of them with standard multiplications at intermediate stages of the NTT computation, we achieve significant computational complexity improvements.

1.1. Contributions

In this paper, we propose a new variant of the NTT with a faster runtime by reducing the number of Montgomery multiplications. The detailed contributions are given as follows:

1. We propose two versions of *LazyNTT* that reduce the number of Montgomery multiplications in the NTT computation by replacing some of them with standard multiplications during the intermediate stages of the NTT.

2. We demonstrate the effectiveness of the proposed algorithm in practical parameters used in Falcon [5] and Kyber [6], showing that this method can significantly reduce the total cycle counts of the NTT computation.
3. The proposed method is applicable to various lattice-based cryptographic schemes, including PQC [7–11], and it can naturally be generalized for the further replacement of Montgomery multiplications.

1.2. Related Works

Extensive research has been conducted on the NTT due to its critical role in cryptography. In the hardware domain, efficient implementations of the NTT on FPGA have been presented in [12,13]. They optimized memory access patterns and minimized the latency of the NTT computation. Additionally, the performance improvement of polynomial multiplication using the NTT by employing a unified butterfly unit has been presented in [14]. In the software domain, performance improvements have been achieved using the Harvey butterfly structure in the radix-4 NTT [15] and optimizing cyclic convolution operations to speed up the NTT [16]. In the context of PQC, efficient NTT implementations using a hardware accelerator and instruction set for Kyber [7,8], parallelism with FPGA for Dilithium [9,10], and twiddle factor generation for Falcon [11,17] have been presented. In the field of homomorphic encryption (HE), some have focused on efficiently implementing the NTT using GPUs [18–20]. Specialized hardware designs to optimize the NTT for HE also have been proposed in [21,22]. Previous works are summarized in Table 1 for better clarity.

Table 1. Related works on number theoretic transform.

Platform	Papers	Cryptosystems	Technique
FPGA	[13]	HE	Parallelization, automatic NTT design
	[21,22]		On-the-fly twiddle factor generation
	[12]	Kyber	Double-bandwidth ping-pong memory access
	[14]	PQC	Unified butterfly unit
GPU	[18]	HE	Optimized memory access patterns
	[19]		Multi-stream asynchronous computation
	[20]		On-the-fly twiddle factor generation
CPU	[15]	HE	Radix-4 Harvey butterfly
	[16]	PQC	Modified cyclic convolution

Our work builds on these advancements by specifically addressing the computational load of modular reduction within the NTT. To the best of our knowledge, a direction focusing on replacing only the type of multiplications used in the NTT has not been explored before. Additionally, previous works on the NTT have been developed across various platforms and cryptosystems, making it difficult to compare them under the same criterion. Therefore, we use the official implementation code of Falcon as a reference, which is selected as the PQC standard by NIST. We believe that comparing computational complexity with this reference code is appropriate for demonstrating the meaningful results of *LazyNTT*. By replacing some Montgomery multiplications with standard multiplications without modular reduction, we achieve a notable computational complexity gain, which complements existing optimization in both hardware and software implementations.

1.3. Organization

The structure of this paper is as follows: Section 2 introduces the basic notations and preliminaries needed for understanding Montgomery multiplication and the NTT. It also includes a detailed explanation of the radix-2 NTT, which is the primary focus of

the proposed method. Section 3 describes the proposed Standard–Montgomery version (*SM-LazyNTT*) and Standard–Standard–Montgomery version (*SSM-LazyNTT*) algorithms. Their designs, operational principles, and generalization are also included in this section. Section 4 provides the implementation results and an analysis of the improved computational complexity achieved by *SM-LazyNTT* and *SSM-LazyNTT* compared to the original NTT. Section 5 concludes the paper with a summary of the proposed method and suggestions for future works.

2. Preliminaries

2.1. Notations

A ciphertext modulus is denoted as q, and n represents both the length of the polynomial and the NTT. Throughout this paper, we assume that n is a power of two. For a polynomial $f(x)$, let f_i denote the i-th coefficient of the polynomial. Vectors are indicated as bold letters, for example $\mathbf{f}$, and it represents a vector containing all the coefficients of $f(x)$, i.e., $\mathbf{f} = (f_0, f_1, \cdots, f_{n-1})$. We denote the j-th component of the vector $\mathbf{f}$ as $\mathbf{f}[j]$. We use $\overline{a}$ to indicate a value in Montgomery space and similarly use $\overline{\mathbf{a}}$ to denote a vector containing values in Montgomery space. The NTT of the polynomial $f(x)$ is denoted as $\hat{\mathbf{f}}$. R is a scaling factor used in Montgomery multiplication. The integer ring is denoted as $\mathbb{Z}_q$ and the polynomial ring with the modulus is $\mathbb{Z}_q[x]$. For a polynomial $\phi(x)$, we define the quotient ring $\mathbb{Z}_q[x]/\phi(x)$. Standard multiplication is denoted by $\cdot$ and Montgomery multiplication is denoted by $*$. Component-wise multiplication between two vectors, $\hat{\mathbf{f}}$ and $\hat{\mathbf{g}}$, is denoted as $\hat{\mathbf{f}} \circ \hat{\mathbf{g}}$, which means that each element in the same index is multiplied as $\hat{f}_i \cdot \hat{g}_i$. When the context is clear, we may omit the modular reduction notation, mod q. Let ω be a primitive n-th root of unity, and let α be a primitive $2n$-th root of unity in $\mathbb{Z}_q$, such that

$$\omega^n \equiv 1 \mod q \quad \text{and} \quad \alpha^{2n} \equiv 1 \mod q,$$

with $\omega^i \not\equiv 1 \mod q$ and $\alpha^j \not\equiv 1 \mod q$ for any integers $0 < i < n$ and $0 < j < 2n$, respectively.

2.2. Montgomery Multiplication

In integer modular arithmetic, while there are simple constant-time methods for addition and subtraction, multiplication is more complex due to the division by the modulus q required in the modular reduction step. Montgomery multiplication [4] is a widely used technique in cryptography for efficient modular multiplication. It allows the modulo multiplication of integers without directly performing division by the modulus, which is computationally expensive. Instead, it first transforms the numbers into a special domain called Montgomery space by multiplying them with a constant, R, and taking the result modulo q as

$$\overline{a} = a \cdot R \mod q.$$

This transformation enables more efficient modular arithmetic operations by replacing the division by q with cheaper operations such as multiplication, addition, and bit-shifting.

An element, a, in $\mathbb{Z}_q$ is mapped to Montgomery space as $\overline{a} = a \cdot R \mod q$. Addition and subtraction in Montgomery space are defined as in $\mathbb{Z}_q$, such that $\overline{a} \pm \overline{b} = a \cdot R \pm b \cdot R = (a \pm b) \cdot R \mod q$. However, multiplication doubles the factor R, and thus Montgomery multiplication $(*)$ is defined as follows:

$$a * b = (a \cdot b \cdot R^2) \cdot R^{-1} \mod q = (a \cdot b) \cdot R \mod q.$$

To implement the steps of multiplying R^{-1} and reducing modulo q efficiently, we use the extended Euclidean algorithm. Algorithm 1 describes the whole procedures. Unlike Montgomery multiplication, which multiplies two values in Montgomery space, we use the standard multiplication in this paper to refer to the multiplication between one value in Montgomery space and the other value in standard space.

Algorithm 1 Montgomery Multiplication

Require: $\overline{a}, \overline{b}, R, q$
Ensure: $\overline{a} \cdot \overline{b} \cdot R^{-1} \mod q$
Pre-computation:
$(R', q') \leftarrow$ Extended Euclidean(R, q)
$R' \leftarrow R' \mod q$ ▷ $RR' \equiv 1 \mod q$
$q' \leftarrow q' \mod R$ ▷ $qq' \equiv 1 \mod R$
Montgomery Multiplication:
$x \leftarrow \overline{a} \cdot \overline{b}$
$m \leftarrow (x \cdot q') \mod R$
$u \leftarrow (x - m \cdot q)/R$
if $u \leq 0$ **then**
    $u \leftarrow u + q$
end if
return u

2.3. Number Theoretic Transform

The NTT is a specialized form of the discrete Fourier transform (DFT) that operates in an integer ring. Unlike the DFT, which handles complex numbers, the NTT utilizes integer modular arithmetic, making it suitable for error-free and efficient computations in cryptography and polynomial multiplication. For efficient NTT computation, a structure analogous to the fast Fourier transform (FFT) [23] can be employed.

Let $\phi(x)$ be the n-th cyclotomic polynomial, whose roots are the primitive n-th roots of unity, and $\phi_1(x) = x - 1$. The NTT can be defined as a ring homomorphism that maps a polynomial in $\mathbb{Z}_q[x]/\phi(x)$ to $(\mathbb{Z}_q[x]/\phi_1(x))^n$. Here, the product of the quotient ring $(\mathbb{Z}_q[x]/\phi_1(x))^n$ can be considered an integer vector $(\mathbb{Z}_q)^n$, which is the vector of evaluations at the powers of a primitive root of unity.

Considering a polynomial multiplication over the quotient ring $\mathbb{Z}_q[x]/(x^n + 1)$ for two polynomials,

$$f(x) = f_0 + f_1 x + \cdots + f_{n-1} x^{n-1}$$

and

$$g(x) = g_0 + g_1 x + \cdots + g_{n-1} x^{n-1},$$

the result $h(x) = (f(x) \times g(x))/(x^n + 1) \in \mathbb{Z}_q[x]/(x^n + 1)$ has coefficients such that

$$h_k = \sum_{i=0}^{k} f_i g_{k-i} - \sum_{i=k+1}^{n-1} f_i g_{n+k-i} \mod q.$$

This requires $O(n^2)$ multiplication complexity using a direct convolution approach. However, modular polynomial multiplication using the NTT achieves $O(n \log n)$ complexity with the following steps:

(1) Calculating the NTT of two polynomials.

$$\hat{\mathbf{f}} = NTT(f(x))$$

$$\hat{\mathbf{g}} = NTT(g(x))$$

(2) The component-wise multiplication of two NTTs.

$$\hat{\mathbf{h}} = \hat{\mathbf{f}} \circ \hat{\mathbf{g}}$$

(3) Calculating the inverse NTT of the result.

$$h(x) = iNTT(\hat{\mathbf{h}})$$

For the NTT, using $\phi(x) = x^n \pm 1$ enhances efficiency due to the presence of the roots of unity $\omega^n \equiv 1 \mod q$ for $x^n - 1$ and $\alpha^{2n} \equiv 1 \mod q$ for $x^n + 1$. These roots of unity enable a structured and cyclic approach to polynomial multiplication, allowing the use of FFT structure that breaks down the problem into smaller subproblems, leading to $O(n \log n)$ complexity. Moreover, the cyclic nature imposed by $x^n \pm 1$ allows for efficient modular arithmetic operations, ensuring that the polynomial degree is maintained within n, thereby simplifying modular reduction steps. For the quotient rings $\mathbb{Z}_q[x]/(x^n - 1)$ and $\mathbb{Z}_q[x]/(x^n + 1)$, the former is known as positive wrapped convolution (PWC) [24] and the latter as negative wrapped convolution (NWC) [25].

In the context of PQC, the NTT usually refers to the NWC version. For example, in Falcon and Dilithium, $\phi(x) = x^n + 1$ is used for the NTT. To perform polynomial multiplication using the NTT, the modulus is chosen to ensure that the primitive $2n$-th root of unity exists in $\mathbb{Z}_q$ such that $q \equiv 1 \mod 2n$. Two examples are as follows:

- Falcon uses $q = 12{,}289 = 6 \cdot 2048 + 1$ for $n = 1024$ or 512 [5].
- Dilithium uses $q = 8{,}380{,}417 = 16{,}368 \cdot 512 + 1$ for $n = 256$ [26].

In Kyber, the parameters $q = 3329 = 13 \cdot 2^8 + 1$ and $n = 256$ are used, but a primitive 512-th root of unity does not exist in $\mathbb{Z}_{3329}$. Therefore, Kyber does not use the NWC-based NTT. Instead, during the NTT, the polynomial $f(x)$ is converted into 128 polynomials with a degree of 1 rather than 256 integers. Although this paper focuses on NWC, the proposed method can be naturally extended to the PWC environment and works similarly well.

The NTT using NWC transforms a polynomial $f(x) \in \mathbb{Z}_q[x]/(x^n + 1)$ to its evaluations at points $x = \alpha^{2i+1}$ for $i = 0, \ldots, n-1$. Since α^{2i} is also a n-th root of unity, we only consider α^{2i+1}, which is a primitive $2n$-th root of unity.

2.4. Radix-2 NTT Using FFT Structure

In the context of the NTT, radix-2 refers to a processing step, where two terms are handled together in a structure commonly known as a butterfly operation. By leveraging the divide-and-conquer strategy inherent in the FFT, the radix-2 NTT recursively splits the polynomial into smaller sub-NTTs, each half the size of the previous one, and applies the butterfly operation to combine the results. We will refer to this single iteration as a "stage". Within a stage, we refer to each spot according to the index as a "position". The butterfly operation involves the pairwise addition and subtraction of polynomial coefficients where one of them is multiplied by a power of the root of unity. This recursive structure ensures a significant reduction in computational complexity, making the transform efficient in both time and space, as in Algorithm 2.

In Algorithm 2, $\overline{\mathbf{G}}$ refers to the vector that stores all odd powers of the root of unity α^{2i+1} for $0 \leq i \leq n-1$, which are converted into Montgomery space. The first "for" loop (line 3) shows that there are a total of $\log n$ stages, with each iteration representing the computation of one stage. The second "for" loop (line 5) handles the computation of the m butterfly structure sets, with each structure using a different element of $\overline{\mathbf{G}}$ in that stage. The third "for" loop (line 7) processes h butterfly structures using the same element of $\overline{\mathbf{G}}$.

Generally, the radix-2 NTT is calculated using following steps.

1. Divide all coefficients into two sets.
2. Perform the NTT of each half.
3. Recursively divide the coefficients in each NTT until a 2-point NTT is reached.
4. Combine the results to compute the final n-point NTT.

This process is conducted recursively, until the 2-point NTT is obtained.

Algorithm 2 Radix-2 NTT

Require: $\bar{\mathbf{f}}$, $\overline{\mathbf{G}}$, $\log n$, q, and R
1: $n \leftarrow 2^{\log n}$
2: $t \leftarrow n$
3: **for** $m \leftarrow 1$ to n **do**
4: $\quad h \leftarrow t/2$
5: $\quad$ **for** $u \leftarrow 0$ to $m-1$ **do**
6: $\quad\quad v \leftarrow u \cdot t$
7: $\quad\quad$ **for** $v \leftarrow 0$ to $h-1$ **do**
8: $\quad\quad\quad \bar{r}_1 \leftarrow \bar{\mathbf{f}}[v]$
9: $\quad\quad\quad \bar{r}_2 \leftarrow \bar{\mathbf{f}}[v+h]$
10: $\quad\quad\quad \overline{x} \leftarrow \bar{r}_1$
11: $\quad\quad\quad \overline{y} \leftarrow \bar{r}_2 * \overline{\mathbf{G}}[m+u]$
12: $\quad\quad\quad \bar{r}_1 \leftarrow \overline{x} + \overline{y} \mod q$
13: $\quad\quad\quad \bar{r}_2 \leftarrow \overline{x} - \overline{y} \mod q$
14: $\quad\quad\quad v \leftarrow v+1$
15: $\quad\quad$ **end for**
16: $\quad$ **end for**
17: $\quad t \leftarrow h$
18: $\quad m \leftarrow 2 \cdot m$
19: **end for**

3. Lazy Modular Reduction for NTT

All multiplications during the NTT computation are carried out using Montgomery multiplication. Although Montgomery multiplication performs integer modular reduction efficiently, it still involves several multiplications, additions, and bit-shift operations. Notably, we observe that the NTT computation can proceed to subsequent stages without completing modular reduction in every intermediate stage. We propose a method, *LazyNTT*, where some of the Montgomery multiplications in intermediate stages of the NTT are replaced with standard multiplications without modular reduction. Although the results in intermediate stages are not reduced modulo, q, our goal is to ensure that the final NTT result is reduced modulo, q, not necessarily in every stage. Therefore, by increasing the value of the parameter R used in Montgomery multiplication, standard multiplications can replace some of the Montgomery multiplications required in the NTT computation. Since the algorithm occasionally skips modular reduction during the NTT computation, we refer to the proposed method as *LazyNTT*.

In this paper, we propose two versions of *LazyNTT*. The first version replaces one out of two Montgomery multiplications with a standard multiplication, using both in an alternating manner during the NTT computation. Based on the order of the standard (S) and Montgomery (M) multiplications used, we refer to this version as *SM-LazyNTT*. The second version replaces the first two of three Montgomery multiplications with standard multiplications. We similarly refer to this version as *SSM-LazyNTT*. Although this paper focuses only on the SM and SSM versions of *LazyNTT*, the approach can be easily generalized to replace more Montgomery multiplications. The generalization technique is described in Section 3.5.

3.1. SM-LazyNTT

A polynomial, $f(x)$, of the length $n = 2^k$ can be expressed in radix-2 form as below.

$$f(x) = \sum_{i_{k-1}=0}^{1} x^{i_{k-1}} \Big(\sum_{i_{k-2}=0}^{1} x^{2 \cdot i_{k-2}} \big(\cdots \big(\sum_{i_0=0}^{1} x^{2^{k-1} \cdot i_0} \cdot f_{2^{k-1} \cdot i_0 + \cdots + 2 \cdot i_{k-2} + i_{k-1}} \big) \big) \Big) \tag{1}$$

For the above polynomial $f(x)$, we can count the number of required multiplications as follows.

Remark 1. *A total of $k \cdot 2^{k-1}$ multiplications are required for the NTT computation of $f(x)$.*

1. From the radix-2 butterfly structure, it is clear that the number of multiplications required in each stage is exactly half of the length, 2^{k-1}.
2. For any $i \in [1, k]$, there are 2^{i-1} sub-NTTs in the ith stage, each of size 2^{k-i+1}. For each sub-NTT, a butterfly operation is applied to pairs of elements, which involves one multiplication by a power of the root of unity.
3. Since there are 2^{k-i+1} elements in each sub-NTT, the number of multiplications required per sub-NTT is 2^{k-i}.
4. Thus, $2^{i-1} \cdot 2^{k-i} = 2^{k-1}$ multiplications are required in the ith stage. In total, $k \cdot 2^{k-1}$ multiplications are required for the entire NTT computation with k stages.

In (1), the ith coefficient f_i is multiplied by $x^{2^{k-1} \cdot i_0}$ in the innermost bracket first, corresponding to the first stage. In the second stage, it is multiplied by $x^{2^{k-2} \cdot i_1}$, and this continues until it is finally multiplied by $x^{i_{k-1}}$. As we unfold the brackets through all k stages, the exponent of x accumulates to the value $2^{k-1} \cdot i_0 + 2^{k-2} \cdot i_1 + \cdots + i_{k-1}$, which is equal to the coefficient index i. This corresponds to the binary representation of i, indicating that each bit is from i_0 to i_{k-1}. Therefore, throughout the NTT computation, f_i encounters Montgomery multiplications with modular reduction when $i_l = 1$ for $0 \leq l \leq k-1$. On the other hand, when $i_l = 0$, there is no multiplication and no modular reduction.

These multiplications can be computed using the proposed *SM-LazyNTT* algorithm based on the following principles:

1. The last multiplication encountered in each index is always performed using a Montgomery multiplication.
2. The second-to-last multiplication, i.e., the one immediately preceding the Montgomery multiplication, is replaced with a standard multiplication.
3. The multiplication preceding the standard multiplication is again performed using a Montgomery multiplication, ensuring that standard and Montgomery multiplications alternate, with the multiplication sequence always ending in a Montgomery multiplication.

We can make the following claim, which is the condition for the replacement of a Montgomery multiplication with a standard multiplication in *SM-LazyNTT*.

Claim 1. *To replace a Montgomery multiplication with a standard multiplication, the condition $R > q^2$ should be satisfied, and there should be at least one Montgomery multiplication in the subsequent stages of the NTT computation.*

Proof. In the original Montgomery multiplication, R is set to be a power of two greater than the modulus q, performing modular reduction within every multiplication.

1. In Montgomery multiplication, the result of the multiplication of two values in $\mathbb{Z}_q$ should be less $R \cdot q$. As described in Section 2.2, this ensures the correct result by subsequent conditional addition. For *SM-LazyNTT*, one standard multiplication will make the value be within the range of $[0, q^2)$, since it is performed without modular reduction. Then, if this value is multiplied by another element in $\mathbb{Z}_q$ using Montgomery multiplication, the result should be less than $R \cdot q$. This leads to the condition that $q^3 < R \cdot q$, which implies that $R > q^2$.
2. After a Montgomery multiplication is replaced by a standard multiplication without modular reduction, there should be at least one Montgomery multiplication remaining in the subsequent stages to ensure that the result is properly reduced modulo, q. Clearly, the last multiplication of each coefficient throughout the NTT should be conducted using Montgomery multiplication to perform modular reduction in the end. At the point of the multiplication just before the last Montgomery multiplication, modular reduction can be skipped because following Montgomery multiplication involves modular reduction.

□

Also, the number of replaceable positions in *SM-LazyNTT* can be determined by the following claim.

Claim 2. *Out of a total of* $k \cdot 2^{k-1}$ *Montgomery multiplications,* $(k-1) \cdot 2^{k-2}$ *can be replaced with standard multiplications.*

Proof. From the Remark 1, there are 2^{k-1} multiplications for each stage.

1. The replaceable positions among all these Montgomery multiplications depend on whether there are any Montgomery multiplications in the subsequent stages.
2. We count the number of replaceable multiplications starting from the final stage. In the final kth stage, all 2^{k-1} multiplications are performed using Montgomery multiplication to ensure that the multiplication results are correctly reduced modulo, q. In the preceding $(k-1)$th stage, half of the Montgomery multiplications (that is, 2^{k-2}) can be replaceable, since they are followed by Montgomery multiplications in the next kth stage. Thus, in the $(k-1)$th stage, the indices with Montgomery multiplication are partitioned exactly in half, depending on whether there is subsequent Montgomery multiplication in the kth stage or not. Only for the indices with subsequent Montgomery multiplication, modular reduction can be skipped and standard multiplication can be used instead of Montgomery multiplication.
3. Similarly, in the $(k-2)$th stage, 2^{k-2} Montgomery multiplications out of the total 2^{k-1} can be replaceable, since there are remaining Montgomery multiplications in the following $(k-1)$th and kth stages.
4. Thus, for every stage except the final one, half of the required multiplications can be replaced with standard multiplications. As a result, a total of $(k-1) \cdot 2^{k-2}$ Montgomery multiplications out of $k \cdot 2^{k-1}$ can be replaced, achieving asymptotically half-reduced multiplicative complexity.

□

The proposed *SM-LazyNTT* algorithm is given in Algorithm 3. The input $\overline{\mathbf{f}}$ is the coefficient vector of $f(x)$. The input $\overline{\mathbf{G}}$ stores all odd powers of the primitive $2n$-th root of unity, α^{2i+1}, in Montgomery space, as in Algorithm 2. The input $\mathbf{G}$ stores all odd powers of the primitive $2n$-th root of unity, but they are in standard space. In lines 11, 37, and 53 (blue colored) of Algorithm 3, Montgomery multiplications are used. However, in lines 21 and 43 (red colored), standard multiplications are used, providing modular reduction-free multiplication. Unlike the original NTT, where computations proceed sequentially with indices increasing by 1 (as in the "for" loop in line 7 of Algorithm 2), the proposed *SM-LazyNTT* requires specifying the positions of standard and Montgomery multiplications. To determine these positions, the index jumping vector $\mathbf{J}$ is used. Details about $\mathbf{J}$ are described in Section 3.3. The two for-loops starting in lines 33 and 49 correspond to the second-to-last stage and the final stage, respectively. In the second-to-last stage, each sub-NTT contains one standard multiplication and one Montgomery multiplication, and thus the jumping index is not used. Additionally, in the final stage, only Montgomery multiplications are performed, which is why a separate loop is created for it.

In *SM-LazyNTT*, larger intermediate values can arise. These larger values are subsequently reduced to the range of $[0, q)$ through Montgomery multiplication. After one standard multiplication, the resulting values range from 0 to q^2. By using the subtraction part in a butterfly operation, the result ranges from $-q^2$ to q^2. The conditional addition (line 8 of Algorithm 1) in Montgomery multiplication is performed to the value but uses q^2 instead of q, ensuring that the final result remains in the range from 0 to q^2.

Algorithm 3 *SM-LazyNTT*

Require: $\bar{\mathbf{f}}, \overline{\mathbf{G}}, \mathbf{G}, \mathbf{J}, \log n, q$

$n' \leftarrow 2^{(\log n - 2)}$
for $m \leftarrow 1$ to n' **do**
  $t \leftarrow n/2$
  $h \leftarrow t/2$
  for $u \leftarrow 0$ to $m-1$ **do**
    $j_1 \leftarrow 0, j_2 \leftarrow 0, v \leftarrow 0$
    for $d \leftarrow 0$ to $h-1$ **do**
      $\bar{r}_1 \leftarrow \bar{\mathbf{f}}[v]$
      $\bar{r}_2 \leftarrow \bar{\mathbf{f}}[v+t]$
      $\bar{x} \leftarrow \bar{r}_1$
      $\bar{y} \leftarrow \bar{r}_2 * \overline{\mathbf{G}}[m+u]$ ▷ Montgomery multiplication
      $\bar{\mathbf{f}}[v] \leftarrow \bar{x} + \bar{y} \mod q$
      $\bar{\mathbf{f}}[v+t] \leftarrow \bar{x} - \bar{y} \mod q$
      $v \leftarrow v + (4 - \mathbf{J}[j_1])$
      $j_1 \leftarrow j_1 + 1$
    end for
    for $d \leftarrow 0$ to $h-1$ **do**
      $\bar{r}_1 \leftarrow \bar{\mathbf{f}}[v+1]$
      $\bar{r}_2 \leftarrow \bar{\mathbf{f}}[v+1+t]$
      $\bar{x} \leftarrow \bar{r}_1$
      $\bar{y} \leftarrow \bar{r}_2 \cdot \mathbf{G}[m+u]$ ▷ Standard multiplication
      $\bar{\mathbf{f}}[v+1] \leftarrow \bar{x} + \bar{y}$
      $\bar{\mathbf{f}}[v+1+t] \leftarrow \bar{x} - \bar{y}$
      $v \leftarrow v + \mathbf{J}[j_2]$
      $j_2 \leftarrow j_2 + 1$
    end for
    $v \leftarrow v + n$
  end for
  $n \leftarrow t$
  $m \leftarrow 2 \cdot m$
end for
$v \leftarrow 0$
for $u \leftarrow 0$ to $m-1$ **do**
  $\bar{r}_1 \leftarrow \bar{\mathbf{f}}[v]$
  $\bar{r}_2 \leftarrow \bar{\mathbf{f}}[v+2]$
  $\bar{x} \leftarrow \bar{r}_1$
  $\bar{y} \leftarrow \bar{r}_2 * \overline{\mathbf{G}}[m+u]$ ▷ Montgomery multiplication
  $\bar{\mathbf{f}}[v] \leftarrow \bar{x} + \bar{y} \mod q$
  $\bar{\mathbf{f}}[v+2] \leftarrow \bar{x} - \bar{y} \mod q$
  $\bar{r}_1 \leftarrow \bar{\mathbf{f}}[v+1]$
  $\bar{r}_2 \leftarrow \bar{\mathbf{f}}[v+3]$
  $\bar{x} \leftarrow \bar{r}_1$
  $\bar{y} \leftarrow \bar{r}_2 \cdot \mathbf{G}[m+u]$ ▷ Standard multiplication
  $\bar{\mathbf{f}}[v+1] \leftarrow \bar{x} + \bar{y}$
  $\bar{\mathbf{f}}[v+3] \leftarrow \bar{x} - \bar{y}$
  $v \leftarrow v + 4$
end for
$v \leftarrow 0, m \leftarrow 2 \cdot m$
for $u \leftarrow 0$ to $m-1$ **do**
  $\bar{r}_1 \leftarrow \bar{\mathbf{f}}[v]$
  $\bar{r}_2 \leftarrow \bar{\mathbf{f}}[v+1]$
  $\bar{x} \leftarrow \bar{r}_1$
  $\bar{y} \leftarrow \bar{r}_2 * \overline{\mathbf{G}}_1[m+u]$ ▷ Montgomery multiplication
  $\bar{\mathbf{f}}[v] \leftarrow \bar{x} + \bar{y} \mod q$
  $\bar{\mathbf{f}}[v+1] \leftarrow \bar{x} - \bar{y} \mod q$
  $v \leftarrow v + 2$
end for

3.2. SSM-LazyNTT

Building on the previous section about *SM-LazyNTT*, we can further consider a method that replaces one more Montgomery multiplication. That is, up to two consecutive preceding multiplications before Montgomery multiplication can be replaced with standard multiplications. In this case, we can state the condition for the replacement as follows.

Claim 3. *To replace up to two consecutive Montgomery multiplications with standard multiplications, the condition $R > q^3$ should be satisfied, and there should be at least one Montgomery multiplication in the subsequent stages of the NTT computation.*

Proof. From the Claim 1, we know that the result of a Montgomery multiplication should be less than $R \cdot q$.

1. After two consecutive standard multiplications without modular reduction, the result lies in the range $[0, q^3)$. To perform the next Montgomery multiplication with an element in $\mathbb{Z}_q$, the condition $q^4 < R \cdot q$ should be satisfied, which implies that $R > q^3$.
2. Similarly, there should be at least one Montgomery multiplication in the subsequent NTT stages to ensure that the computed result is properly modular reduced.

□

For a polynomial, $f(x)$, in (1), the number of standard multiplications and Montgomery multiplications in the *SSM-LazyNTT* algorithm are counted by the following claim, where $HW(i)$ represents the Hamming weight of i.

Claim 4. *In the proposed SSM-LazyNTT algorithm, the number of Montgomery multiplications is*

$$\sum_{i=0}^{2^k-1} \left\lceil \frac{HW(i)}{3} \right\rceil$$

and the number of standard multiplications is

$$\sum_{i=0}^{2^k-1} \left(HW(i) - \left\lceil \frac{HW(i)}{3} \right\rceil \right).$$

Proof. We know that multiplication is performed in the positions where the bit is 1 in the binary representation of an index, i.

1. Since the final multiplication of each index should be performed by Montgomery multiplication, the possible patterns of the multiplications are combinations of M, S-M, and S-S-M.
2. Since Montgomery multiplication is only performed in the last of every three multiplications, the number of Montgomery multiplications required in each index i is given by $\left\lceil \frac{HW(i)}{3} \right\rceil$. The remaining multiplications are performed as standard multiplications, and, thus, the number of replaceable multiplications in each index is $\left(HW(i) - \left\lceil \frac{HW(i)}{3} \right\rceil \right)$.

□

The proposed *SSM-LazyNTT* algorithm is presented in Algorithm 4. It is similar to Algorithm 3, but with some key differences. First, unlike in *SM-LazyNTT*, where standard and Montgomery multiplications are equally partitioned, here, there are more standard multiplications. Therefore, we define the vectors **S** and **M**, specifying the number of multiplications required in each stage. Second, unlike in *SM-LazyNTT*, where a single jumping vector **J** is sufficient to indicate both positions of the standard multiplication and Montgomery multiplication, here, the separate index jumping vectors $\mathbf{J}_S$ and $\mathbf{J}_M$ for standard and Montgomery multiplication are required, respectively.

Algorithm 4 *SSM-LazyNTT*

Require: $\bar{\mathbf{f}}, \overline{\mathbf{G}}, \mathbf{G}, \mathbf{J}_S, \mathbf{J}_M, \log n, q$

$n' \leftarrow 2^{(\log n - 3)}$
$\mathbf{M} = \{170, 85, 43, 22, 11, 5, 2\}$
$\mathbf{S} = \{342, 171, 85, 42, 21, 11, 6\}$
for $m \leftarrow 1$ to n' **do**
  $t \leftarrow n/2$
  $i_M \leftarrow (10 - \log n), i_S \leftarrow (10 - \log n)$
  for $u \leftarrow 0$ to $m - 1$ **do**
    $j_1 \leftarrow 0, j_2 \leftarrow 0, v \leftarrow 0$
    for $d \leftarrow 0$ to $M[i_M]$ **do**
      $\bar{r}_1 \leftarrow \bar{\mathbf{f}}[v]$
      $\bar{r}_2 \leftarrow \bar{\mathbf{f}}[v + t]$
      $\bar{x} \leftarrow \bar{r}_1$
      $\bar{y} \leftarrow \bar{r}_2 * \overline{\mathbf{G}}[m + u]$ ▷ Montgomery multiplication
      $\bar{\mathbf{f}}[v] \leftarrow \bar{x} + \bar{y} \mod q$
      $\bar{\mathbf{f}}[v + t] \leftarrow \bar{x} - \bar{y} \mod q$
      $v \leftarrow v + \mathbf{J}_M[j_1]$
      $j_1 \leftarrow j_1 + 1$
    end for
    for $d \leftarrow 0$ to $S[i_S]$ **do**
      $\bar{r}_1 \leftarrow \bar{\mathbf{f}}[v + 1]$
      $\bar{r}_2 \leftarrow \bar{\mathbf{f}}[v + 1 + t]$
      $\bar{x} \leftarrow \bar{r}_1$
      $\bar{y} \leftarrow \bar{r}_2 \cdot \mathbf{G}[m + u]$ ▷ Standard multiplication
      $\bar{\mathbf{f}}[v + 1] \leftarrow \bar{x} + \bar{y}$
      $\bar{\mathbf{f}}[v + 1 + t] \leftarrow \bar{x} - \bar{y}$
      $v \leftarrow v + \mathbf{J}[j_2]$
      $j_2 \leftarrow j_2 + 1$
    end for
    $v \leftarrow v + n$
  end for
  $n \leftarrow t, m \leftarrow 2 \cdot m$
end for
$v \leftarrow 0$
for $u \leftarrow 0$ to $m - 1$ **do**
  $\bar{r}_1 \leftarrow \bar{\mathbf{f}}[v]$
  $\bar{r}_2 \leftarrow \bar{\mathbf{f}}[v + 4]$
  $\bar{x} \leftarrow \bar{r}_1$
  $\bar{y} \leftarrow \bar{r}_2 * \overline{\mathbf{G}}[m + u]$ ▷ Montgomery multiplication
  $\bar{\mathbf{f}}[v] \leftarrow \bar{x} + \bar{y} \mod q$
  $\bar{\mathbf{f}}[v + 4] \leftarrow \bar{x} - \bar{y} \mod q$
  for $d \leftarrow 1$ to 3 **do**
    $\bar{r}_1 \leftarrow \bar{\mathbf{f}}[v + d]$
    $\bar{r}_2 \leftarrow \bar{\mathbf{f}}[v + d + 4]$
    $\bar{x} \leftarrow \bar{r}_1$
    $\bar{y} \leftarrow \bar{r}_2 \cdot \mathbf{G}[m + u]$ ▷ Standard multiplication
    $\bar{\mathbf{f}}[v + d] \leftarrow \bar{x} + \bar{y}$
    $\bar{\mathbf{f}}[v + d + 4] \leftarrow \bar{x} - \bar{y}$
  end for
  $v \leftarrow v + 8$
end for
$v \leftarrow 0, m \leftarrow 2 \cdot m$
for $u \leftarrow 0$ to $m - 1$ **do**
  $\bar{r}_1 \leftarrow \bar{\mathbf{f}}[v]$
  $\bar{r}_2 \leftarrow \bar{\mathbf{f}}[v + 2]$
  $\bar{x} \leftarrow \bar{r}_1$
  $\bar{y} \leftarrow \bar{r}_2 * \overline{\mathbf{G}}_1[m + u]$ ▷ Montgomery multiplication
  $\bar{\mathbf{f}}[v] \leftarrow \bar{x} + \bar{y} \mod q$
  $\bar{\mathbf{f}}[v + 2] \leftarrow \bar{x} - \bar{y} \mod q$
  $\bar{r}_1 \leftarrow \bar{\mathbf{f}}[v + 1]$
  $\bar{r}_2 \leftarrow \bar{\mathbf{f}}[v + 3]$
  $\bar{x} \leftarrow \bar{r}_1$
  $\bar{y} \leftarrow \bar{r}_2 \cdot \mathbf{G}[m + u]$ ▷ Standard multiplication
  $\bar{\mathbf{f}}[v + 1] \leftarrow \bar{x} + \bar{y}$
  $\bar{\mathbf{f}}[v + 3] \leftarrow \bar{x} - \bar{y}$
end for
$v \leftarrow 0, m \leftarrow 2 \cdot m$
for $u \leftarrow 0$ to $m - 1$ **do**
  $\bar{r}_1 \leftarrow \bar{\mathbf{f}}[v]$
  $\bar{r}_2 \leftarrow \bar{\mathbf{f}}[v + 1]$
  $\bar{x} \leftarrow \bar{r}_1$
  $\bar{y} \leftarrow \bar{r}_2 * \overline{\mathbf{G}}_1[m + u]$ ▷ Montgomery multiplication
  $\bar{\mathbf{f}}[v] \leftarrow \bar{x} + \bar{y} \mod q$
  $\bar{\mathbf{f}}[v + 1] \leftarrow \bar{x} - \bar{y} \mod q$
  $v \leftarrow v + 2$
end for

However, since these jumping vectors are determined deterministically once the size of the NTT, n, is fixed, they can be precomputed and stored before the algorithm begins, and, thus, this does not pose any issue.

Similarly, the third-to-last, second-to-last, and final stages are also separated into their own for-loops. In the for-loop starting in line 34, i.e., the third-to-last stage, there is only one Montgomery multiplication per sub-NTT, and, thus, the jumping vector is not used here.

In *SSM-LazyNTT*, larger intermediate values can arise, as in *SM-LazyNTT*. A notable point is that the ranges of values differ after the first and second standard multiplications; the first standard multiplication results in values ranging from 0 to q^2, while the second results in values ranging from 0 to q^3. To ensure that these values remain positive and within correct ranges, different conditional additions (adding q^2 and q^3, respectively) are applied for each case. This ensures that the results are always positive and modularly reduced during subsequent Montgomery multiplications.

3.3. Jumping Vector

Unlike the original NTT, which uses only Montgomery multiplication, the proposed algorithms mix standard multiplication and Montgomery multiplication. Therefore, it is necessary to determine in each stage which indices use standard multiplication and which use Montgomery multiplication. In Algorithms 3 and 4, the jumping vectors $\mathbf{J}$, $\mathbf{J}_M$, and $\mathbf{J}_S$ indicate in which indices standard or Montgomery multiplication should be applied. These vectors are deterministic and can be predetermined once the NTT size, n, is defined.

As shown in Figures 1 and 2, by considering Montgomery and standard multiplications from the final stage, we can determine all the multiplications to be performed as either Montgomery multiplications or standard multiplications. Red lines represent standard multiplications, black lines represent Montgomery multiplications, and dashed lines represent no multiplication. Clearly, we do not generate or compute these vectors in each execution. Instead, they can be precomputed and stored for use.

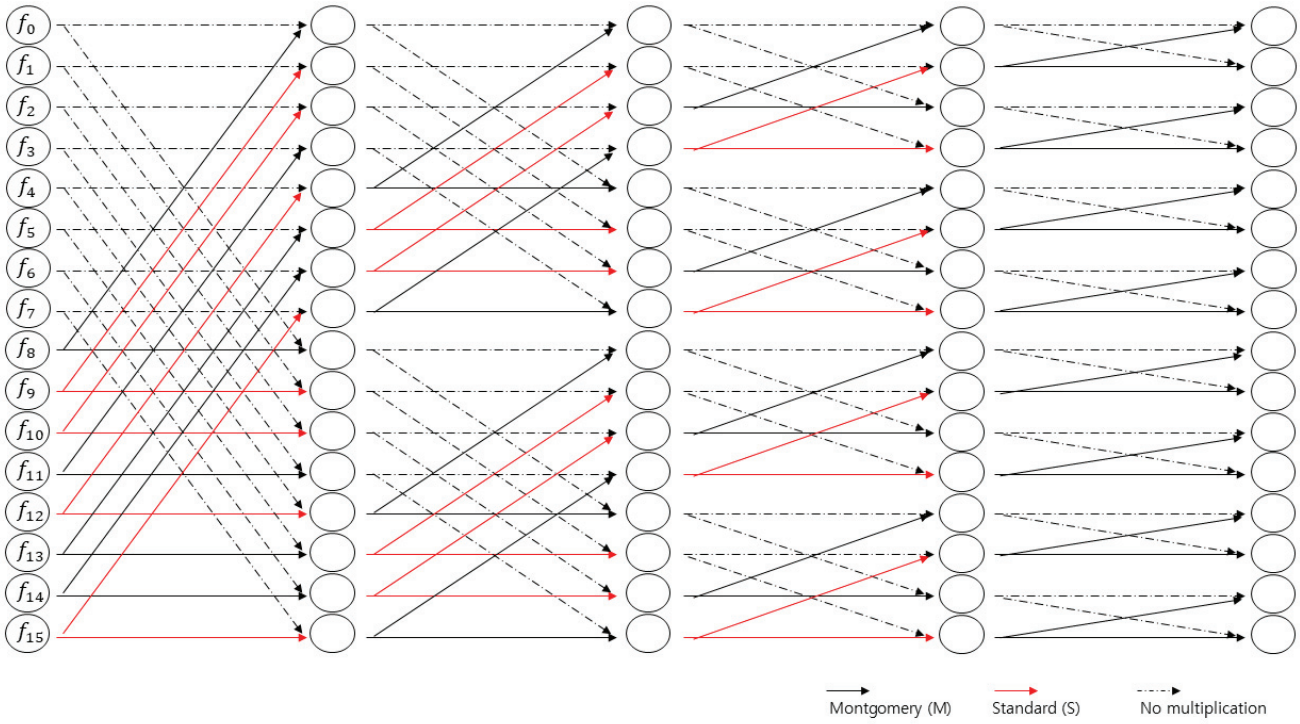


Figure 1. The positions taking standard and Montgomery multiplications in the *SM-LazyNTT* computation of the length $n = 16$.

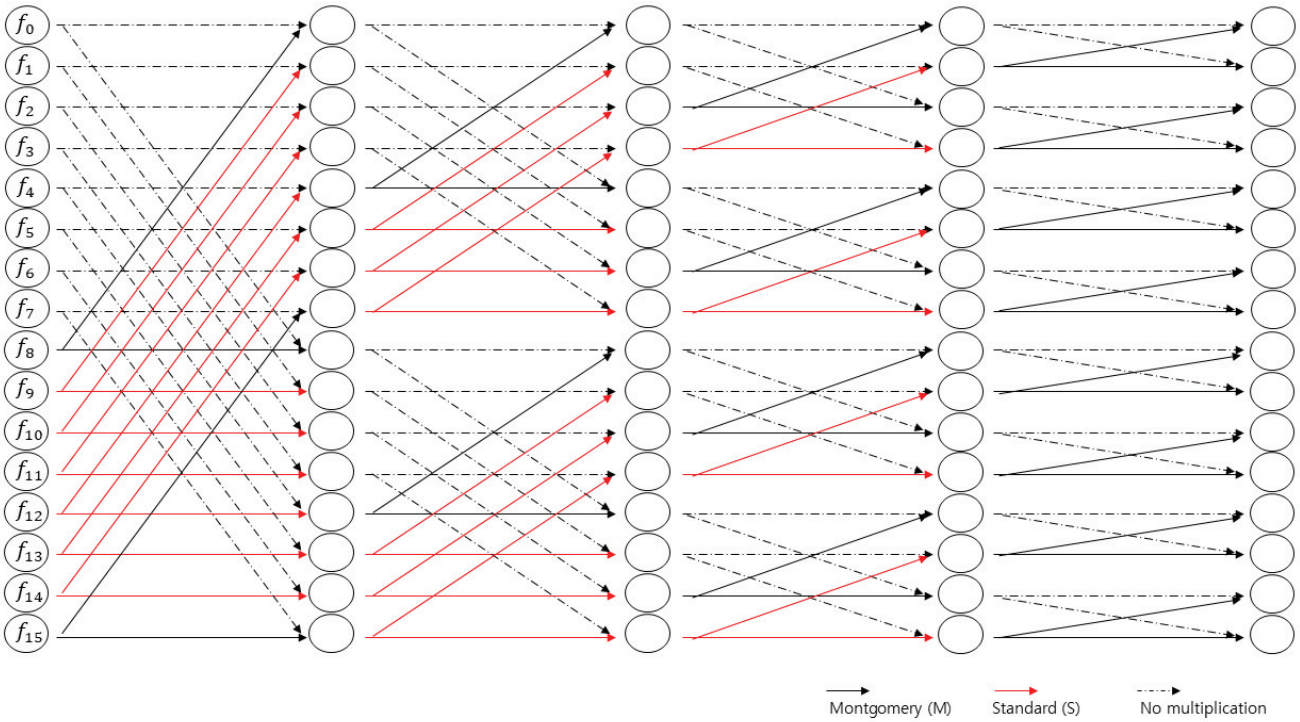


Figure 2. The positions taking standard and Montgomery multiplications in the *SSM-LazyNTT* computation of the length $n = 16$.

3.4. Memory Usage Analysis

LazyNTT may produce larger intermediate values compared to the original NTT and requires slightly more memory, as it utilizes not only $\overline{\mathbf{G}}$, the powers of the root of unity in Montgomery space, but also $\mathbf{G}$, those in standard space. However, the additional memory consumption is not expensive. For cases where $q = 12{,}289$ or $q = 3329$, a 16-bit variable is used, requiring a total of $n \times 16$ bits. Furthermore, since the jumping index values range from 1 to 3 for *SM-LazyNTT* and from 1 to 7 for *SSM-LazyNTT* with a length of $(n/2 - 1)$, eight-bit variables are sufficient to store the indices. This requires $(n/2 - 1) \times 8$ bits in software implementation, but two or three bits are actually enough, respectively. In Algorithm 4, the arrays M and S are needed to store the number of standard and Montgomery multiplications for running multiplications for loops of varying lengths. In summary, the additional memory required compared to the original NTT is not substantial, and this trade-off of using a small amount of additional memory to achieve speed improvement is quite meaningful.

3.5. Generalization

The proposed *LazyNTT* algorithm is not limited to only one or two replacements of Montgomery multiplication but can be extended to more replacements. When multiplying a total of $(t + 1)$ elements in $\mathbb{Z}_q$, we can proceed with $(t - 1)$ standard multiplications instead of Montgomery multiplications. We apply Montgomery multiplication only in the final tth multiplication to ensure that the result is in the range of $\mathbb{Z}_q$. The condition for this to hold is $R > q^t$, since the result x should satisfy $0 \leq x < q^{t+1} \leq q \cdot R$.

In general, for an algorithm that uses $t - 1$ standard multiplications out of a total of t required multiplications, referred to as $S^{t-1}M$*-LazyNTT*, the following claim holds.

Claim 5. *In the $S^{t-1}M$-LazyNTT algorithm, the number of Montgomery multiplications is*

$$\sum_{i=0}^{2^k-1} \left\lceil \frac{HW(i)}{t} \right\rceil,$$

and the number of standard multiplications is

$$\sum_{i=0}^{2^k-1}\left(HW(i)-\left\lceil\frac{HW(i)}{t}\right\rceil\right).$$

Proof. Similarly to Claim 4, we know that Montgomery multiplication is only performed in the last of all consecutive multiplications in each index. Thus, the number of required Montgomery multiplications in each index, i, is given by $\left\lceil\frac{HW(i)}{t}\right\rceil$. The remaining multiplications can be performed by standard multiplications, and, thus, the number of replaceable multiplications in each index is $\left(HW(i)-\left\lceil\frac{HW(i)}{t}\right\rceil\right)$. □

In Falcon [5], the prime modulus and Montgomery scaling factor are $q = 12{,}289$ and $R = 2^{16}$. In the original Falcon implementation [27], since each multiplication is performed using Montgomery multiplication, R only needs to be a power of two larger than q. However, to apply the proposed *SM-LazyNTT* and *SSM-LazyNTT* algorithms, given that the size of q is 14 bits, R should be at least 2^{28} and 2^{42}, respectively. For the Kyber parameter $q = 3329$, given that the size of q is 12 bits, R should be at least 2^{24} and 2^{36}, respectively.

In the original NTT, modular reduction is performed after each multiplication, ensuring that no value exceeds the modulus q. However, in the proposed algorithm, modular reductions are skipped in some of the intermediate stages, requiring the storage of larger values, which sometimes necessitates a 64-bit variable. For *SM-LazyNTT* with $q = 12{,}289$, after one standard multiplication, the result can reach up to q^2, which corresponds to a maximum of 28 bits. To perform the subsequent Montgomery multiplication of this value, a variable capable of holding up to 42 bits is required. In the case of *SSM-LazyNTT*, after two standard multiplications, the value can reach up to 42 bits, and to perform the subsequent Montgomery multiplication, a variable capable of holding up to 56 bits is necessary.

S^{t-1}*M-LazyNTT* can be the generalized version of the proposed *LazyNTT*; however, there are some practical obstacles for its implementation. First, variables larger than 64 bits are required for going beyond *SSM-LazyNTT*. For example, in *SSSM-LazyNTT*, R should be 2^{56} for $q = 12{,}289$ to perform three consecutive standard multiplications, and following Montgomery multiplication would demand a variable capable of $70 = 56 + 14$ bits for its computation. Thus, while it is theoretically possible to extend the algorithm, the need for higher bit-width processing may introduce overhead, potentially diminishing the performance benefits. Additionally, memory usage would increase accordingly to store these larger values. These are more seriously considered in both hardware and software implementations of the generalized version of *LazyNTT*.

4. Implementation Results

We implemented two versions of the proposed algorithm, *SM-LazyNTT* and *SSM-LazyNTT*, using the NTT function in Falcon implementation [27]. The experiments were conducted on a Linux server with an AMD Ryzen 9 5950X 16-Core Processor CPU with 64 GB of memory. We used a "GCC" compiler, and no compiler optimization options were applied. The number of iterations was one million. We compared the proposed algorithms with the original NTT algorithm in terms of the number of cycles for different parameters, q and n. The implementation source code is publicly available at https://github.com/Yongwoo-Lee-ccl/fast-ntt (accessed on 24 October 2024).

Figures 3 and 4 show the performance comparison results of the proposed algorithms, *SM-LazyNTT* and *SSM-LazyNTT*, with the original NTT in Falcon [27]. Clearly, the required cycle counts were reduced in the proposed algorithms. As shown in Tables 2 and 3, we achieved a speedup gain of around 28% with $q = 12{,}289$ and 24% with $q = 3329$ for *SSM-LazyNTT* and around 9% with both parameters for *SM-LazyNTT*.

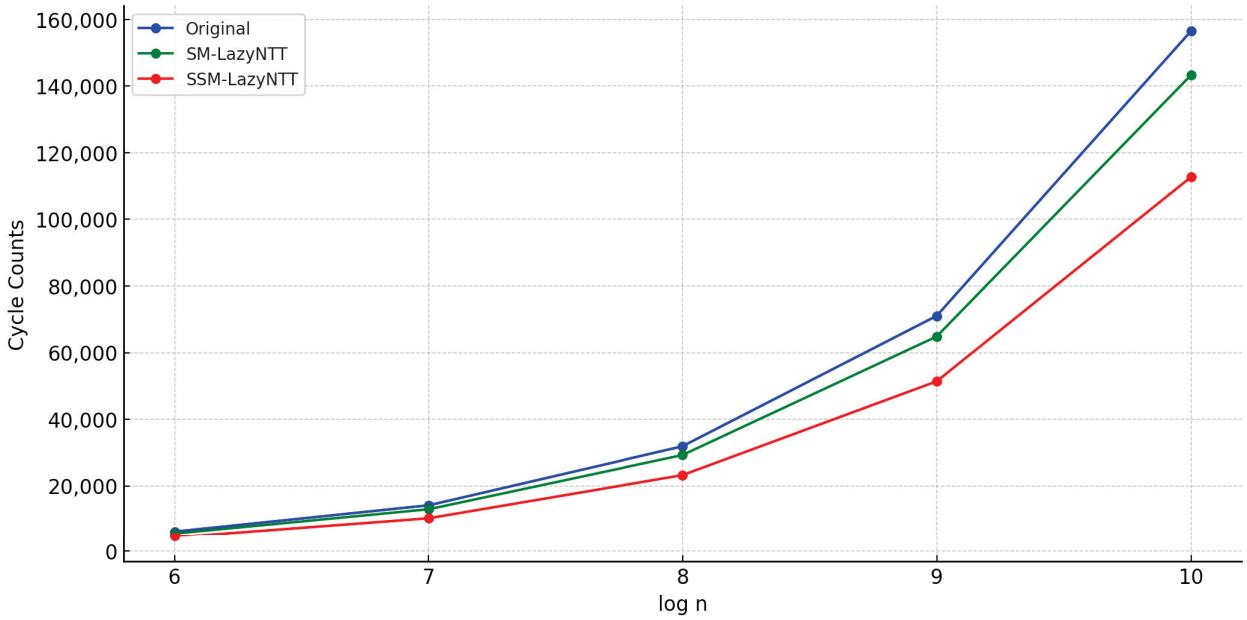


Figure 3. A performance comparison of the original and the proposed *SM-LazyNTT* and *SSM-LazyNTT* algorithms with $q = 12{,}289$.

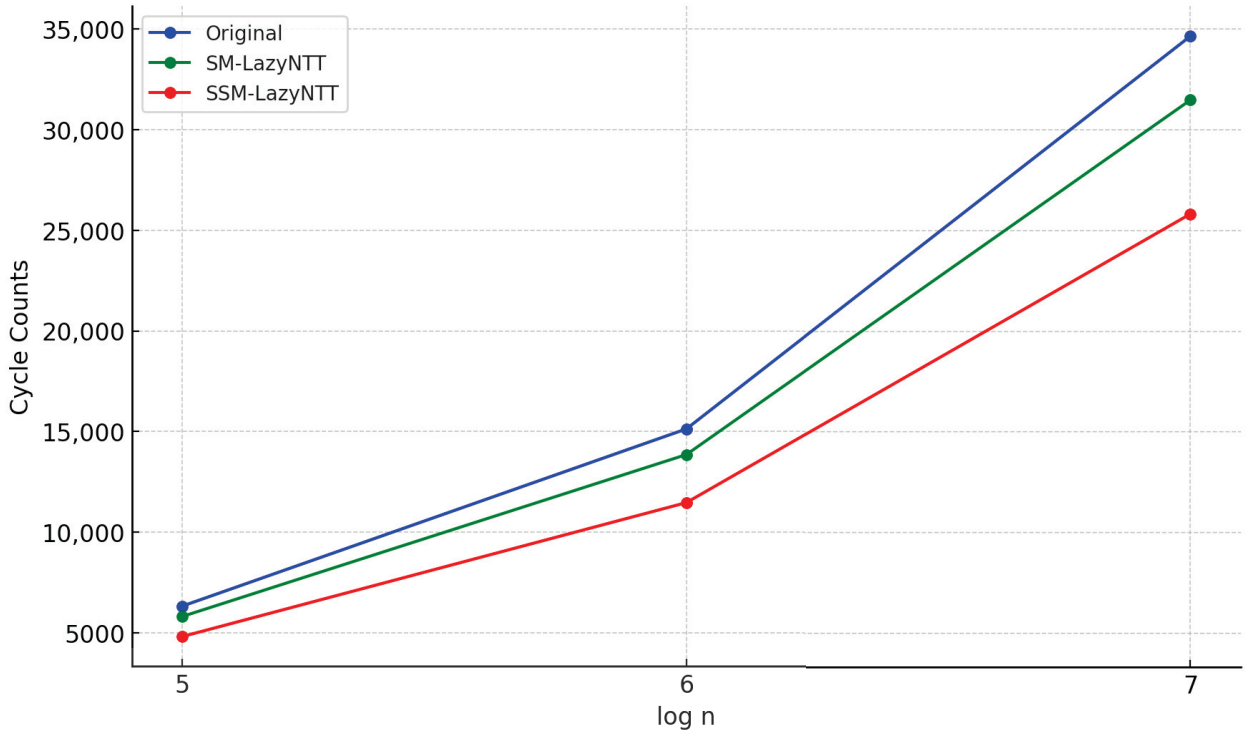


Figure 4. A performance comparison of the original and the proposed *SM-LazyNTT* and *SSM-LazyNTT* algorithms with $q = 3329$.

Table 2. A performance comparison of the original and the proposed *SM-LazyNTT* and *SSM-LazyNTT* algorithms based on the number of cycles with $q = 12{,}289$.

	$q = 12{,}289$				
$\log n$	**6**	**7**	**8**	**9**	**10**
SM-LazyNTT	5660	13,000	29,258	64,843	143,297
SSM-LazyNTT	4514	10,287	23,230	51,384	112,701
Original	6255	14,187	31,889	71,041	156,509
Speedup (SM/Original)	0.905	0.916	0.918	0.913	0.916
Speedup (SSM/Original)	0.722	0.725	0.728	0.723	0.720

Table 3. A performance comparison of the original and the proposed *SM-LazyNTT* and *SSM-LazyNTT* algorithms based on the number of cycles with $q = 3329$.

	$q = 3329$		
$\log n$	**5**	**6**	**7**
SM-LazyNTT	5823	13,860	31,469
SSM-LazyNTT	4826	11,476	25,816
Original	6339	15,135	34,636
Speedup (SM/Original)	0.919	0.916	0.909
Speedup (SSM/Original)	0.761	0.758	0.745

5. Discussion and Conclusions

As the NTT plays an important role in lattice-based cryptosystems, reducing its computational load can be an important contribution. We proposed a new and faster NTT algorithm, *LazyNTT*, and its two versions of implementation. The implementation results showed that *SSM-LazyNTT* and *SM-LazyNTT* were practically efficient, improving the runtime of the NTT by up to 28% and 9%, respectively. Moreover, the proposed technique could be naturally generalized to replace more than two Montgomery multiplications with standard integer multiplications. Even with the replacement of just one or two multiplications, we observed meaningful computational complexity improvements compared to the original NTT and confirmed the proper functionality of the proposed methods. Furthermore, the proposed *LazyNTT* only replaces some Montgomery multiplications with standard multiplications, without changing any other structures of the NTT; the runtime improvement can be achieved without compromising security.

In this paper, *LazyNTT* was implemented for PQC parameters, especially Falcon and Kyber. However, it can also be extended to other cryptographic schemes, such as HE. For example, HE typically employs larger parameters and longer NTT lengths compared to PQC. Consequently, *LazyNTT* could replace a larger number of multiplications, potentially leading to a higher reduction in the number of cycles. Furthermore, when implementing software with large parameters requiring 128-bit variables or more, although there may be additional costs, improved speed with *LazyNTT* compensates the costs and introduces a suitable implementation for homomorphic encryption.

Some future works can be considered in the following directions. First, implementing the generalized version of the proposed *LazyNTT* can be a promising research topic. Second, this approach can be extended not only to the radix-2 NTT but also by mixing NTTs that use various sizes of radix. Third, this technique can be applied to the inverse NTT, thereby improving the overall polynomial multiplication process. Fourth, it can be applied to various lattice-based cryptographic schemes, such as digital signatures, homomorphic encryption, and zero knowledge proofs. Finally, implementing *LazyNTT* on various platforms, such as GPUs or FPGAs, can be a research direction for broader applications.

Author Contributions: G.K. developed and improved the overall idea of this paper, wrote the C code implementation, and derived the experimental results. E.S. initially proposed the idea of this paper. E.S. and Y.L. contributed to the programming and discussed the idea together. Y.-S.K. and J.-S.N. guided the paper's overall direction and structure. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Institute of Information and communications Technology Planning and Evaluation (IITP) grant funded by the Korean government(MSIT) (RS-2024-00399401, Development of Quantum-Safe Infrastructure Migration and Quantum Security Verification Technologies).

Data Availability Statement: The data presented in this study are openly available in [fast-ntt] at https://github.com/Yongwoo-Lee-ccl/fast-ntt.git (accessed on 24 October 2024).

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of this study; in the collection, analysis, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Rivest, R.L.; Shamir, A.; Adleman, L. A method for obtaining digital signatures and public-key cryptosystems. *Commun. ACM* **1978**, *21*, 120–126. [CrossRef]
2. Shor, P.W. Algorithms for quantum computation: Discrete logarithms and factoring. In Proceedings of the 35th Annual Symposium on Foundations of Computer Science, Santa Fe, NM, USA, 20–22 November 1994.
3. Alagic, G.; Apon, D.; Cooper, D.; Dang, Q.; Dang, T.; Kelsey, J.; Lichtinger, J.; Miller, C.; Moody, D.; Peralta, R.; et al. *Status Report on the Third Round of the NIST Post-Quantum Cryptography Standardization Process*; US Department of Commerce: Washington, DC, USA, **2022**.
4. Montgomery, P.L. Modular multiplication without trial division. *Math. Comput.* **1985**, *44*, 519–521. [CrossRef]
5. Fouque, P.A.; Hoffstein, J.; Kirchner, P.; Lyubashevsky, V.; Pornin, T.; Prest, T.; Ricosset, T.; Seiler, G.; Whyte, W.; Zhang, Z.; et al. Falcon: Fast-Fourier lattice-based compact signatures over NTRU. *Post-Quantum Cryptogr. Stand.* **2018**, *36*, 1–75.
6. Bos, J.; Ducas, L.; Kiltz, E.; Lepoint, T.; Lyubashevsky, V.; Schanck, J.M.; Schwabe, P.; Seiler, G.; Stehlé, D. Crystals-Kyber: A CCA-secure module-lattice-based KEM. In Proceedings of the 2018 IEEE European Symposium on Security and Privacy (EuroS&P), London, UK, 24–26 April 2018; pp. 353–367.
7. Yaman, F.; Mert, A.C.; Öztürk, E.; Savaş, E. A hardware accelerator for polynomial multiplication operation of Crystals-Kyber PQC scheme. In Proceedings of the 2021 Design, Automation & Test in Europe Conference & Exhibition (DATE), Grenoble, France, 1–5 February 2021; pp. 1020–1025.
8. Nannipieri, P.; Di Matteo, S.; Zulberti, L.; Albicocchi, F.; Saponara, S.; Fanucci, L. A RISC-V post quantum cryptography instruction set extension for number theoretic transform to speed-up Crystals algorithms. *IEEE Access* **2021**, *9*, 150798–150808. [CrossRef]
9. Nguyen, T.H.; Kieu-Do-Nguyen, B.; Pham, C.K.; Hoang, T.T. High-speed NTT Accelerator for Crystals-Kyber and Crystals-Dilithium. *IEEE Access* **2024**, *12*, 34918–34930. [CrossRef]
10. Li, B.; Yan, Y.; Wei, Y.; Han, H. Scalable and parallel optimization of the number theoretic transform based on FPGA. *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.* **2023**, *32*, 291–304. [CrossRef]
11. Alsuhli, G.; Saleh, H.; Al-Qutayri, M.; Mohammad, B.; Stouraitis, T. Efficient twiddle factor generation for post quantum cryptography Falcon-based number theoretic transform. *Authorea Prepr.* **2024**. [CrossRef]
12. Zhang, C.; Liu, D.; Liu, X.; Zou, X.; Niu, G.; Liu, B.; Jiang, Q. Towards efficient hardware implementation of NTT for Kyber on FPGAs. In Proceedings of the 2021 IEEE International Symposium on Circuits and Systems (ISCAS), Daegu, Republic of Korea, 22–28 May 2021; pp. 1–5.
13. Yang, Y.; Kuppannagari, S.R.; Kannan, R.; Prasanna, V.K. NTTGen: A framework for generating low latency NTT implementations on FPGA. In Proceedings of the 19th ACM International Conference on Computing Frontiers, Turin, Italy, 17–22 May 2022; pp. 30–39.
14. Derya, K.; Mert, A.C.; Öztürk, E.; Savaş, E. CoHA-NTT: A configurable hardware accelerator for NTT-based polynomial multiplication. *Microprocess. Microsystems* **2022**, *89*, 104451. [CrossRef]
15. Bradbury, J.; Drucker, N.; Hillenbrand, M. NTT software optimization using an extended Harvey butterfly. *Cryptol. Arch.* **2021**, 1396. Available online: https://eprint.iacr.org/2021/1396 (accessed on 24 October 2024).
16. Longa, P.; Naehrig, M. Speeding up the number theoretic transform for faster ideal lattice-based cryptography. In Proceedings of the Cryptology and Network Security: 15th International Conference, CANS 2016, Milan, Italy, 14–16 November 2016; Proceedings 15; Springer: Berlin/Heidelberg, Germany, 2016; pp. 124–139.
17. Lyubashevsky, V.; Seiler, G. NTTRU: Truly fast NTRU using NTT. *IACR Trans. Cryptogr. Hardw. Embed. Syst.* **2019**, 180–201. [CrossRef]
18. Özerk, Ö.; Elgezen, C.; Mert, A.C.; Öztürk, E.; Savaş, E. Efficient number theoretic transform implementation on GPU for homomorphic encryption. *J. Supercomput.* **2022**, *78*, 2840–2872. [CrossRef]

19. Goey, J.Z.; Lee, W.K.; Goi, B.M.; Yap, W.S. Accelerating number theoretic transform in GPU platform for fully homomorphic encryption. *J. Supercomput.* **2021**, *77*, 1455–1474. [CrossRef]
20. Kim, S.; Jung, W.; Park, J.; Ahn, J.H. Accelerating number theoretic transformations for bootstrappable homomorphic encryption on GPUs. In Proceedings of the 2020 IEEE International Symposium on Workload Characterization (IISWC), Beijing, China, 27–30 October 2020; pp. 264–275.
21. Kim, S.; Lee, K.; Cho, W.; Nam, Y.; Cheon, J.H.; Rutenbar, R.A. Hardware architecture of a number theoretic transform for a bootstrappable RNS-based homomorphic encryption scheme. In Proceedings of the 2020 IEEE 28th Annual International Symposium on Field-Programmable Custom Computing Machines (FCCM), Fayetteville, AR, USA, 3–6 May 2020; pp. 56–64.
22. Duong-Ngoc, P.; Kwon, S.; Yoo, D.; Lee, H. Area-efficient number theoretic transform architecture for homomorphic encryption. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2022**, *70*, 1270–1283. [CrossRef]
23. Cooley, J.W.; Tukey, J.W. An algorithm for the machine calculation of complex Fourier series. *Math. Comput.* **1965**, *19*, 297–301. [CrossRef]
24. Agarwal, R.C.; Burrus, C.S. Number theoretic transforms to implement fast digital convolution. *Proc. IEEE* **1975**, *63*, 550–560. [CrossRef]
25. Chu, E.; George, A. *Inside the FFT Black Box: Serial and Parallel Fast Fourier Transform Algorithms*; CRC Press: Boca Raton, FL, USA, 1999.
26. Ducas, L.; Kiltz, E.; Lepoint, T.; Lyubashevsky, V.; Schwabe, P.; Seiler, G.; Stehlé, D. Crystals-Dilithium: A lattice-based digital signature scheme. *IACR Trans. Cryptogr. Hardw. Embed. Syst.* **2018**, *2018*, 238–268. [CrossRef]
27. NTT Code from the Falcon Project. Available online: https://falcon-sign.info/impl/falcon.h.html (accessed on 24 October 2024).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

electronics

MDPI

Article

An Applied Analysis of Securing 5G/6G Core Networks with Post-Quantum Key Encapsulation Methods

Paul Scalise, Robert Garcia, Matthew Boeding, Michael Hempel * and Hamid Sharif

Department of Electrical and Computer Engineering, University of Nebraska-Lincoln, Lincoln, NE 68588, USA; pscalise2@huskers.unl.edu (P.S.); rgarcia12@huskers.unl.edu (R.G.); mboeding@huskers.unl.edu (M.B.); hsharif@unl.edu (H.S.)

* Correspondence: mhempel@unl.edu

Abstract: Fifth Generation (5G) cellular networks have been adopted worldwide since the rollout began around 2019. It brought with it many innovations and new services, such as Enhanced Mobile Broadband (eMBB), Ultra Reliable and Low-Latency Communications (URLLC), and Massive Internet of Things (mIoT). Furthermore, 5G introduced a more scalable approach to network operations using fully software-based Virtualized Network Functions (VNF) in Core Networks (CN) rather than the prior hardware-based approach. However, while this shift towards a fully software-based system design provides numerous significant benefits, such as increased interoperability, scalability, and cost-effectiveness, it also brings with it an increased cybersecurity risk. Security is crucial to maintaining trust between vendors, operators, and consumers. Cyberattacks are rapidly increasing in number and sophistication, and we are seeing a shift towards zero-trust approaches. This means that even communications between VNFs inside a 5G core must be scrutinized and hardened against attacks, especially with the advent of quantum computers. The National Institute of Standards and Technology (NIST), over the past 10 years, has led efforts to standardize post-quantum cryptography (PQC) to protect against quantum attacks. This paper covers a custom implementation of the open-source free5GC CN, to expand its HTTPS capabilities for VNFs by introducing PQC Key Encapsulation Methods (KEM) for Transport Layer Security (TLS) v1.3. This paper provides the details of this integration with a focus on the latency of different PQC KEMs in initial handshakes between VNFs, on packet size, and the implications in a 5G environment. This work also conducts a security comparison between the PQC-equipped free5GC and other open-source 5G CNs. The presented results indicate a negligible increase in UE connection setup duration and a small increase in connection setup data requirements, strongly indicating that PQC KEM's benefits far outweigh any downsides when integrated into 5G and 6G core services. To the best of our knowledge, this is the first work incorporating PQC into an open-source 5G core. Furthermore, the results from this effort demonstrate that employing PQC ciphers for securing VNF communications results in only a negligible impact on latency and bandwidth usage, thus demonstrating significant benefits to 5G cybersecurity.

Keywords: post-quantum cryptography; key encapsulation methods; CRYSTALS-Kyber; 5G; 6G; core network; open source; free5GC

1. Introduction

In recent years, 5G has revolutionized cellular communications by providing higher connection throughputs, higher User Equipment (UE) density, lower latencies, improved network configurability, the software virtualization of network services, and a resulting shift towards more affordable commodity hardware platforms for 5G operations and deployments [1]. These improvements have also brought with them a rapidly expanding ecosystem of open-source 5G implementations and tools and even allow them to run or

experiment with personal or private 5G networks. Currently, the available free and open-source 5G tools also include full-stack implementations that provide all the functionality needed for implementing UE, gNodeB (gNB) software implementations using Universal Software Radio Peripherals, and all the required Core Network (CN) software. Some solutions also provide support for an Open Radio Access Network (O-RAN), allowing researchers and customers to mix and match various components to meet their needs, with all of these solutions complying with 3rd Generation Partnership Program (3GPP) standards [2].

Within this realm of commercial and open-source solutions around 5G, security is an ever-evolving topic with many focus areas and areas of research interest. This paper, with its focus on security aspects of open-source CNs, is part of this domain. Reviewing the available 5G open-source CN software packages, three notable solutions are at the forefront: OpenAirInterface, Open5GS, and free5GC. All of these powerful tools leverage the standards set in place by 3GPP but have significant differences when looking at the security capabilities they provide to their users. In an escalating climate of cyber hacks and data leaks, security is ever more important [3]. Users on a public or private 5G network rely on the systems in place to safeguard their data. As our review has shown, free5GC has the most robust security measures in place among these three prevalent open-source 5G projects, incorporating TLS-enabled HTTPS requests between its 5G core's Virtualized Network Functions (VNF).

While this is a great capability considering the current security standards, the feasibility, availability, and capability of quantum computers are rising. Quantum computing represents a revolutionary leap in computational power, leveraging the principles of quantum mechanics to perform complex calculations at speeds unattainable by traditional non-quantum computers. However, this advancement poses a significant threat to current cryptographic systems, as quantum computers could potentially break widely used encryption methods by recovering cryptographic keys that are currently infeasible to be recovered by classical machines.

In cryptography, we consider two primary types of encryption: symmetric and asymmetric encryption. A common example of symmetric encryption, where the same cryptographic key is used for both encryption and decryption, is the Advanced Encryption Standard (AES). Asymmetric encryption utilizes key pairs comprised of a public and a private key that are related to each other. It is primarily used for digital signatures but can also be used for data exchanges. For example, it can be found in key exchange algorithms used with symmetric encryption to ensure the same key is available to both parties. These challenges have been explored by the European Telecommunications Standards Institute (ETSI) [4] and others [5,6]. One common example of that role is found in TLS, which, in turn, is the key security mechanism behind HTTPS and other secure communications approaches. In TLS, symmetric encryption using AES is used for the actual data exchange, but asymmetric methods are used for TLS session establishment and the key exchange. Here, it is important to know that while AES-256 is generally considered quantum-safe, there is an increased risk specifically for asymmetric cryptographic methods, which in turn could pose the risk of quantum computing attacks, revealing the symmetric keys being exchanged during TLS handshaking and thus allowing an attacker to decrypt all the data being exchanged by both parties.

To counter this threat, post-quantum cryptography (PQC) is being developed as a suite of cryptographic algorithms designed to be secure against the capabilities of quantum computers. PQC algorithms are based on mathematical problems that are believed to be resistant to quantum attacks, such as lattice-based, hash-based, and code-based cryptographic methods. Implementing PQC is crucial for future-proofing secure communications and ensuring that sensitive data remain protected as quantum technology evolves. The National Institute of Standards and Technology (NIST)'s efforts in this domain focus on standardizing algorithms to address threats to key encapsulation and digital signature algorithms from quantum computing [6].

PQC is also crucial for securing 5G core networks due to the introduction of VNFs in the 3GPP 5G standards. While 3GPP security standards do not explicitly list support for PQC in their supported TLS guidelines [7], expansion of security profiles Subscription Permanent Identifier (SUPI) protection allows the introduction of quantum-resistant protection [8]. VNFs decentralize and abstract network functions into virtualized software services running on standard hardware, increasing flexibility and efficiency but also expanding the attack surface if the communication between all of these software services, which predominantly use HTTP-based REpresentational State Transfer (REST) API calls between themselves, is not properly secured. In a typical perimeter-based cyber defense posture, where we focus on keeping attackers out of the 5G Core Network, any successful intruder that can establish a foothold inside of the 5G core poses a grave threat to any and all 5G core activities and its users. With current security measures, all 5G CNs are at risk of Store Now Decrypt Later (SNDL) attacks. This is the idea of collecting encrypted data and decrypting it later using, for example, a quantum computer. In the event an attack like this succeeds, sensitive customer connection data will be leaked. In order to protect against these possible attacks, CNs must be equipped with strong Key Encapsulation Methods (KEM).Current 5G CNs that have TLS capabilities, like free5GC, are still vulnerable to SNDL attacks. Specifically, connections between the VNFs within the CN are at risk if these data are ever leaked and must adopt PQC features to mitigate this vulnerability.

The remainder of this paper is organized as follows. Section 2 outlines an overview of our approach, Section 3 provides an overview of related research efforts, Section 4 outlines current open-source 5G cores and compares their implemented security features, and Section 5 gives an overview of our integration of PQC algorithms into the open-source free5GC package. Section 6 will outline the results of our 5G core efforts, including our obtained bandwidth and latency results for connecting UEs to 5G services when different PQC KEM algorithms are being employed, and Section 7 will conclude our work and provide future research directions.

2. PQC Protections for 5G Core Networks

Open Quantum Safe (OQS) [9] has assembled open-source packages to address this specific issue, integrating quantum-resistant KEMs into OpenSSL. With the current 5G CNs being susceptible to SNDL attacks, our research, and development of a quantum-safe CN, show that it is absolutely feasible for production systems, with only a small amount of associated overhead. Figure 1 showcases the main systematic components of a 5G network, including the focus of this paper—a quantum-safe 5G CN. Our research is already continuing to expand these security provisions to the communications between the Radio Access Network (RAN) gNodeB base stations and the 5G Core, as well as for end-to-end protections for the user plane traffic, especially when it involves multi-access edge computing integrated with the 5G infrastructure. These efforts are also shown in Figure 1.

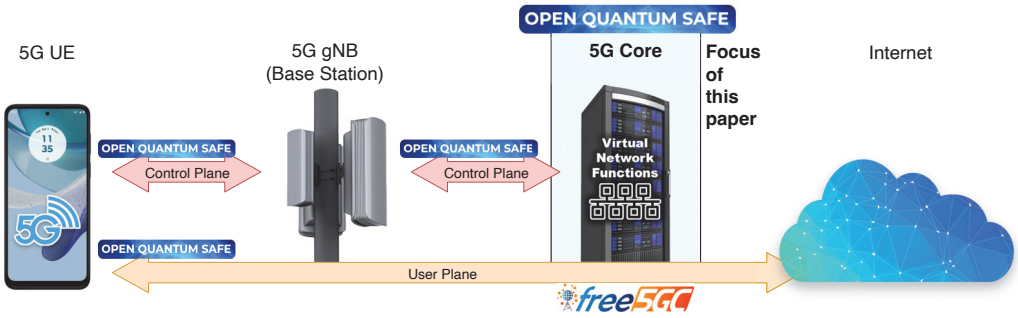


Figure 1. A 5G system overview and the inclusion of post-quantum cryptography for secure end-to-end and control communications.

In Figure 2 below, we are providing a more in-depth look into the 5G core architecture and its set of VNFs, all of which communicate with each other using HTTP REST API calls, which for security would utilize TLS to secure the transport layer. TLS, as mentioned earlier, uses asymmetric cryptography for the session establishment, and it is here that we focused our research in order to provide PQC protections for the TLS key exchange process.

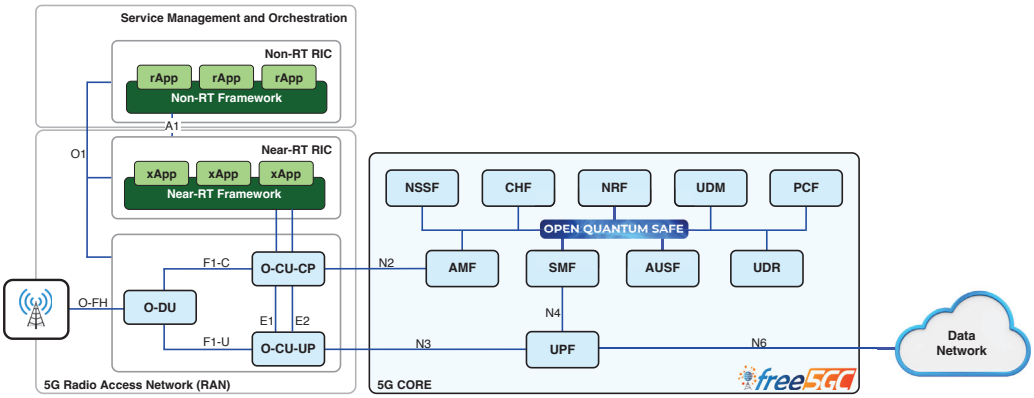


Figure 2. Addition and analysis of Post-Quantum Cryptography in a 5G Core Network.

This paper introduces an integration of PQC algorithms within an open-source 5G core network, free5GC. A detailed analysis of these algorithms and their effect on latency and bandwidth during a UE attachment to the CN is conducted. From the results presented in this paper, there are existing PQC implementations that achieve similar latency to traditional security methods, while having limited effect on a network's utilized bandwidth.

3. Related Works

As quantum computing (QC) continues to evolve, finds practical applicability, and is more widely available, the threat of current encryption methods becoming vulnerable to quantum algorithms able to determine the encryption key and thus enabling the unauthorized decryption of the protected information grows as well. The discovery of various quantum algorithms (QA) has also shown that current TLS is at risk given a sufficiently powerful quantum computer infrastructure. Current cryptography techniques leverage math problems that would be too resource- and time-intensive to solve with a typical computer [10]. With a traditional Turing-based computer, the difficulty of factorization of a number grows exponentially with the size of the number. In 1999, Peter Shor developed Shor's Factorization Algorithm—a quantum computing algorithm that can factorize any given number in logarithmic time, potentially reducing a factorization that would take thousands of years down to a few minutes [11]. Similarly, in 1996, Grover's algorithm was discovered, which leverages quantum computing in order to take the problem of finding a random record in N records from O(N) difficulty down to O(log(N)) [12]. Due to their ease of solving the mathematical problems behind current encryption methods, quantum computers have shown to be relatively effective at circumventing different types of currently employed encryption methods such as Rivest–Shamir–Adleman (RSA), Diffie–Hellman (DH), and Elliptic Curve Cryptography (ECC) [13]. Therefore, an urgent need to create and implement algorithms that cannot be easily solved by either a traditional or quantum computer has emerged.

Much of the current internet infrastructure relies on one or more of these encryption algorithms, and due to the widespread nature of the problem, a need was identified to find quantum-safe cryptographic algorithms. In 2016, the United States NIST called for proposals of alternative Public Key Encryption Schemes, Key Encapsulation Methods, and Digital Signature Schemes, which would be resistant to a quantum attack [14]. The goal of this process is to identify the PQC algorithm best suited for widespread adoption and standardization. In the first round of submissions, a total of 69 algorithms were

submitted. These were then evaluated based on three major characteristics: security, cost and performance, and algorithm implementation characteristics. NIST completed three rounds of evaluation, and encryption methods were improved and resubmitted as the evaluation progressed [15]. As of 13 August 2024, NIST has released its first three finalized encryption standards, which will significantly reduce the threat posed by quantum attacks on systems that implement PQC approaches. NIST has also stated that it will continue to evaluate alternative encryption methods, but these should be considered backups to the finalized selections [16].

While these algorithms can be applied to regular internet traffic, they are far more advisable for consideration in addressing specific complexities and needs in areas such as 5G security. First, 5G UEs must communicate with a base station, referred to as a gNB by 3GPP, which then communicates with a 5G CN. This 5G CN subsequently performs authentication and provides the UE with access to the internet. One of the main features of 5G is that many of the prior generations' hardware-based network operations are now represented by software services and are able to be virtualized [17]. These network functions use HTTP REST API calls to communicate with each other [18]. HTTP communications do not provide any secure encrypted communications. For that, HTTPS was created, facilitating security through Transport Layer Security (TLS) capabilities. Whenever one of the VNFs requires the services of another VNF, it will first need to establish a HTTP/HTTPS connection to that VNF service before exchanging data securely. The capabilities of software virtualization in 5G mean that there are no restrictions about where each VNF could be running. A 5G core's entire set of VNFs could be deployed on the same physical host or be distributed across a number of physical hosts. But regardless of where they are deployed, they do require network communications to function properly, and for those communications to be secure. One of the goals of the research presented in this paper is thus to strengthen the secure communication between these network functions, ensuring that any information exchanged is quantum-safe.

There is existing research on applying PQC algorithms to the 5G domain. In [19], Figlarz and Hessel apply a PQC signature method to enhance the security of the 5G-AKA key exchange. Hoque, Abdullah, and Zeydan present a PQC key distribution framework for Internet of Things (IoT) networks that utilize 5G networks in [20]. In [21], the authors Minhas and Mansoor propose securing communication between edge-computing devices for e-Health applications in IoT, which often rely on a cellular connection through 5G. One of the key performance indicators of 5G networks is to be able to support multi-access edge computing (MEC), or a form of computing where applications and data are stored at the network edge, greatly reducing latency and improving scalability, compared to a cloud-based service, and allowing for low-latency computation to be performed off of constrained devices [22]. Previous works have created an implementation of this service, such as a 2016 report that presents an architecture for MEC and, through their tests, has shown that MEC is capable of applications with a network delay of as little as 17 ms [23]. Other works have shown that it is practical and have shown a minimum latency of as little as 10 ms in the network [24]. However, there is also existing concern about the speed at which PQC can be utilizes, especially in TLS. In [25], Zheng et al. evaluated PQ-KEM algorithms and found hardware advancements speed up PQC key generation by 3.5 to 4.9 times. In this paper, we expand on the current 5G security research by implementing PQC key exchange for communication between each of the VNFs within the 5G core network and present the resulting performance impact to latency and bandwidth compared to traditional cryptographic algorithms.

In [26], Zeydan et al. introduce a novel architecture that integrates Quantum Key Distribution (QKD) with Blockchain-based Self-Sovereign Identity (SSI) within the Open RAN (O-RAN) framework for 6G networks, aiming to enhance security and identity management. It explores the practical applications, challenges, and future directions of this integration, emphasizing its potential to transform telecommunication networks into more secure and decentralized systems. The authors of [27] explore the integration of

quantum computing, federated learning, and 6G networks to enhance security and privacy in the Internet of Things (IoT) ecosystem, addressing the unique challenges posed by the proliferation of IoT devices. They also provide an overview of recent developments, a conceptual framework for implementing these technologies, and insights into future advancements that could significantly improve IoT security. Nakka et al. [28] address the threat of quantum computing in distributed energy resources (DER), explore the adoption of PQC within the IEEE 2030.5 standard for DER networks, and present a real-time hardware-in-the-loop co-simulation testbed to validate the feasibility of implementing PQC in DER systems. Additional work as been conducted in evaluating PQC for IP Security (IPSec), which achieved 100 Gbps throughput in a high-speed data center and 0.486 Gbps in a local-to-cloud architecture [29].

4. Current Open-Source 5G Core Security

While the focus of this paper is on implementing post-quantum cryptography for a custom implementation of free5GC and the resulting performance considerations, we first present our comparative analysis of prominent, currently available open-source implementations of 5G cores in order to compare and contrast what security measures are currently in place for each of them. There are currently five implementations of open-source 5G networks available. Of these five, there are three that have complete core network implementations: Open5GS, free5GC, and Open Air Interface (OAI) [30]. Current published research analyzing the security of network functions focuses on these three. The study of these three core networks (CN) identified and briefly described any existing vulnerabilities. In our own review, outlined in Table 1, it was found that the OAI CN did not implement TLS at all. Every REST API call between 5G core VNFs is performed using unencrypted HTTP. Furthermore, it is critical to analyze the projects utilizing TLS to determine the vulnerabilities in the TLS communication used throughout the network. The OAI CN was excluded from this review because it did not implement TLS [31].

Table 1. Comparison of open-source 5G core networks (CNs) with TLS and PQC support.

Open-Source 5G CNs	TLS	Total Ciphers	Weak Ciphers	Vulnerabilities/ Weaknesses	Cipher Ordering	PQC KEMs
OAI	No	N/A	N/A	N/A	N/A	N/A
Open5GS	v1.2–v1.3	16	10 [31]	None Found	No	No
free5GC	v1.0–v1.3	14	8	Sweet32, TLS Version Intolerance [31]	Yes	No
Our PQC free5GC	v1.3	31	TBD	TBD	Yes	Yes

TLS is recommended by 3GPP to be used for any 5G core communications to improve security. Hence, we can see that OAI does not fully meet 3GPP recommendations, at least in the state it was in during the research [2]. The free5GC project currently implements TLS versions 1.0–1.3, while Open5gs implements versions 1.2 and 1.3. Having these outdated versions of TLS opens free5GC up to a downgrade attack where, if a TLS version cannot be agreed upon, then the version will be downgraded to a less secure one until the two communicating parties both agree on a version. Limiting this to only recent TLS versions decreases the risks of downgrading attacks in Open5gs, giving it an advantage in this regard. In terms of TLS vulnerabilities and weaknesses, the tools used in this research, TLS-Scanner [32], SSLyze [33], and PySSLScan [34], were unable to find any vulnerabilities in Open5GS, while free5GC is vulnerable to the "sweet32" attack and has the "version intolerance" bug.

Additionally, we observed that in both implementations that have TLS, a default key and certificate were used. As stated in the research [31], this could have catastrophic consequences if these cores were ever commercially deployed unmodified, as anyone could retrieve the public key from GitHub and then would be able to decrypt messages being sent

between network functions. It was found that free5GC makes use of 14 different ciphers, all of which are ordered by strength. The application then uses the most complex first, going down the list if the strongest cannot be used. In research, it has been found that 8 of these 14 cipher suites are weak, with some being broken as far back as 2017 [31]. Open5GS also uses weak cipher suites, with 10 of the 16 ciphers used being considered weak. Additionally, Open5GS does not have cipher ordering, meaning there are no guarantees that the strongest available cipher will be used in communication. Open5GS and OAI can also benefit from the addition of PQC. OAI would first need to implement TLS capabilities, and since OAI is written in the C language, direct access to OpenSSL at the system level can be integrated. Like OAI, Open5GS could also benefit from the integration of PQC, utilizing a similar migration process as suggested for OAI.

However, there are some challenges and risks associated with integrating PQC algorithms into projects such as these, as PQC implementations are inherently unproven and can have potentially negative security ramifications if implemented incorrectly. While this is a concern in general, several projects already use the OQS implemented we selected successfully, specifically to mitigate this risk. Additionally, there is no straightforward approach for empirically evaluating the security improvements provided by such PQC implementations, which we consider out of the scope of this paper. It is also possible for malicious parties to subvert open-source third-party packages in what is referred to as a software supply chain attack, in order to introduce vulnerabilities, then propagating to dependent packages. While not detailed in this paper, we performed due diligence and reviewed all utilized packages for any reported vulnerabilities.

There are also computational trade-offs for utilizing PQC in a system. Additional processing for encryption and decryption, higher bandwidth utilization, and increased latency are all potential concerns when utilizing PQC instead of more traditional cryptographic approaches. During our testing, no noticeable increase in CPU usage was observed within the system. Thus, the primary focus of our study is on the implication of our PQC integration on the utilized bandwidth and latency, our study of which is presented in Section 6 of this paper.

While the aim of this research is on the current implementations of 5G CNs, research towards 6G networks identifies additional capabilities. Current research suggests 6G implementations will include Holographic MIMO, which leverages large intelligent surfaces (LIS) [35], the implementation of AI/ML for radio access [36], AI/ML for more efficient receivers [37], and Reconfigurable Intelligent Surfaces (RIS) for improved energy efficiency and adaptive beam forming [38]. Core network improvements will focus on quality of service improvements and may have lower latency and bandwidth tolerances than presented in this paper [36]. MEC is also predicted to experience increased adoption with the introduction of 6G standards, which often have different requirements for latency and processing time. However, there are concerns with adding PQC methods to fog and cloud computing applications, as key size, encryption processing time, and data overhead may increase latency [39].

5. Methodology

As NIST continues to standardize PQC, the expectation is that this technology will experience a rapid acceleration of its integration into a wide variety of existing and future software and infrastructure projects in an effort to maintain a high level of security and privacy protections. To aid in this effort, the open-source community started a new project called "Open Quantum Safe" (OQS). OQS's efforts focus on creating and maintaining a free implementation of PQC algorithms in the form of "liboqs", as well as integration of this capability into common providers and architectures, such as OpenSSL [40], BoringSSL [41], and more. Specifically, enabling the OQS extension for OpenSSL allows for a more straightforward approach to testing and researching these algorithms in lab and production systems before, during, and after their standardization by NIST. The specific

repository used in this research, oqs-provider, lists its included PQC KEM and signature algorithms. A list of these KEM and signature algorithm groups is shown in Table 2.

Table 2. KEM and signature algorithm groups.

Category	Main Groups
KEM Algorithms	BIKE, CRYSTALS-Kyber, FrodoKEM, HQC, ML-KEM
Signature Algorithms	CRYSTALS-Dilithium, ML-DSA, Falcon, SPHINCS-SHA2, SPHINCS-SHAKE, MAYO

Specifically, the main focus of testing is put on CRYSTALS-Kyber, as it is the main KEM that is being considered for standardization by NIST. The full list of KYBER PQC and hybrid (a combination of traditional methods with PQC) KEMs included in oqs-provider is detailed in Table 3.

Table 3. CRYSTALS-Kyber algorithms.

KEM Algorithm	Subgroups
CRYSTALS-Kyber	kyber512, p256_kyber512, X25519_kyber512, kyber768, p384_kyber768, x448_kyber768, X25519_kyber768, p256_kyber768, kyber1024, p521_kyber1024

The full incorporation of PQC into a 5G CN does require careful consideration of understanding the components and interfaces that require changes. If there are any missed in the process, essential CN functions will no longer work, causing fatal system errors.

The process outlined in Figure 3 details the high-level steps taken to add PQC into a 5G CN. This methodology can also be used as a roadmap to incorporate PQC into other 5G systems, as the underlying structure of different CNs is comparable in accordance with 3GPP standards.

In order to utilize oqs-provider, it needs to be compiled from source, which ultimately generates the necessary shared library file that OpenSSL utilizes to incorporate the PQC KEMs and signature algorithms. In Listing 1 in Appendix A, we are demonstrating the instructions for this procedure that we have incorporated inside each of the Dockerfiles associated with all the NFs included in free5GC.

To fully enable the features provided by this extension in a variety of environments, OQS provides detailed instructions in their documentation as well. For the purpose of this research, the CRYSTALS-Kyber KEMs were activated at the system level by adding their values to the 'openssl.cnf' configuration file. Figure 4 shows the lines that were added or updated in order to enable OQS.

The free5GC project is a free and open-source 5G CN that is currently compliant with 3GPP Release 15, with ongoing development aiming at compliance with 3GPP Release 17 [42]. In our research, we cloned the original free5GC repositories and subsequently modified them to incorporate PQC capabilities using oqs-provider. Free5GC has a "docker-compose" deployment method that allows for seamless integration and use with systems that support Docker. Figure 5 illustrates the Docker images that are running after updating the free5GC codebase to allow PQC.

In our custom implementation of free5GC, each VNF container is running a lightweight version of Ubuntu 22.04 Jammy with OpenSSL v3.0.2 installed. After compiling oqs-provider from source for the respective platform, the resulting shared library file is installed in each VNF Docker container, enabling PQC integration at a system level.

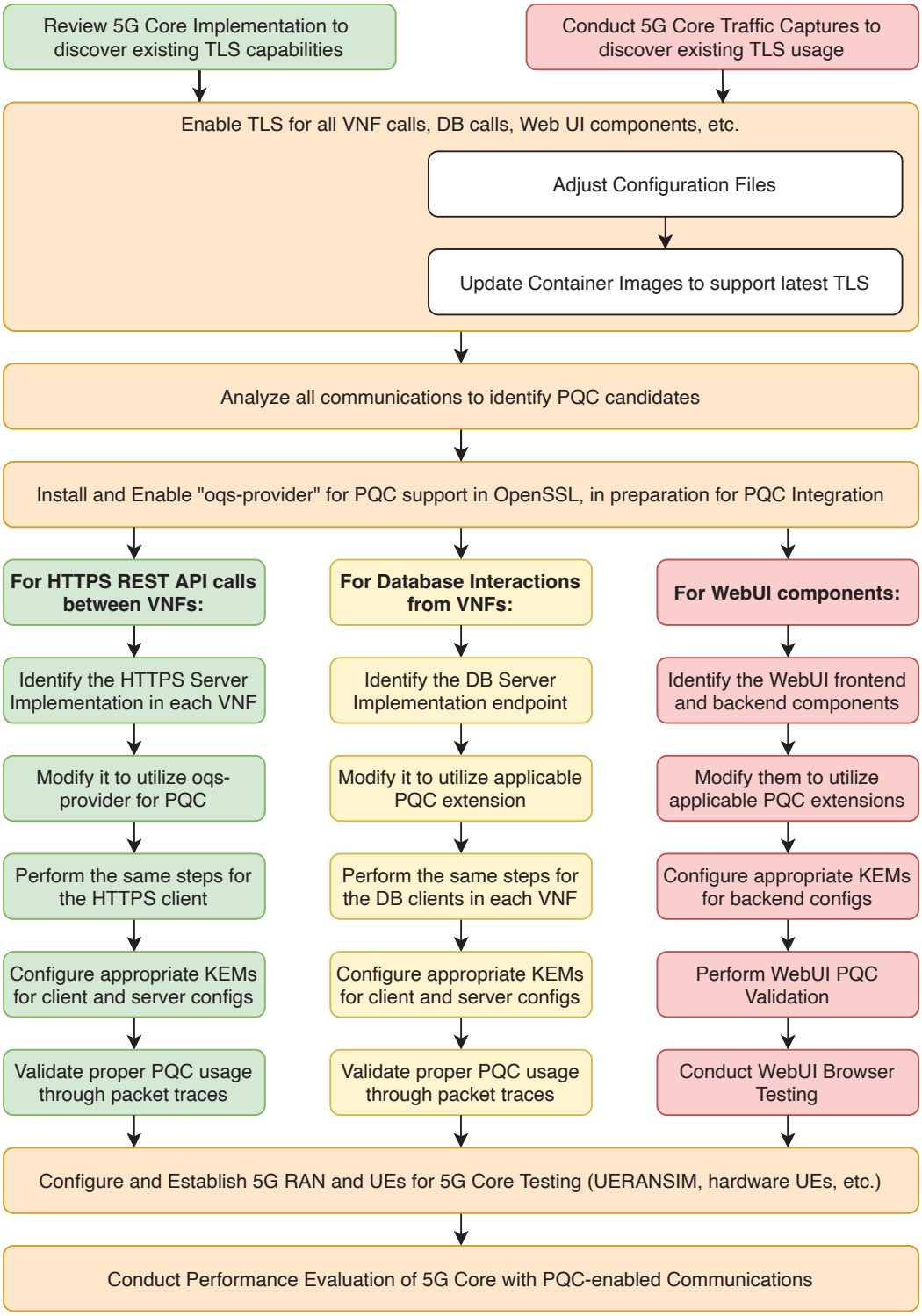


Figure 3. Steps taken in the discovery, integration, validation, and analysis of PQC in 5G core networks.

```
55  ssl_conf = ssl_sect
56
57  # List of providers to load
58  [provider_sect]
59  default = default_sect
60  oqsprovider = oqsprovider_sect
61
74  [default_sect]
75  activate = 1
76
77  [oqsprovider_sect]
78  activate = 1
79
80
81  ####################################################################
82  [ ca ]
83  default_ca  = CA_default        # The default ca section
398 [ssl_sect]
399 system_default = system_default_sect
400
401 [system_default_sect]
402 CipherString = DEFAULT:@SECLEVEL=2
403 Groups = kyber1024
404
```

Figure 4. Openssl.cnf with the edited files highlighted to enable OQS.

```
tel-mfoss-core@telmfosscore:~/free5gc-compose$ docker ps
CONTAINER ID   IMAGE                          COMMAND                  CREATED          STATUS          PORTS                                                  NAMES
3fd31542139 3  free5gc-compose-free5gc-n3iwf  "sh -c './n3iwf-ipse…"   25 seconds ago   Up 23 seconds                                                          n3iwf
bdf60e32637a   free5gc-compose-free5gc-chf    "./chf -c ./config/c…"   25 seconds ago   Up 23 seconds   2122/tcp, 8000/tcp                                     chf
d68f6dd75aa8   free5gc-compose-free5gc-pcf    "./pcf -c ./config/p…"   25 seconds ago   Up 24 seconds   8000/tcp                                               pcf
6d3bac1d3324   free5gc-compose-free5gc-udm    "./udm -c ./config/u…"   25 seconds ago   Up 24 seconds   8000/tcp                                               udm
6211d78496e5   free5gc-compose-free5gc-amf    "./amf -c ./config/a…"   25 seconds ago   Up 24 seconds   8000/tcp                                               amf
cfe547e0683d   free5gc-compose-free5gc-udr    "./udr -c ./config/u…"   25 seconds ago   Up 24 seconds   8000/tcp                                               udr
8c6d72eb61c4   free5gc-compose-free5gc-webui  "./webui -c ./config…"   25 seconds ago   Up 24 seconds   2121/tcp, 0.0.0.0:5000->5000/tcp, :::5000->5000/tcp    webui
f328745ce0ec   free5gc-compose-free5gc-smf    "./smf -c ./config/s…"   25 seconds ago   Up 24 seconds   8000/tcp                                               smf
e9fad4753faf   free5gc-compose-free5gc-ausf   "./ausf -c ./config/…"   25 seconds ago   Up 24 seconds   8000/tcp                                               ausf
8af0c1ff8ab3   free5gc-compose-free5gc-nssf   "./nssf -c ./config/…"   25 seconds ago   Up 24 seconds   8000/tcp                                               nssf
4e6d358fdf33   free5gc-compose-free5gc-nrf    "./nrf -c ./config/n…"   25 seconds ago   Up 24 seconds   8000/tcp                                               nrf
461ba83b2d3b   pqc-mongo                      "docker-entrypoint.s…"   25 seconds ago   Up 25 seconds   27017/tcp                                              mongodb
f2a75e2ac510   free5gc-compose-free5gc-upf    "bash -c './upf-ipta…"   25 seconds ago   Up 25 seconds                                                          upf
```

Figure 5. Essential PQC free5GC CN VNFs up during runtime.

Additionally, some modifications to the source of free5GC are necessary in order to utilize OpenSSL. Virtually all aspects of free5GC are written in GoLang [43], which lacks direct native support for using OpenSSL. A Go program traditionally uses Go's own TLS and Crypto libraries to handle TLS communications [44]. To address this challenge, we selected the open-source Go-Openssl wrapper from 'pexip' [45] and incorporated it within each VNF to enable the use of OpenSSL for cryptography within free5GC rather than Go's crypto library. At the time of publication of this paper, the custom CN implementation is built using free5GC v3.4.2. The integration of Go-Openssl focuses on two main areas within each VNF in free5GC's source code: server.go and the OpenAPI system. Within server.go, the network function is typically utilized to create a TCP listener, which subsequently utilizes the Go TLS to secure the HTTP connection on the server side. Listing 2 in Appendix A shows a code snippet of the two functions that are amended within the code base, in this specific case for 5G's the "Access and Mobility Management Function" (AMF) VNF. In our PQC-enabled implementation, if encryption (https) is enabled in free5GC's YAML configuration file, then an OpenSSL-enabled listener is created within the 'NewServer' function. The newly created server uses the OpenSSL context that is stored in the server structure from the logic in the NewServer function. The built-in functions are used to obtain the path to the certificates used for context creation, and certificate verification is set to "none" for testing purposes.

Next, our focus was on enabling each of the HTTPS clients used throughout free5GC for API calls to similarly be enabled for OpenSSL usage. This process was not as straightforward as the corresponding server-side effort. Free5GC utilizes an OpenAPI-driven approach to API implementation and handling of client HTTP requests. Analyzing the call flow across multiple different files that include client attributes revealed that, ultimately, they all utilize a call to the CallAPI() function within the free5GC OpenAPI repository,

which is located in the client.go file. In this function, an OpenSSL context is created using two specific functions: CreateOpenSSLClientCtx() and GetCurrentNetworkFunctionUsingAPI(). The OpenSSL context is managed in the same manner as the server. However, in the case of a client, there are no built-in functions to obtain the path to the client certificates. Client certificates are not strictly required for TLS connectivity. However, they are needed when mutual TLS authentication (mTLS) is desired, which can provide a significant improvement in security between both client and server. Thus, GetCurrentNetworkFunctionUsingAPI() obtains the specific network function that is making a request by parsing the stack trace. Returned from this function is effectively the name of the VNF making the HTTP request, which is appended to the path where the certificates are stored inside the Docker containers. In either scenario, if the configuration includes a custom HTTPClient setup requirement, a custom client transport method that includes OpenSSL support is added. The logic sequence for this process is shown in CreateOpenSSLInnerClientTransport(). The dialTLS function that handles the transport protocols for an HTTP client in Go is overridden to use the OpenSSL dial wrapper. The request is then made using the OpenSSL-enabled dialTLS() function during the invocation of the Do() method. This process is shown in the code Listing 3 contained in Appendix A.

One major goal of this customization was to make it as seamless as possible for a free5GC administrator to install and operate it with support for PQC. With the extensible nature of Docker, this goal was met while maintaining the same basic structure for downloading and installing our PQC-enabled free5GC build. The user only needs to run one additional command to install MongoDB from source, which for convenience, has been incorporated into a single shell script. The steps to clone, build, install, and run PQC-enabled free5GC that are used by this script are also shown in Listing 4 in Appendix A.

Once we established a full integration of PQC into all HTTPS REST API endpoints (server code) and all corresponding client calls to these APIs, we could proceed to collect results that demonstrate the feasibility of this integration of PQC into free5GC operations and also quantify the impact on latency and bandwidth usage, which are shown and discussed in the next section.

6. Results

The results from this study aim to evaluate the impact of using standard Elliptic Curve Cryptography (ECC), PQC, and Hybrid PQC KEMs within the communications of the VNFs in a 5G CN using free5GC. All of the CRYSTALS-Kyber KEMs displayed in Table 3 are included in this evaluation, along with the classical ECC KEM X25519. Furthermore, these different methods are compared to a CN that does not use TLS to secure the 5G CN API calls and instead uses unencrypted HTTP calls. This acts as the baseline for our tests. All of them were conducted on a Dell (Dell Inc., Round Rock, TX, USA) Server equipped with an AMD (AMD Inc., Santa Clara, CA, USA) Epyc 7313p 16-Core, 32-Thread CPU, running the aforementioned custom implementation of free5GC. Included in this test environment is another computer, which is responsible for running UERANSIM [46], a tool included in the free5GC package to simulate a gNB as well as multiple UE connections in accordance with 3GPP standards. Any performance metrics of the UERANSIM host machine are inconsequential to this evaluation since all the signaling requests and communications activities are measured relative to the TLS activity within the CN machine. Both machines are running Ubuntu 22.04 LTS.

6.1. Testing Scenarios

Each test is conducted with 100 total UE connections in order to maintain statistically relevant, consistent, and comparable results. Each UE is scheduled with a seven-second cadence between connections, allowing the CN sufficient time to complete each client attachment and to be able to adequately capture and measure each TLS handshake. The data gathered from these studies represent the varying total UE connection times between the different KEMs available in this system. Prior hypotheses were that the PQC KEMs

would result in a significant additional latency during a UE connection, but as our results clearly show, this is not the case. In our testing, all connections use a consistent cipher, AES_256_GCM_SHA384, to maintain high security and uniform data communication latency results. Due to OpenSSL typically only accepting signed certificates from a trusted Certificate Authority (CA), all tests are conducted with OpenSSL set to skip certificate verification. This is due to the fact that free5GC includes self-signed certificates. However, it is highly recommended for production systems to use certificates that are signed by a trusted CA. The use of certificates from a CA could increase the latency of the UE connections, due to the short additional time required for verification, but this additional time would be true for any KEM that is used—traditional or PQC.

Wireshark is utilized to capture packets within the CN, displaying the communications between the VNFs. Data from these packets can be extrapolated to further understand connection times, bytes sent, etc. Prior to each test, one KEM is enabled in each of the VNFs openssl.cnf file, ensuring proper control over the test environment. The KEM that is under test is also validated in the final Wireshark capture, which can be seen during the client or server hello message under the 'key_share' extension field. Then, the CN is brought up, and the gNB connection is established. Once all of these processes are complete, Wireshark begins capturing packets on the free5GC-br network interface. Once Wireshark begins capturing, the UERANSIM program is executed, with inputs configured for 100 total UE connections and a seven-second cadence between each subsequent connection. Each UE connects in its own time interval in order to gain a large dataset to understand the average UE connection time with each respective KEM active in the CN, and other information which is analyzed in the next section.

6.2. Data Analysis

One important test we conducted focuses on the measurement and evaluation of the total client connection setup duration, which represents the time from the UE initially requesting a connection via the gNodeB until the UE can send user data through the 5G connection. Figure 6 shows a box plot of the different KEMs and their respective UE connection times that we obtained. The families of algorithms are color coded in Figures 6–9. We observe that the baseline connection time when using no encryption (plain-text HTTP requests) is slightly less than 0.5 seconds. This is expected due to the fact that there are significantly fewer computations required when we are not encrypting and decrypting each communication request.

It is important to stress, however, that we conducted this analysis solely for baseline purposes. In no way do we recommend implementing or operating a 5G Core Network using unencrypted communications, as this would pose a significant and unnecessary risk to user data and control data. As a secondary baseline, therefore, we are utilizing the X25519 KEM. This method is a commonly used ECC KEM known for its high performance and adequate security, utilizing a key size of 256 bits. It can be seen that the hybrid X25519 algorithms are nearly identical in total connection time compared to the regular Kyber KEMs, in terms of additional latency. The P256, P384, and P521 hybrid variants, however, do add significantly more latency to the handshake and total connection process times.

Figure 7 outlines box plots that represent the amount of data, in Megabytes, transmitted in the CN during an attachment of a UE. Showing a similar outcome to the results in Figure 6, the tests conducted and shown here for unencrypted communications and the X25519 KEM once again show the lowest amount of bytes exchanged. With the understanding that encryption by nature adds more data to a packet, along with more packets overall being required due to the addition of the TLS handshakes, it is no surprise that there is significantly less data transferred during the unencrypted baseline. Likewise, X25519 uses a 32-byte key, which is the smallest key size compared to other PQC and Hybrid PQC KEMs. It is obvious from these results that with this increase in key size we observe an increase in the required data being exchanged. For example, kyber512 has a median data volume of 1.12 MB, and kyber1024, with double the public key and cipher text size as seen

in Table 4, is responsible for an increase of nearly 200 KB of data, which are sent during the UE connection process.

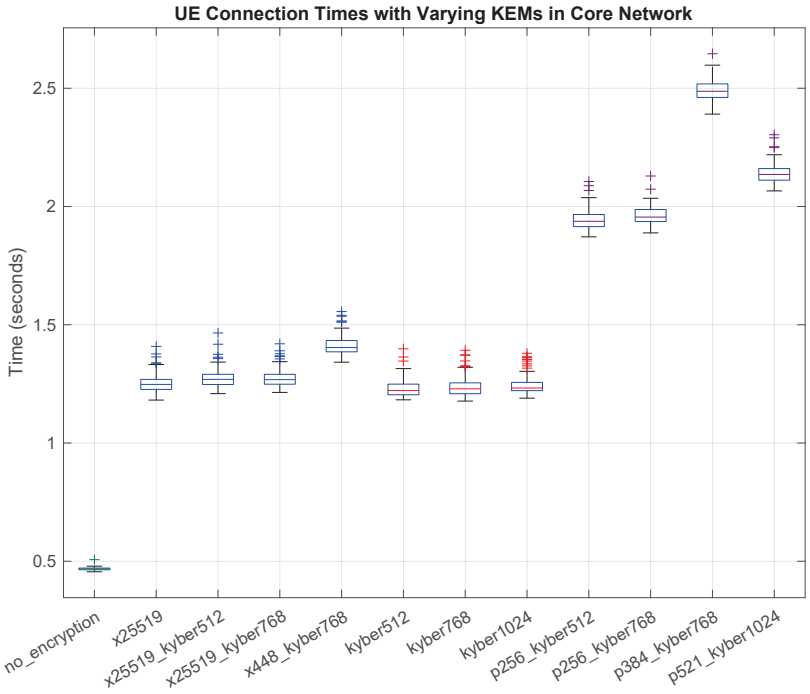


Figure 6. UE Connection times measured with no encryption, X25519, PQC and Hybrid PQC KEMs in the CN.

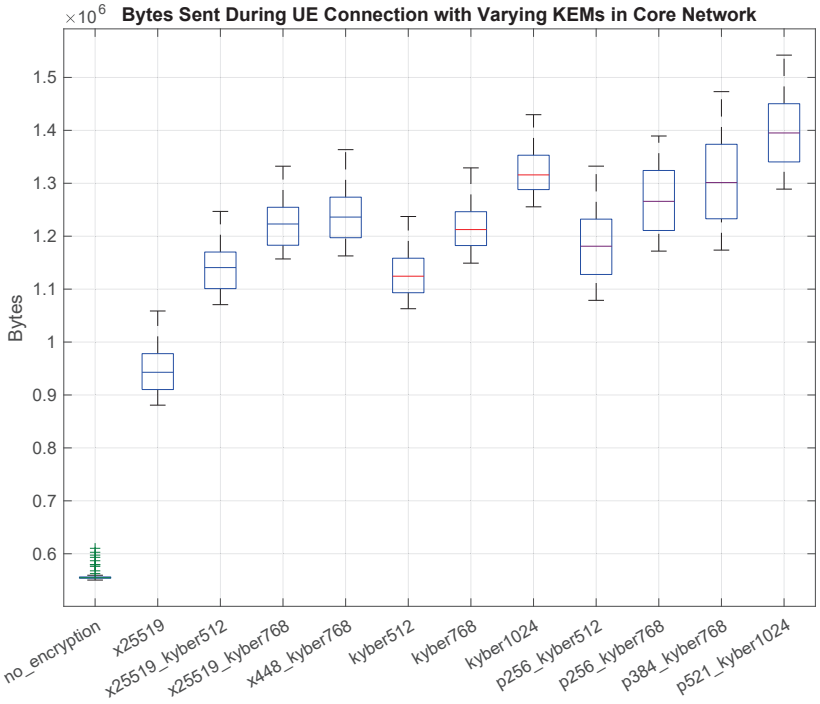


Figure 7. Measurement of Megabytes of data transmitted in the CN during a UE connection while varying the KEM used in TLS 1.3 within the CN.

Table 4. Key sizes and cycle counts for different versions of Kyber on Haswell CPUs [47].

Kyber-512	**Sizes (In Bytes)**	**Haswell Cycles (Ref)**	**Haswell Cycles (avx2)**
sk:	1632	gen: 122,684	gen: 33,856
pk:	800	enc: 145,424	enc: 45,200
ct:	768	dec: 187,960	dec: 34,572
Kyber-768	**Sizes (In Bytes)**	**Haswell Cycles (Ref)**	**Haswell Cycles (avx2)**
sk:	2400	gen: 199,408	gen: 52,732
pk:	1184	enc: 235,260	enc: 67,624
ct:	1088	dec: 274,900	dec: 53,156
Kyber-1024	**Sizes (In Bytes)**	**Haswell Cycles (Ref)**	**Haswell Cycles (avx2)**
sk:	3168	gen: 307,148	gen: 73,544
pk:	1568	enc: 346,648	enc: 97,324
ct:	1568	dec: 396,584	dec: 79,128

Legend: sk—Secret Key, pk—Public Key, ct—Ciphertext.

Further correlation between larger key sizes and bytes exchanged can also be observed throughout the dataset collected for the Hybrid PQC KEM results. We hypothesize that the observed variance in the data is due to ongoing internal CN communication that can vary from time to time, thus creating some tests that record a differing amount of bytes during a UE connection operation. Hybrid PQC KEMs are specifically important in the role of quantum-resistant algorithms. In a blog post from Google engineers, the authors state the two main rationale points for using Hybrid KEMs is 'Mitigating risks associated with new algorithms' and 'A safe transition' to PQC [48]. Hybrid KEMs provide the security of PQC combined with the trust afforded to ECC. Thus, many systems are expected to adopt the hybrid approach, and the data presented on the network performance of hybrid PQC KEMs in 5G are promising.

Figures 8 and 9 illustrate the relative differences in time between the different KEM algorithms used during UE connection operations in relation to the unencrypted operation and to the ECC X25519 KEM, respectively. For example, Figure 9 takes the average UE connection time of each tested PQC and hybrid PQC KEM and subtracts the average UE connection time of the standard ECC X25519 KEM. A similar approach is used in Figure 8, subtracting the connection time we observed for unencrypted 5G core operations from the obtained PQC and hybrid PQC results. The previously mentioned trends are extrapolated here to give a clearer view of the relative time differences. We would also like to specifically point out the performance shown for p384_kyber768, as it is nearly 50% slower than its counterparts. Our hypothesis is that the cause of this outcome is the lack of optimization capability for the specific P-384 algorithm in software, combined with additional computational complexities when working with that specific KEM. Figure 9 notes the time differences compared to the classical ECC X25519 KEM. Interestingly, all of the standalone Kyber KEMs performed better than the baseline ECC X25519 KEM. This could be due to heightened optimizations within the OQS OpenSSL library, along with using an AVX2-enabled machine.

In order to better understand the complexity of a UE connection operation, we wanted to show the entire process and all API calls required to complete this process. Thus, in Figure A1, shown in Appendix A, we are showing a detailed call sequence for 1 entire UE connection sequence, observing a total of 129 TLSv1.3 handshakes. Contained in this capture is each control signal sent or received by the gNB, along with a condensed version of the TLS client hello messages. Column 'Qty' details how many duplicate consecutive packets were sent. For example, during the first TLS handshakes of the UE connection, the AMF sends three separate requests to the NRF. This figure is helpful in that it quantifies

how many TLS handshakes occur during an attachment and also provides more insights into the duration of each handshake.

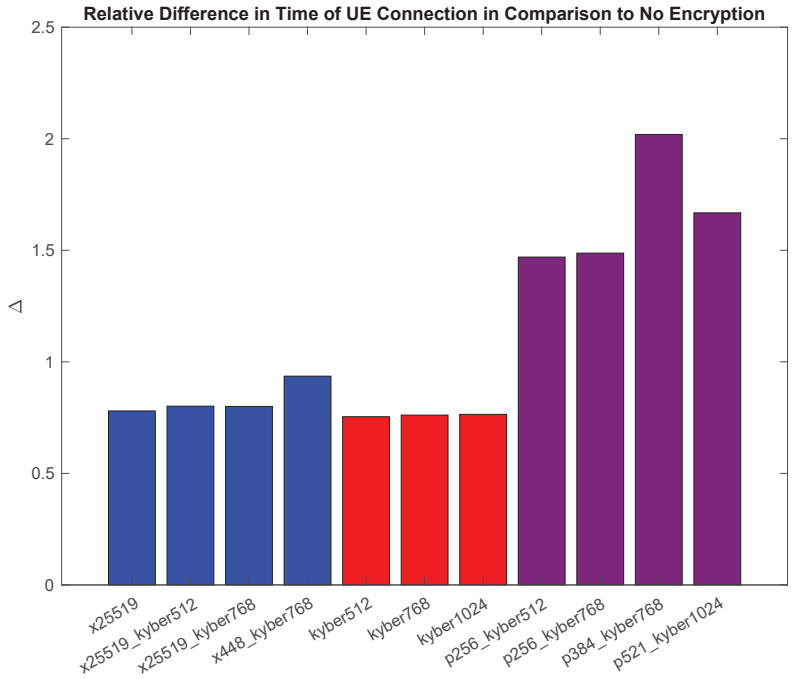


Figure 8. Difference in unit time of the KEM latency compared to no encryption.

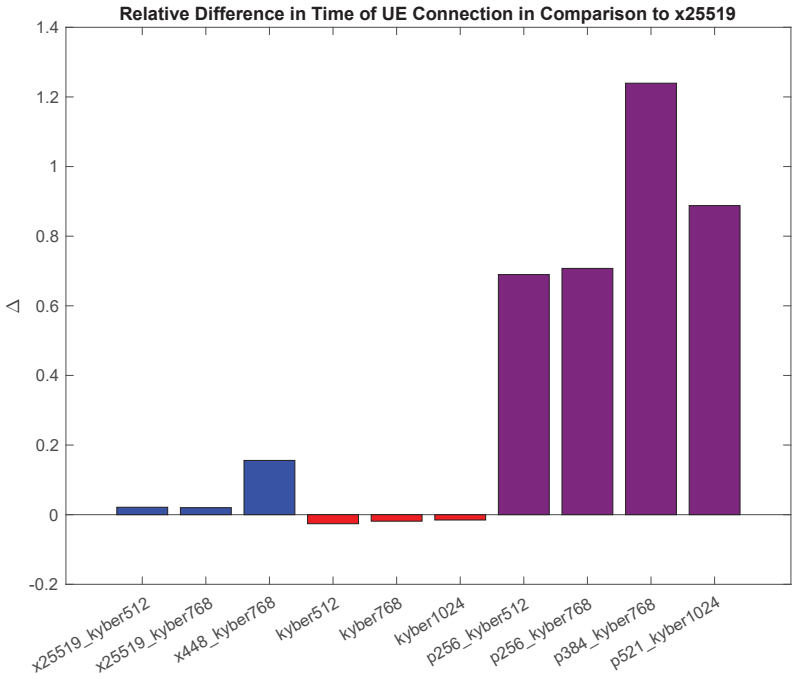


Figure 9. Difference in unit time of the KEM latency compared to X25519.

Further information on the time it takes to complete a transaction is illustrated in Figure 10. The differences in handshake time between p521_kyber1024 and plain kyber1024 demonstrate that the additional latency shown in Figure 9 is indeed attributable solely to the TLSv1.3 handshake process, which includes the KEM.

Burst connections of multiple UEs simultaneously are part of our future work, where the focus is on the scalability of both the 5G core provided by free5Gc and also of the 5G core VNFs' communications using PQC KEMs. With the understanding that Massive Machine-Type Communications (mMTC) could potentially need to support one million devices per km^2, situations in which devices are connecting and disconnecting quickly could have a real impact on performance in the CN with PQC enabled [49]. Kyber1024 has nearly 400 KB or 3.2 Mb more data transferred during a single UE connection, and in the extreme case of 1 million devices connecting, 400 GB or 3.2 Tb of data would be transferred in this scenario. Therefore, the systems employed for CN operations would need to withstand these connection bursts, with adequate load balancing and compute resources. Scalability is of particular focus also in commercial deployments, where the concern is not about adequate resource availability but about the cost for those resources, thereby driving the need for efficiency.

p521_kyber1024 TLSv1.3 Handshake

No.	Time	Source	Destination	Protocol	Length	Info
1	0.000000000	AMF	NRF	TLSv1.3	2026	Client Hello (SNI=nrf.free5gc.org)
2	0.000020930	NRF	AMF	TCP	66	irdmi(8000) > 50264 [ACK] Seq=1 Ack=1961 Win=501 Len=0 TSval=1634323114 TSecr=603203809
3	0.003669746	NRF	AMF	TLSv1.3	3164	Server Hello, Change Cipher Spec, Application Data, Application Data, Application Data, Application Data
4	0.003691688	AMF	NRF	TCP	66	50264 > irdmi(8000) [ACK] Seq=1961 Ack=3099 Win=497 Len=0 TSval=603203812 TSecr=1634323117
5	0.008226219	AMF	NRF	TLSv1.3	146	Change Cipher Spec, Application Data
6	0.008327794	AMF	NRF	TLSv1.3	429	Application Data

kyber1024 TLSv1.3 Handshake

No.	Time	Source	Destination	Protocol	Length	Info
1	0.000000000	AMF	NRF	TLSv1.3	1893	Client Hello (SNI=nrf.free5gc.org)
2	0.000016713	NRF	AMF	TCP	66	irdmi(8000) > 56616 [ACK] Seq=1 Ack=1828 Win=501 Len=0 TSval=1541772042 TSecr=510652737
3	0.000952064	NRF	AMF	TLSv1.3	3031	Server Hello, Change Cipher Spec, Application Data, Application Data, Application Data, Application Data
4	0.000967093	AMF	NRF	TCP	66	56616 > irdmi(8000) [ACK] Seq=1828 Ack=2966 Win=497 Len=0 TSval=510652738 TSecr=1541772043
5	0.001857607	AMF	NRF	TLSv1.3	146	Change Cipher Spec, Application Data
6	0.001940036	AMF	NRF	TLSv1.3	429	Application Data

Figure 10. Example packet captures with relative timestamps of TLS handshakes using p521_kyber1024 (**top**) and plain kyber1024 (**bottom**).

The results we collected and analyzed in this section clearly indicate that we can accomplish the integration and use of post-quantum cryptography KEM for TLS with minimal additional overhead. Our results also demonstrate that it is feasible and practical to incorporate PQC into a 5G core network, for use in the communication between VNFs, databases, etc. Our evaluation methodology successfully validates this conclusion through accurate, precise, and verifiable data. This novel assessment of PQC KEMs within a 5G CN environment—to the best of our knowledge, the first of its kind—thus makes a strong case for the inclusion of these PQC algorithms within 5G networks to strengthen its cybersecurity.

7. Conclusions

Currently, 5G continues its worldwide rollout, allowing millions of people worldwide to have access to high-speed cellular service. At the same time, the need for customer data privacy protections and strong operational cybersecurity is rapidly increasing due to the growing sophistication and reach of cyberattacks in recent years. As such, we need to protect all aspects of today's technology systems, especially 5G cellular services and their importance to our modern information society. We also need to be looking ahead at the cybersecurity requirements for the upcoming 6G, and beyond. Future work in this area can expand upon the tests presented in this paper, with further testing using physical UEs, different channel bandwidths, testing with commercial grade equipment, and expanding PQC to end-to-end communication channels. Utilizing physical UEs with differing channel bandwidths may provide an interesting insight into the device implementation differences between different devices, which may be resource-constrained.

With the rising possibility of usable quantum computers in the near future and the threats of SNDL attacks, 5G and 6G must adapt to PQC in order to keep customer data

quantum-safe. The research we present in this paper demonstrates that the modifications we conducted to integrate PQC into free5GC resulted in only a negligible latency impact and a minimal increase in data transfer volume during the UE connection phase to the 5G network from the inclusion of certain PQC or hybrid PQC KEMs. Our efforts to integrate and analyze PQC into a 5G CN is a new contribution to the cybersecurity domain during the ongoing migration from classical KEMs to PQC KEMs, and provides important insights into vendors and cybersecurity professionals alike. These positive test outcomes were collected by capturing and analyzing a large number of 5G core network communications utilizing TLS v1.3 with PQC KEMs during numerous UE attachments using a variety of different KEMs. The results indicate that standalone Kyber KEMs can actually outperform the traditional ECC X25519 KEM in terms of the UE client connection duration tests. Thus, we can conclude that without any UE latency impact, we can achieve a significantly strengthened cybersecurity posture for the entire 5G core network. With these modifications, every API call between VNFs is not only AES-256 protected, which is currently considered quantum-safe, but we also achieved strong protections for the key exchange during the TLS handshaking for quantum safety. Thus, based on our tests, we anticipate similarly encouraging results regarding the latency and bandwidth impact of introducing PQC into TLS connections for all control plane and user plane operations in 5G and beyond. Our future work encompasses significant steps in that direction, with efforts already underway to secure RAN communications, multi-access edge communications in 5G, and end-to-end user data protections for 5G UEs.

Author Contributions: Investigation, P.S., R.G., M.B., M.H. and H.S.; writing—original draft preparation, P.S., R.G., M.B. and M.H.; writing—review and editing, P.S., R.G., M.B., M.H. and H.S.; supervision, H.S. and M.H.; project administration, H.S. and M.H.; funding acquisition, H.S. and M.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research has been supported by the University of Nebraska-Lincoln's Advanced Telecommunications Engineering Laboratory (TEL) and Belcan Government Solutions (BGS) under grant# TEL-Belcan-5G23.

Data Availability Statement: The original data presented in the study are openly available in: https://git.unl.edu/tel-papers/2024_pqc_for_securing_5gcn_mdpi (accessed on 26 September 2024).

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

Listing 1. Dockerfile snippet of commands for building free5GC containers with OQS.

```
RUN apt update && apt upgrade -y
RUN apt install -y git gcc cmake ninja-build openssl libssl-dev
RUN git clone https://github.com/open-quantum-safe/oqs-provider.git
RUN cd oqs-provider
WORKDIR oqs-provider
RUN sh ./scripts/fullbuild.sh
RUN cmake --install _build
WORKDIR /free5gc
```

Listing 2. Creation of HTTP server using Go-OpenSSL Wrapper.

```
import "github.com/pexip/go-openssl"
...
func NewServer(amf ServerAmf, tlsKeyLogPath string) (*Server, error){
...
s.tlsCtx, err = openssl.NewCtxFromFiles(cfg.GetCertPemPath(), cfg.GetCertKeyPath
    ())
if err != nil {
logger.InitLog.Errorf("Initialize OpenSSL context failed: %v", err)
return nil, err
}
s.tlsCtx.SetVerify(openssl.VerifyNone, nil)
...
}
...
func (s *Server) startServer(wg *sync.WaitGroup){
...
} else if scheme == "https" {
listener, err := openssl.Listen("tcp", cfg.GetSbiBindingAddr(), s.tlsCtx)
if err != nil {
logger.SBILog.Errorf("Failed to create OpenSSL listener: %v", err)
return
}
err = s.httpServer.Serve(listener)
} else {
err = fmt.Errorf("no support this scheme[%s]", scheme)
}
...
}
```

Listing 3. Creation of HTTP client using Go-OpenSSL Wrapper.

```
import "github.com/pexip/go-openssl"
...
var (
innerOpenSSLClient     *http.Client
innerOpenSSLClientCtx *openssl.Ctx
)
...
func GetCurrentNetworkFunctionUsingAPI() string {
var buf bytes.Buffer
buf.Write(debug.Stack())
var trace = buf.String()
traceLines := strings.Split(trace, "\n")
pathPattern := regexp.MustCompile('/go/src/free5gc/NFs/[^/]+/')
nfPath := regexp.MustCompile('/go/src/free5gc/NFs/([^/]+)/')
webuiPath := "/go/src/free5gc/webconsole"

for _, line := range traceLines {
isolatedPath := pathPattern.FindString(line)
if isolatedPath != "" {
nf := nfPath.FindStringSubmatch(isolatedPath)
if len(nf) > 1 {
return nf[1]
}
}
}

// if it is not a NF using the API, check for the webui
if strings.Contains(trace, webuiPath) {
return "webui"
}
// if nothing is found:
fmt.Println("Couldn't find the specified path or network function. Trace:")
fmt.Println(trace)
return ""
}
```

```
...
func CreateOpenSSLInnerClientTransport() error {
// Custom dial function to use OpenSSL for TLS connections
dialTLS := func(network, addr string) (net.Conn, error) {
conn, err := openssl.Dial(network, addr, innerOpenSSLClientCtx, openssl.
    InsecureSkipHostVerification)
if err != nil {
return nil, err
}
return conn, nil
}
// Create a custom transport using the custom dial function
tr := &http.Transport{
TLSClientConfig: &tls.Config{
InsecureSkipVerify: true, // Skip certificate verification,
},
TLSNextProto: make(map[string]func(authority string, c *tls.Conn) http.
    RoundTripper),
DialTLS:      dialTLS,
}
innerOpenSSLClient = &http.Client{Transport: tr}
return nil
}
...
func CallAPI(cfg Configuration, request *http.Request) (*http.Response, error) {
opensslCtx := CreateOpenSSLClientCtx(GetCurrentNetworkFunctionUsingAPI())
if opensslCtx == nil {
return nil, fmt.Errorf("Error Creating OpenSSL Context")
}
if cfg.HTTPClient() != nil {
CreateOpenSSLClientTransportWithCfg(cfg, opensslCtx)
return cfg.HTTPClient().Do(request)
}
if request.URL.Scheme == "https" {
innerOpenSSLClientCtx = opensslCtx
CreateOpenSSLInnerClientTransport()
return innerOpenSSLClient.Do(request)
...
return nil, fmt.Errorf("unsupported scheme[%s]", request.URL.Scheme)
}
```

Listing 4. Shell commands for building and installing PQC enabled free5GC.

```
# Clone the project
git clone https://git.unl.edu/tel-papers/2024_pqc_for_securing_5gcn_mdpi/free5gc
    -pqc.git
cd free5gc-pqc
# Build the mongodb from source
./install_mongo.sh
# Build the images
make all
docker compose -f docker-compose-build.yaml build
# Alternatively you can build specific NF image e.g.:
make amf
docker compose -f docker-compose-build.yaml build free5gc-amf
```

#	Time	Qty	Client	Dest	Protocol	Info
1	0.000	1	gNB	AMF	NGAP/ NAS-5GS	InitialUEMessage, Registration request
2	0.006	3	AMF	NRF	TLSv1.3	Client Hello
5	0.032	1	AMF	AUSF	TLSv1.3	Client Hello
6	0.039	3	AUSF	NRF	TLSv1.3	Client Hello
9	0.065	1	AUSF	UDM	TLSv1.3	Client Hello
10	0.074	3	UDM	NRF	TLSv1.3	Client Hello
13	0.100	2	UDM	UDR	TLSv1.3	Client Hello
15	0.114	1	AMF	gNB	NGAP/ NAS-5GS	SACK (Ack=0, Arwnd=106496) , DownlinkNASTransport , Authentication request
16	0.115	1	gNB	AMF	NGAP/ NAS-5GS	SACK (Ack=0, Arwnd=106496) , UplinkNASTransport, Authentication response
17	0.123	1	AMF	NRF	TLSv1.3	Client Hello
18	0.132	1	AMF	AUSF	TLSv1.3	Client Hello
19	0.139	1	AUSF	NRF	TLSv1.3	Client Hello
20	0.148	1	AUSF	UDM	TLSv1.3	Client Hello
21	0.156	1	UDM	NRF	TLSv1.3	Client Hello
22	0.165	1	UDM	UDR	TLSv1.3	Client Hello
23	0.169	1	AMF	gNB	NGAP/ NAS-5GS	SACK (Ack=1, Arwnd=106496) , DownlinkNASTransport , Security mode command
24	0.170	1	gNB	AMF	NGAP/ NAS-5GS	SACK (Ack=1, Arwnd=106496) , UplinkNASTransport, Security mode complete, Registration request
25	0.176	3	AMF	NRF	TLSv1.3	Client Hello
28	0.203	1	AMF	UDM	TLSv1.3	Client Hello
29	0.210	1	UDM	NRF	TLSv1.3	Client Hello
30	0.219	1	UDM	UDR	TLSv1.3	Client Hello
31	0.229	3	AMF	NRF	TLSv1.3	Client Hello
34	0.255	1	AMF	UDM	TLSv1.3	Client Hello
35	0.262	1	UDM	NRF	TLSv1.3	Client Hello
36	0.271	1	UDM	UDR	TLSv1.3	Client Hello
37	0.280	1	AMF	NRF	TLSv1.3	Client Hello
38	0.292	1	AMF	UDM	TLSv1.3	Client Hello
39	0.299	1	UDM	NRF	TLSv1.3	Client Hello
40	0.308	1	UDM	UDR	TLSv1.3	Client Hello
41	0.317	1	AMF	NRF	TLSv1.3	Client Hello
42	0.326	1	AMF	UDM	TLSv1.3	Client Hello
43	0.333	1	UDM	NRF	TLSv1.3	Client Hello
44	0.343	1	UDM	UDR	TLSv1.3	Client Hello
45	0.351	1	AMF	NRF	TLSv1.3	Client Hello
46	0.361	1	AMF	UDM	TLSv1.3	Client Hello
47	0.368	1	UDM	NRF	TLSv1.3	Client Hello
48	0.377	1	UDM	UDR	TLSv1.3	Client Hello
49	0.385	1	AMF	NRF	TLSv1.3	Client Hello
50	0.394	1	AMF	UDM	TLSv1.3	Client Hello
51	0.401	1	UDM	NRF	TLSv1.3	Client Hello
52	0.411	1	UDM	UDR	TLSv1.3	Client Hello
53	0.418	3	AMF	NRF	TLSv1.3	Client Hello
56	0.445	1	AMF	PCF	TLSv1.3	Client Hello
57	0.454	3	PCF	NRF	TLSv1.3	Client Hello
60	0.480	1	PCF	UDR	TLSv1.3	Client Hello
61	0.489	3	PCF	NRF	TLSv1.3	Client Hello
64	0.515	1	PCF	AMF	TLSv1.3	Client Hello
65	0.518	1	AMF	gNB	NGAP/ NAS-5GS	InitialContextSetupRequest, Registration accept
66	0.518	1	gNB	AMF	NGAP	SACK (Ack=2, Arwnd=106496) , InitialContextSetupResponse
67	0.723	1	gNB	AMF	NGAP/ NAS-5GS	UplinkNASTransport, Registration complete, UplinkNASTransport, UL NAS transport, PDU session establishment request, UplinkNASTransport, UL NAS transport, PDU session establishment request
68	0.723	1	AMF	gNB	NGAP/ NAS-5GS	Downlink NAS Transport, Configuration update command
69	0.729	3	AMF	NRF	TLSv1.3	Client Hello
72	0.787	1	AMF	NSSF	TLSv1.3	Client Hello
73	0.795	3	AMF	NRF	TLSv1.3	Client Hello
76	0.819	1	AMF	SMF	TLSv1.3	Client Hello
77	0.829	3	SMF	NRF	TLSv1.3	Client Hello
80	0.857	1	SMF	UDM	TLSv1.3	Client Hello
81	0.865	1	UDM	NRF	TLSv1.3	Client Hello
82	0.874	1	UDM	UDR	TLSv1.3	Client Hello
83	0.884	5	SMF	NRF	TLSv1.3	Client Hello
88	0.931	1	SMF	PCF	TLSv1.3	Client Hello
89	0.940	1	PCF	NRF	TLSv1.3	Client Hello
90	0.950	1	PCF	UDR	TLSv1.3	Client Hello
91	0.972	1	PCF	NRF	TLSv1.3	Client Hello
92	0.981	1	PCF	UDR	TLSv1.3	Client Hello
93	0.989	1	PCF	NRF	TLSv1.3	Client Hello
94	0.998	1	PCF	UDR	TLSv1.3	Client Hello
95	1.006	2	PCF	NRF	TLSv1.3	Client Hello
97	1.024	3	SMF	NRF	TLSv1.3	Client Hello
100	1.053	1	SMF	CHF	TLSv1.3	Client Hello
101	1.069	1	AMF	NRF	TLSv1.3	Client Hello
102	1.072	1	SMF	NRF	TLSv1.3	Client Hello
103	1.079	1	AMF	NSSF	TLSv1.3	Client Hello
104	1.081	1	SMF	AMF	TLSv1.3	Client Hello
105	1.084	1	AMF	gNB	NGAP/ NAS-5GS	PDU session establishment accept
106	1.086	1	gNB	AMF	NGAP	PDU Session Resource Setup Response
107	1.087	3	AMF	NRF	TLSv1.3	Client Hello
110	1.112	1	AMF	SMF	TLSv1.3	Client Hello
111	1.119	3	SMF	NRF	TLSv1.3	Client Hello
114	1.146	1	SMF	UDM	TLSv1.3	Client Hello
115	1.153	1	UDM	NRF	TLSv1.3	Client Hello
116	1.162	1	UDM	UDR	TLSv1.3	Client Hello
117	1.171	5	SMF	NRF	TLSv1.3	Client Hello
122	1.217	1	SMF	PCF	TLSv1.3	Client Hello
123	1.224	1	PCF	NRF	TLSv1.3	Client Hello
124	1.234	1	PCF	UDR	TLSv1.3	Client Hello
125	1.251	1	PCF	NRF	TLSv1.3	Client Hello
126	1.260	1	PCF	UDR	TLSv1.3	Client Hello
127	1.267	1	PCF	NRF	TLSv1.3	Client Hello
128	1.276	1	PCF	UDR	TLSv1.3	Client Hello
129	1.283	2	PCF	NRF	TLSv1.3	Client Hello
131	1.299	3	SMF	NRF	TLSv1.3	Client Hello
134	1.324	1	SMF	CHF	TLSv1.3	Client Hello
135	1.336	1	AMF	NRF	TLSv1.3	Client Hello
136	1.339	1	SMF	NRF	TLSv1.3	Client Hello
137	1.345	1	AMF	SMF	TLSv1.3	Client Hello
138	1.349	1	SMF	AMF	TLSv1.3	Client Hello
139	1.351	1	AMF	gNB	NGAP/ NAS-5GS	PDU session establishment accept
140	1.353	1	gNB	AMF	NGAP	PDU Session Resource Setup Response
141	1.358	2	AMF	NRF	TLSv1.3	Client Hello

Figure A1. All stages of a typical UE connection with a total of 129 TLSv1.3 handshakes.

References

1. Sultan, A. 5G System Overview—3gpp.org. 2022. Available online: https://www.3gpp.org/technologies/5g-system-overview (accessed on 23 September 2024).
2. 3rd Generation Partnership Program. 3GPP Release 18. 2024. Available online: https://www.3gpp.org/specifications-technologies/releases/release-18 (accessed on 7 May 2024).
3. Neto, N.N.; Madnick, S.; de Paula, A.M.G.; Borges, N.M. Developing a Global Data Breach Database and the Challenges Encountered. *J. Data Inf. Qual.* **2021**, *13*, 1–33. [CrossRef]
4. European Telecommunications Standards Institute. ETSI GR QSC 006 v1.1.1: Quantum-Safe Cryptography (QSC); Limits to Quantum Computing Applied to Symmetric Key Sizes. 2017. Available online: https://www.etsi.org/deliver/etsi_gr/QSC/001_099/006/01.01.01_60/gr_qsc006v010101p.pdf (accessed on 22 September 2024).
5. McNeely, D. Transitioning to Quantum-Safe Encryption. 2024. Available online: https://delinea.com/blog/quantum-safe-encryption (accessed on 22 September 2024).
6. National Institute for Standards and Technology. Post-Quantum Cryptography. 2017. Available online: https://csrc.nist.gov/projects/post-quantum-cryptography (accessed on 21 June 2024).
7. 3rd Generation Partnership Program. TS 33.210 V18.1.0 Technical Specification Group Services and System Aspects; Network Domain Security (NDS); IP Network Layer Security (Release 18). 2024. Available online: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2279 (accessed on 14 September 2024).
8. 3rd Generation Partnership Program. TS 33.501 V19.0.0 Technical Specification Group Services and System Aspects; Security Architecture and Procedures for 5G System (Release 19). 2024. Available online: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3169 (accessed on 14 September 2024).
9. Open Quantum Safe. Software for the Transition to Quantum-Resistant Cryptography. 2024. Available online: https://openquantumsafe.org (accessed on 10 June 2024).
10. Mounica, G.R.; Manimaran, G.; Jerome, L.B.; Bhattacharjee, P. Implementation of 5-Qubit approach-based Shor's Algorithm in IBM Qiskit. In Proceedings of the 2021 IEEE Pune Section International Conference (PuneCon), Pune, India, 16–19 December 2021; pp. 1–6. [CrossRef]
11. Shor, P.W. Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer. *SIAM J. Comput.* **1997**, *26*, 1484–1509. [CrossRef]
12. Grover, L.K. A fast quantum mechanical algorithm for database search. In Proceedings of the Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing, STOC '96, Philadelphia, PA, USA, 22–24 May 1996; pp. 212–219. [CrossRef]
13. Zeydan, E.; Turk, Y.; Aksoy, B.; Ozturk, S.B. Recent Advances in Post-Quantum Cryptography for Networks: A Survey. In Proceedings of the 2022 Seventh International Conference on Mobile and Secure Services (MobiSecServ), Gainesville, FL, USA, 26–27 February 2022; pp. 1–8. [CrossRef]
14. Dam, D.T.; Tran, T.H.; Hoang, V.P.; Pham, C.K.; Hoang, T.T. A Survey of Post-Quantum Cryptography: Start of a New Race. *Cryptography* **2023**, *7*, 40. [CrossRef]
15. Bavdekar, R.; Jayant Chopde, E.; Agrawal, A.; Bhatia, A.; Tiwari, K. Post Quantum Cryptography: A Review of Techniques, Challenges and Standardizations. In Proceedings of the 2023 International Conference on Information Networking (ICOIN), Bangkok, Thailand, 11–14 January 2023; pp. 146–151. [CrossRef]
16. National Institute for Standards and Technology. NIST Releases First 3 Finalized Post-Quantum Encryption Standards —nist.gov. 2024. Available online: https://www.nist.gov/news-events/news/2024/08/nist-releases-first-3-finalized-post-quantum-encryption-standards (accessed on 24 September 2024).
17. Scalise, P.; Boeding, M.; Hempel, M.; Sharif, H.; Delloiacovo, J.; Reed, J. A Systematic Survey on 5G and 6G Security Considerations, Challenges, Trends, and Research Areas. *Future Internet* **2024**, *16*, 67. [CrossRef]
18. Clancy, T.C.; McGwier, R.W.; Chen, L. Post-quantum cryptography and 5G security: Tutorial. In Proceedings of the Proceedings of the 12th Conference on Security and Privacy in Wireless and Mobile Networks, WiSec '19, Miami, FL, USA, 15–17 May 2019; p. 285. [CrossRef]
19. Rossi Figlarz, G.; Passuelo Hessel, F. Enhancing the 5G-AKA Protocol with Post-quantum Digital Signature Method. In *Advanced Information Networking and Applications*; Barolli, L., Ed.; Springer: Cham, Switzerland, 2024; pp. 99–110.
20. Hoque, S.; Aydeger, A.; Zeydan, E. Exploring Post Quantum Cryptography with Quantum Key Distribution for Sustainable Mobile Network Architecture Design. In Proceedings of the 4th Workshop on Performance and Energy Efficiency in Concurrent and Distributed Systems, PECS '24, New York, NY, USA, 3 June 2024; pp. 9–16. [CrossRef]
21. Minhas, N.N.; Mansoor, K. Edge Computing-Based Scheme for Post-Quantum IoT Security for e-Health. *IEEE Internet Things J.* **2024**, *11*, 31331–31337. [CrossRef]
22. Hu, Y.C.; Patel, M.; Sabella, D.; Sprecher, N.; Young, V. Mobile Edge Computing A Key Technology Towards 5G. 2015. Available online: https://yucianga.info/wp-content/uploads/2015/11/Ref02-2015-09-etsi_wp11_mec_a_key_technology_towards_5g.pdf (accessed on 22 June 2024).
23. Zhang, J.; Xie, W.; Yang, F.; Bi, Q. Mobile edge computing and field trial results for 5G low latency scenario. *China Commun.* **2016**, *13*, 174–182. [CrossRef]
24. Chen, Q.; Wang, Z.; Su, Y.; Fu, L.; Wei, Y. Educational 5G Edge Computing: Framework and Experimental Study. *Electronics* **2022**, *11*, 2727. [CrossRef]

25. Zheng, J.; Zhu, H.; Dong, Y.; Song, Z.; Zhang, Z.; Yang, Y.; Zhao, Y. Faster Post-quantum TLS 1.3 Based on ML-KEM: Implementation and Assessment. In Proceedings of the Computer Security—ESORICS 2024, Bydgoszcz, Poland, 16–20 September 2024; Garcia-Alfaro, J.; Kozik, R.; Choraś, M.; Katsikas, S., Eds.; Springer: Cham, Switzerland, 2024; pp. 123–143.
26. Zeydan, E.; Blanco, L.; Mangues-Bafalluy, J.; Aydeger, A.; Arslan, S.; Turk, Y. Integrating Quantum-Secured Blockchain Identity Management in Open RAN for 6G Networks. In Proceedings of the 2024 IEEE 49th Conference on Local Computer Networks (LCN), Normandy, France, 8–10 October 2024; pp. 1–7. [CrossRef]
27. Javeed, D.; Saeed, M.S.; Ahmad, I.; Adil, M.; Kumar, P.; Islam, A.N. Quantum-empowered federated learning and 6G wireless networks for IoT security: Concept, challenges and future directions. *Future Gener. Comput. Syst.* **2024**, *160*, 577–597. [CrossRef]
28. Nakka, K.; Ahmad, S.; Kim, T.; Atkinson, L.; Ammari, H.M. Post-Quantum Cryptography (PQC)-Grade IEEE 2030.5 for Quantum Secure Distributed Energy Resources Networks. In Proceedings of the 2024 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), Washington, DC, USA, 19–22 February 2024; pp. 1–5. [CrossRef]
29. Lawo, D.C.; Abu Bakar, R.; Cano Aguilera, A.; Cugini, F.; Imaña, J.L.; Tafur Monroy, I.; Vegas Olmos, J.J. Wireless and Fiber-Based Post-Quantum-Cryptography-Secured IPsec Tunnel. *Future Internet* **2024**, *16*, 300. [CrossRef]
30. open5G. Core Network | Open, Programmable and Virtualized 5G Networks. 2024. Available online: https://open5g.info/core/ (accessed on 8 July 2024).
31. Linh, A.B.N. Analysing Open-Source 5G Core Networks for TLS Vulnerabilities and 3GPP Compliance. 2023. Available online: https://cs.ru.nl/bachelors-theses/2023/Alex_Bui_Nhat_Linh___1040308___Analysing_open-source_5G_core_networks_for_TLS_vulnerabilities_and_3GPP_compliance.pdf (accessed on 8 July 2024).
32. TLS-Attacker. TLS-Scanner. 2023. Available online: https://github.com/tls-attacker/TLS-Scanner (accessed on 8 July 2024).
33. Diquet, A. SSLyze. 2024. Available online: https://github.com/nabla-c0d3/sslyze (accessed on 8 July 2024).
34. DinoTools. PySSLScan. 2022. Available online: https://github.com/DinoTools/pysslscan (accessed on 8 July 2024).
35. Adhikary, A.; Deb Raha, A.; Qiao, Y.; Saad, W.; Han, Z.; Seon Hong, C. Holographic MIMO with Integrated Sensing and Communication for Energy-Efficient Cell-Free 6G Networks. *IEEE Internet Things J.* **2024**, *11*, 30617–30635. [CrossRef]
36. Chen, W.; Lin, X.; Lee, J.; Toskala, A.; Sun, S.; Chiasserini, C.F.; Liu, L. 5G-Advanced Toward 6G: Past, Present, and Future. *IEEE J. Sel. Areas Commun.* **2023**, *41*, 1592–1619. [CrossRef]
37. Boeding, M.; Hempel, M.; Sharif, H. Novel Approach towards a Fully Deep Learning-Based IoT Receiver Architecture: From Estimation to Decoding. *Future Internet* **2024**, *16*, 155. [CrossRef]
38. Chen, R.H.; Zhou, J.; Zhu, Y.; Zhang, K. When to Simply Use Passive RIS as Beamformer: An Information-Theoretic Analysis and A Novel Single-RF MIMO Transceiver Architecture. *IEEE Trans. Wirel. Commun.* **2024**, 1. [CrossRef]
39. Karakaya, A.; Ulu, A. A survey on post-quantum based approaches for edge computing security. *Wiley Interdiscip. Rev. Comput. Stat.* **2024**, *16*, e1644. [CrossRef]
40. OpenSSL. OpenSSL. 2024. Available online: https://www.openssl.org (accessed on 28 May 2024).
41. BoringSSL. BoringSSL—Github. 2024. Available online: https://github.com/google/boringssl/ (accessed on 29 May 2024).
42. free5GC. free5GC Roadmap. 2024. Available online: https://free5gc.org/#next-steps (accessed on 27 May 2024).
43. Go. Build Simple, Secure, Scalable Systems with Go. 2024. Available online: https://go.dev (accessed on 5 June 2024).
44. Go. Crypto/Tls Go Library. 2024. Available online: https://pkg.go.dev/crypto/tls (accessed on 6 July 2024).
45. Pexip. Go-Openssl Wrapper. 2024. Available online: https://pkg.go.dev/github.com/pexip/go-openssl?utm_source=godoc (accessed on 9 July 2024).
46. Güngör, A. UERANSIM. 2022. Available online: https://github.com/aligungr/UERANSIM (accessed on 12 August 2024).
47. Schwabe, P. Performance Overview of CRYSTALS-Kyber. Technical Report, PQ-CRYSTALS. 2020. Available online: https://pq-crystals.org/kyber/ (accessed on 17 September 2024).
48. Kölbl, S.; Schmieg, S.; Endignoux, G. Why Hybrid Deployments Are Key to Secure PQC Migration. 2024. Available online: https://bughunters.google.com/blog/5266882047639552/why-hybrid-deployments-are-key-to-secure-pqc-migration (accessed on 14 October 2024).
49. Chen, X.; Ng, D.W.K.; Yu, W.; Larsson, E.G.; Al-Dhahir, N.; Schober, R. Massive access for 5G and beyond. *IEEE J. Sel. Areas Commun.* **2020**, *39*, 615–637. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

electronics

MDPI

Article

Provably Quantum Secure Three-Party Mutual Authentication and Key Exchange Protocol Based on Modular Learning with Error

Hyewon Park [1], Seunghwan Son [2], Youngho Park [2] and Yohan Park [1,*]

[1] School of Computer Engineering, Keimyung University, Daegu 42601, Republic of Korea; wldnjsfuf@stu.kmu.ac.kr
[2] School of Electronic and Electrical Engineering, Kyungpook National University, Daegu 41566, Republic of Korea; sonshawn@knu.ac.kr (S.S.); parkyh@knu.ac.kr (Y.P.)
* Correspondence: yhpark@kmu.ac.kr; Tel.: +82-53-580-5229

Abstract: With the rapid development of quantum computers, post-quantum cryptography (PQC) has become critical technology in the security field. PQC includes cryptographic techniques that are secure against quantum-computer-based attacks, utilizing methods such as code-based, isogeny-based, and lattice-based approaches. Among these, lattice-based cryptography is the most extensively studied due to its ease of implementation and efficiency. As quantum computing advances, the need for secure communication protocols that can withstand quantum computer-based threats becomes increasingly important. Traditional two-party AKE protocols have a significant limitation: the security of the entire system can be compromised if either of the communicating parties behaves maliciously. To overcome this limitation, researchers have proposed three-party AKE protocols, where a third party acts as an arbiter or verifier. However, we found that a recently proposed three-party AKE protocol is vulnerable to quantum-computer-based attacks. To address this issue, we propose a provably quantum secure three-party AKE protocol based on MLWE. The proposed scheme leverages the user's biometric information and the server's master key to prevent the exposure of critical parameters. We analyzed the security of the protocol using simulation tools such as the Burrows–Abadi–Needham (BAN) logic, Real-or-Random (RoR) model, and Automated Validation of Internet Security Protocols and Applications (AVISPA). Furthermore, comparative analysis with similar protocols demonstrates that our protocol is efficient and suitable.

Keywords: post-quantum cryptography; learning with error (LWE); modular-LWE (MLWE); three-party authentication

1. Introduction

With the rapid advancement of quantum computing, the security of traditional cryptographic systems is increasingly under threat. The security of widely used cryptographic algorithms, such as Rivest–Shamir–Adleman (RSA) and elliptic curve cryptography (ECC), largely relies on the difficulty of the factorization problem (FP) [1] and the discrete logarithm problem (DLP) [2]. However, as demonstrated by Shor [3] and Grover [4], both FP and DLP can be solved in polynomial time by quantum computers. This development makes the foundational mathematical problems underlying RSA and ECC vulnerable to quantum-based attacks. In response to this emerging threat, extensive research has been conducted on post-quantum cryptography (PQC), which aims to develop cryptographic methods resilient to quantum attacks [5–9].

There are several approaches, such as multivariate-based [5], hash-based [6], code-based [7], isogeny-based methods [8], and lattice-based [9], to defend these attacks. Lattice-based cryptography algorithms, first proposed by Ajtai in 1996 [10], are built on complex mathematical problems defined within a lattice. Notable advantages of these lattice-based

algorithms are easy to implement, efficient in computation, and suitable for lightweight devices, because all operations can be executed using matrix multiplications and additions.

PQC is gaining attention in the research of three-party authentication and key exchange (AKE) protocols. While current three-party AKE schemes may resist classical attacks, they remain vulnerable to quantum-based threats. Unlike two-party AKE, which assumes honest behavior from both participants, three-party AKE protocols introduce a third party as a mediator, enhancing system reliability and reducing the risk of data leaks. Mutual authentication among all three parties is essential to prevent malicious impersonation and data manipulation, ensuring secure and trustworthy communication. However, many existing quantum-secure three-party AKE protocols still face performance and security challenges.

One popular cryptosystem used in three-party authentication and key exchange protocols is the lattice-based encryption algorithm learning with error (LWE). LWE is considered more secure than other cryptosystems because it relies on a worst-case hardness assumption, which greatly improves security. In 2009, Regev [11] was the first to propose using LWE problems to design public-key cryptosystems. However, LWE has the disadvantage of incurring a high overhead. In 2010, Lyubashevsky et al. [12] introduced the concept of ideal lattices and proposed ring-LWE (RLWE) based on it. RLWE is based on the worst-case intractability of the ideal lattice and allows for a more efficient implementation of public-key cryptosystems. However, RLWE is generally considered to be less secure than LWE. To address this, in 2015, Langlois et al. [13] proposed modular-LWE (MLWE), which combines the strengths of LWE and RLWE. MLWE has lower overhead compared to LWE and higher security compared to RLWE by introducing rings while maintaining the matrix structure. In addition, MLWE has flexible parameterization that can be adapted to different security levels by modifying the matrix size. Therefore, we conducted research on three-party AKE protocols with a focus on MLWE schemes that combine the strengths of LWE and RLWE.

While researching MLWE-based three-party AKE protocols, we came across a paper by Guo et al. in 2023 [14]. Guo et al. presented a three-party AKE protocol based on MLWE and claimed that their scheme is secure, efficient, and mutually authentic. However, we found that the scheme is vulnerable to replay attacks, user and server impersonation, authentication table leaks, insider attacks, and DoS attacks. In addition, mutual authentication was not fully guaranteed. Their scheme used a password-based AKE scheme. In this paper, we analyze which attacks Guo et al.'s scheme is vulnerable to and propose a three-factor AKE scheme that uses the user's biometric information to improve the security and efficiency of the scheme.

1.1. Contribution

- We identified that the authentication and key exchange scheme proposed by Guo et al. exhibits correctness issues and security vulnerabilities, particularly in the mutual authentication process. To address these challenges, we propose a three-party authentication and key exchange scheme based on MLWE that incorporates biometric data.
- We provide a formal proof of the security and robustness of the proposed scheme using BAN logic, the Real-or-Random (RoR) model, and Automated Validation of Internet Security Protocols and Applications (AVISPA). Moreover, we demonstrate that the scheme is secure against various attacks, including insider, impersonation, replay, and man-in-the-middle attacks. Furthermore, we establish that our scheme ensures mutual authentication.
- We evaluate the security features and computational costs of the proposed scheme. Through comparative analysis, we demonstrate its superiority over the methods proposed by Guo et al. and other related schemes.

1.2. Paper Organization

In Section 2, we review the foundational works on LWE, RLWE, and MLWE, along with related research on public-key cryptography and key exchange protocols based on LWE. In Section 3, we describe MLWE, the mathematical techniques utilized in our proposed

scheme, and introduce the error reconciliation scheme, biohash, attacker model, and notation. Next, in Section 4, we revisit the scheme proposed by Guo et al., and in Section 5, we detail the vulnerabilities identified in their scheme. In Section 6, we introduce our proposed scheme, followed by our security analysis using BAN logic, RoR models, and AVISPA tool in Section 7. The performance analysis, which includes security features, communication costs, and computational efficiency, is presented in Section 8. Finally, in Section 9, we conclude the paper and outline future research directions.

2. Related Works

Lattice-based cryptography is a promising field for achieving post-quantum security, particularly due to its reliance on worst-case hardness assumptions, which provide resistance against quantum attacks. In 2009, Regev [11] introduced the LWE problem, which became a foundation for constructing secure public-key cryptographic schemes based on the hardness of lattice problems. In 2010, Lyubashevsky et al. [12] introduced the RLWE problem, which incorporated ideal lattices and algebraic structures to enable more efficient public-key cryptosystems based on worst-case hardness assumptions in ideal lattices. Further advancing this field, Langlois et al. [13] proposed the MLWE problem in 2015, which bridges the gap between LWE and RLWE by combining ring polynomials with matrix structures. MLWE offers lower computational overhead than LWE and greater security compared to RLWE, along with the flexibility to adjust matrix dimensions for different security levels.

Building on these advances in lattice-based cryptography, various researchers have focused on developing practical and secure key exchange protocols, often inspired by the Diffie–Hellman approach, using LWE, RLWE, and their variants. In 2012, Ding et al. [15] introduced the first provably secure key exchange protocol based on the LWE problem, which is both computationally efficient and extendable to the RLWE framework. However, the shared key generated in their scheme is not uniformly random, leading them to recommend the use of an extractor, which unfortunately reduces the effective key length. While the protocol proposed by Ding et al. resembles the Diffie–Hellman key exchange, it does not inherently provide reliability. The key contribution of their work lies in presenting the first LWE and RLWE key exchange frameworks incorporating a novel error reconciliation mechanism. In 2014, Peikert [16] proposed a straightforward, low-bandwidth coordination technique that allows two parties to "roughly agree" on a secret value, ultimately enabling them to reach an exact agreement. However, this method, while grounded in LWE-based cryptographic techniques, is computationally more intensive. Essentially, this approach is a form of key encapsulation and does not guarantee reliability for users. The primary distinction between Ding et al. and Peikert is that the former employs a biased coordination mechanism, whereas the latter utilizes an unbiased one.

Building on the DH-like structures discussed earlier, various researchers have proposed authentication and key exchange schemes to enhance security in different environments. Recently, Wang et al. [17] introduced a quantum-resistant two-factor authentication protocol optimized for mobile devices. Although this protocol is effective, it does introduce some overhead. In 2022, Dharminder et al. [18] proposed a security reconciliation mechanism based on lattice cryptography to facilitate key exchange in Internet of Things (IoT) environments. The following year, Kumar et al. [19] introduced a post-quantum key exchange protocol grounded in RLWE, a variant of lattice-based cryptography. Their protocol supports authenticated key exchange (AKE) with just two message exchanges, significantly enhancing efficiency.

However, all of the aforementioned protocols are two-party AKE schemes, which inherently suffer from certain limitations. In a two-party authentication protocol, two entities, commonly referred to as Alice and Bob, exchange messages to authenticate each other. This setup assumes that both parties will act honestly and securely. However, if either party is compromised or behaves maliciously, the security of the entire system can be jeopardized. This vulnerability underscores the need for a more robust solution. Introducing a third party, typically called Charlie, transforms the process into a three-party AKE protocol. In this setup, Charlie acts as a trusted intermediary, enhancing the overall

security of the authentication process by mediating between Alice and Bob. Charlie's role is critical, as it helps to mitigate risks associated with direct communication between the two parties and provides an additional layer of verification and trust. This added complexity not only strengthens the authentication process but also ensures that even if one party is compromised, the security of the system remains intact [20].

Given the inherent limitations of two-party AKE schemes, recent research has shifted focus towards three-party AKE protocols to enhance security and resilience. In 2023, Chaudhary et al. [21] proposed a three-party key agreement protocol using ECC encryption, providing authentication, anonymity, and forward secrecy. Rewal et al. [22] introduced a lattice-based key establishment protocol that enhances efficiency and resists password guessing but lacks anonymous communication and full three-factor security. Guo et al. [14] proposed a three-party AKE protocol based on MLWE. However, the scheme was found to be vulnerable to attacks such as replay, impersonation, insider, and DoS. In addition, mutual authentication was not fully guaranteed. As a result, designing a secure three-party quantum key agreement protocol that addresses these vulnerabilities remains a significant challenge.

Table 1 summarizes cryptographic technologies and contributions of various authentication schemes using LWE.

Table 1. Post-quantum key exchange protocols overview.

Authors	Cryptographic Technique	Contribution
Regev [11]	LWE	Introduced the LWE problem for building public-key cryptographic schemes based on worst-case hardness assumptions.
Lyubashevsky et al. [12]	RLWE	Introduced ideal lattice and RLWE problem, allowing for more efficient public-key cryptosystems.
Langlois et al. [13]	MLWE	Proposed MLWE as a compromise between LWE and RLWE, offering flexibility and higher security with adjustable parameters.
Ding et al. [15]	LWE & RLWE	First provably secure key exchange protocol based on LWE, introduced error reconciliation mechanism.
Peikert [16]	LWE	Proposed low-bandwidth coordination method for key agreement with unbiased mechanism.
Wang et al. [17]	RLWE	Proposed efficient quantum-resistant protocol for mobile devices.
Dharminder et al. [18]	RLWE	Proposed lattice-based security coordination for IoT environments.
Kumar et al. [19]	RLWE	Proposed new post-quantum key exchange ensuring both AKE.
Chaudhary et al. [21]	RLWE & ECC	Proposed protocol ensuring authentication, anonymity, and forward secrecy.
Rewal et al. [22]	RLWE	Proposed protocol to resist password guessing attacks and enhance efficiency.
Guo et al. [14]	MLWE	Proposed a three-party PAKE protocol using MLWE and Peikert's error reconciliation.

3. Preliminaries

3.1. Module Learning with Error

The MLWE problem extends the LWE and RLWE problems by introducing a more flexible structure. Specifically, let R_q^d represent a vector containing d ring polynomials in $R_q = \mathbb{Z}[X]/(X^n + 1)$. Additionally, let $R_q^{k \times d}$ denote a matrix of ring polynomials with rank k and dimension d. The MLWE problem generalizes both LWE and RLWE: when the ring R_q is set to $\mathbb{Z}_q$, it reduces to the LWE problem, and when $d = k = 1$, it reduces to the RLWE problem.

In 2015, Langlois and Stehlé [13] introduced the MLWE problem as a compromise between LWE and RLWE. By retaining the matrix format while incorporating algebraic structure, MLWE offers greater flexibility and simplicity in designing lattice-based cryptographic schemes for various security scenarios. The decision version of the MLWE problem is defined as follows:

Definition 1. *(MLWE problem). The MLWE problem is defined with parameters (n, q, k, η), where n represents the dimension of the vector, q is the modulus, k is the dimension of the polynomial matrix, and η specifies the distribution parameters for the error and the secret. Given these parameters, a polynomial matrix $A \in R_q^{k \times k}$ is selected randomly and uniformly. Subsequently, a secret $s \in \beta_\eta^k$ and an error vector $e \in \beta_\eta^k$ are also chosen randomly and uniformly. Using these, the vector $b = As + e \in R_q^{k \times k}$ is computed.*

The MLWE problem then asks one to distinguish between two distributions: one where A and s are chosen uniformly at random and e follows the distribution β_η, and another where A and b are chosen uniformly at random.

Definition 2. *(Decision MLWE problem). The decision MLWE problem involves determining whether a given pair (A, b) originates from one of two specified distributions: one where $b = As + e$ for a uniformly chosen secret s and an error e sampled from the distribution β_η, or another where b is uniformly random.*

More specifically, given $(A, X, Y, \sigma, K) \in R_q^{k \times k} \times R_q^k \times R_q^k \times 0, 1^n \times 0, 1^n$, where $\sigma = dbl(X^T s_y + e_z)_{2q,2} \in \{0,1\}^n$ and $K = dbl(Z)_{2q,2} \in \{0,1\}^n$, the decision MLWE problem requires deciding whether $Y = A^T s_y + e_y \in R_q^k$ comes from the first distribution or whether Y is uniformly random in R_q^k.

3.2. Reconciliation Mechanism

An error reconciliation mechanism helps ensure that two parties, each holding slightly differing values, can converge on the same final value through a combination of information exchange and computation. In 2014, Peikert introduced an enhanced error reconciliation mechanism, building on the earlier work of Ding. Suppose Alice and Bob possess approximate values $w \in \mathbb{Z}_q$ and $v \in \mathbb{Z}_q$, respectively. To achieve reconciliation, Bob applies a modular rounding function and a cross-rounding function. The output of the modular rounding function serves as the shared key, while the cross-rounding function generates the signal value. The definitions of these two functions are as follows:

Definition 3. *Let q be an even modulus. The modular rounding function is defined as*

$$\lfloor x \rceil_{q,2} := \left\lfloor \frac{2}{q} x \right\rceil \bmod 2$$

and the cross-rounding function is defined as

$$\langle x \rangle_{q,2} := \left\lfloor \frac{4}{q} x \right\rfloor \bmod 2.$$

Once Alice receives the signal value from Bob, she generates the shared key using the reconciliation function.

Definition 4. *(The reconciliation function). Given an element $w \in \mathbb{Z}_q$ and the cross-rounding value $h = \langle v \rangle_{q,2}$ of a close element $v \in \mathbb{Z}_q$, define the sets $I_0 = \left\{0, 1, \cdots, \lfloor \frac{q}{4} \rceil - 1\right\}$ and $I_1 = \left\{-\lfloor \frac{q}{4} \rfloor, \cdots, -1\right\}$. Let $E = \left[-\frac{q}{8}, \frac{q}{8}\right)$, and the reconciliation function is*

$$Rec(w, h) = \begin{cases} 0, & \text{if } w \in I_h + E \bmod q \\ 1, & \text{otherwise.} \end{cases}$$

3.3. Biohashing

The utilization of a user's biometric information in an authentication system is an effective method for verifying the identity of the actual user. In 2004, Jin et al. [23] introduced biohashing, a technique that combines pseudorandom numbers with biometric features to enhance security. This approach, particularly in the context of fingerprint data, makes it challenging for an attacker to replicate the biohash code without access to the correct hash key.

- Feature extraction: Biometric data (e.g., a face image) are captured and processed to extract a feature vector $x \in \mathbb{R}^n$. For instance, discrete cosine transform (DCT) coefficients can be utilized as features.
- Pseudo-random vector generation: A secret seed (hash key) K is used to generate a sequence of pseudorandom vectors $r_i \in \mathbb{R}^n, (i = 1, \ldots, n)$. These vectors are ensured to be linearly independent. An example of an algorithm that generates such pseudorandom numbers is the Blum–Blum–Shub method.
- Orthonormalization: The generated vectors r_i are transformed into an orthonormal set or_i using the Gram–Schmidt orthonormalization procedure.
- Scalar product and threshold binarization: The inner product between the biometric feature vector x and each orthonormal vector or_i is calculated: $\langle x \mid or_i \rangle$. Each result is compared to a threshold τ to produce a bit:

$$b_i = \begin{cases} 0, & \text{if } \langle x \mid o_i \rangle \leq v \\ 1, & \text{if } \langle x \mid o_i \rangle \end{cases}$$

- BioHash code: The resulting bit vector $b_i \in 0, 1^m$ is the BioHash code, which is unique to the user's biometric data and the hash key.
- Matching: During authentication, the BioHash code is compared using Hamming distance for similarity matching.

3.4. Adversary Model

The adversary's capabilities are defined based on the "Dolev–Yao (DY) model" [24], which is commonly used for protocol security analysis [25,26]. The adversary is assumed to possess the following abilities:

- The adversary can eavesdrop on, intercept, modify, delete, and fabricate messages transmitted over public channels.
- The adversary can make educated guesses regarding either the legitimate user's identity or password but cannot correctly guess both at the same time.
- The adversary can launch various attacks, including impersonation, replay, and man-in-the-middle attacks.
- The adversary may also operate as an insider within the system.

3.5. Notations

Table 2 provides the symbols and their descriptions used throughout this paper.

Table 2. Notations.

Symbol	Descryption
U_i	User
S_j	Server
ID_i	Identity of U_i
PW_i	Password of U_i
Bio_i	Biometric information of U_i
T_1, T_2, T_3, T_4	Timestamp
n_1, n_2, n_3, n_4, n_5	Random number
$dbl(v)$	Randomized doubling function
$\langle \bar{v} \rangle_{2q,2}$	Cross-rounding function
$\lfloor \bar{v} \rceil_{2q,2}$	Modular rounding function
$Rec(2v, \sigma)$	Reconciliation function
b_i	Public key of U_i in session
s_i	Secret key of U_i in session
sk	Session key
$h(\cdot)$	Hash function
$h_b(\cdot)$	Biohash function
$\oplus$	Exclusive-or operation
$\|\|$	Concatenation operation

4. Revisiting Guo et al.'s Scheme

4.1. System Initialization Phase

In the password-based AKE scheme proposed by Guo et al., R and R_q are used to denote $\mathbb{Z}[x]/(x^n+1)$ and $\mathbb{Z}q[x]/(x^n+1)$, respectively. In this context, lowercase italics are used to denote column vectors whose elements belong to the ring R or R_q, while uppercase italics are used to denote matrices. The open parameters are (n, q, k), where n is a power of 2, and q is an odd prime number that satisfies $q \equiv 1 \mod n$. The notation $A \leftarrow R_q^{k\times k}$ represents a matrix sampled uniformly from $R_q^{k\times k}$. Similarly, $(s, e) \leftarrow \beta_\eta^k \times \beta_\eta^k$ represents a private key and noise sampled uniformly from β_η^k.

4.2. Registration Phase

Before generating a session key for communication, users must register with the server through a secure channel. The detailed procedure is outlined below:

Step 1: U_i chooses ID_i and PW_0, then calculates $PW_1 = H(ID_i||PW_0)$. U_i subsequently sends ID_i and PW_1 to the server as a registration request through a secure channel.

Step 2: Upon receiving ID_i and PW_1 from the user, the server generates a random number n_0 and calculates $TID_i = H(ID_i||n_0||PW_1)$. The server then sends TID_i to U_i through a secure channel.

4.3. Authentication and Key Exchange Phase

For User A and User B to securely share a session key for communication, they must first mutually authenticate with a trusted server to verify the legitimacy of their identities. The detailed process is outlined below and is illustrated in Figure 1.

Step 1: U_A uniformly samples $(s_{A0}, e_{A0}) \leftarrow \beta_\eta^k \times \beta_\eta^k$ and generates a random number n_1. Then, U_A computes $b_{A0} = As_{A0} + e_{A0}$. After finding its own TID_A and the counterpart TID_B it wants to communicate with, U_A creates the message $M_0 = (TID_A, TID_B, n_1, b_{A0})$ and sends it to the server.

Step 2: Upon receiving M_0 from U_A, the server verifies the user's identity using the TID_A and TID_B. The server then samples $(s_{S0}, e_{S0}, e_{S1}) \leftarrow \beta_\eta^k \times \beta_\eta^k \times \beta_\eta$ and $(s_{S1}, e_{S2}) \leftarrow \beta_\eta^k \times \beta_\eta^k$ and generates the random number N_2 and computes $b_{S0} = A^T s_{S0} + e_{S0}$, $b_{S1} = As_{S1} + e_{S2}$, and $v_0 = b_{A0}^T s_{S1} + e_{S1}$. Finally, the server uses the randomized function, cross-rounding function, and modular rounding function to compute $\bar{v_0} \leftarrow dbl(v_0)$, $\sigma_0 \leftarrow \langle \bar{v_0} \rangle_{2q,2}$, $k_0 \leftarrow \lfloor \bar{v_0} \rceil_{2q,2}$, and sends the message $M_1 = (TID_A, TID_B, n_1, n_2, b_{A0}, b_{S1})$ to U_B.

Step 3: After receiving M_1 from the server, U_B samples $(s_{B0}, e_{B0}, e_{B1}, e_{B2}) \leftarrow \beta_\eta^k \times \beta_\eta^k \times \beta_\eta \times \beta_\eta$, generates two random numbers, n_2 and n_4, computes $b_{B0} = A^T s_{B0} + e_{B0}$, $v_1 = b_{S1}^T s_{B0} + e_{B1}$, $v_2 = b_{A0}^T s_{B0} + e_{B2}$, and calculates $\bar{v}_1 \leftarrow dbl(v_1)$, $\bar{v}_2 \leftarrow dbl(v_2)$, $\sigma_1 \leftarrow \langle \bar{v}_1 \rangle_{2q,2}$, $\sigma_2 \leftarrow \langle \bar{v}_2 \rangle_{2q,2}$, $k_1 \leftarrow \lfloor \bar{v}_1 \rceil_{2q,2}$, $k_2 \leftarrow \lfloor \bar{v}_2 \rceil_{2q,2}$ using the randomized function, cross-rounding function, and modular rounding function. U_B computes $Auth_0 = H_1(TID_A, TID_B, n_2, n_3, k_1, PW_{B1})$, the authentication information between U_B and the server, and the shared key $sk_0 = H_2(TID_A, TID_B, n_2, n_3, k_1, PW_{B1})$, and computes $Auth_1 = H_1(TID_A, TID_B, b_{A0}, b_{B0}, n_1, n_4, k_2)$, the authentication information between U_B and U_A, and the shared key $sk_B = H_2(TID_A, TID_B, b_{A0}, b_{B0}, n_1, n_4, k_2)$. Finally, U_B encrypts $(\sigma_2, n_4, Auth_1)$ with sk_0 and generates the message $M_2 = (TID_A, Auth_0, \sigma_1, n_3, c_0, b_{B0})$ and sends it to the server.

Step 4: The server computes $v_1' = b_{B0}^T s_{S1}$ and $k_1' = Rec(2v_1', \sigma_1) \in \{0,1\}^n$ using Peikert error reconciliation function, and calculates $Auth_0' = H_1(TID_A, TID_B, n_2, n_3, k_1', PW_{B1})$, the authentication information between the server and U_B. By checking $Auth_0' ? = Auth_0$, the server verifies that U_B is a legitimate user, then computes

$sk_0 = H_2(TID_A, TID_B, n_2, n_3, k_1, PW_{B1})$, the shared key between the server and U_B, and uses sk_0 to decrypt c_0 to obtain $(\sigma_2, n_4, Auth_1)$. The server then computes $Auth_2 = H_1(TID_A, TID_B, n_1, n_2, k_0, PW_{A1})$, the authentication information between the server and U_A, and computes the shared key $sk_1 = H_2(TID_A, TID_B, n_1, n_2, k_0, PW_{A1})$. Finally, the server encrypts $(\sigma_2, n_4, Auth_1)$ with sk_1 and generates a message $M_3 = (TID_B, n_2, \sigma_0, c_1, Auth_2, b_{B0}, b_{S0})$ and sends it to U_A.

Step 5: U_A computes $v_0' = b_{S0}^T s_{A0}$ and $k_0' = Rec(2v_0', \sigma_0) \in \{0,1\}^n$ using the Peikert error reconciliation function, and calculates $Auth_2' = H_1(TID_A, TID_B, n_1, n_2, k_0', PW_{A1})$, the authentication information between the server and U_A. By verifying $Auth_2' ? = Auth_2$, U_A confirms the server's credibility. Then U_A computes $sk_1 = H_1(TID_A, TID_B, n_1, n_2, k_0', PW_{A1})$, the shared key between the server and U_A, and uses sk_1 to decrypt c_1 to obtain $(\sigma_2, n_4, Auth_1)$. Using the decrypted information, U_A computes $v_2' = b_{B0}^T s_{A0}$ and $k_2' = Rec(2v_2', \sigma_2) \in \{0,1\}^n$ with the Peikert error reconciliation function. U_A then computes $Auth_1' = H_1(TID_A, TID_B, b_{A0}, b_{B0}, n_1, n_4, k_2')$ and compares it with $Auth_1$. If $Auth_1' ? = Auth_1$, then U_B is verified, and both U_A and U_B obtain the same session key $sk_A = H_2(TID_A, TID_B, b_{A0}, b_{B0}, n_1, n_4, k_2')$.

Figure 1. AKE phase of Guo et al.'s scheme.

5. Cryptanalysis of Guo et al.'s Scheme

As described in Section 3, Attacker $\mathcal{A}$ can eavesdrop on, intercept, and forge transmitted messages, as well as extract parameters from the server. This information can be leveraged to launch various security attacks, including replay attacks, user impersonation, server impersonation, verification table leakage attacks, privileged insider attacks, and DoS attacks, all of which are described in more detail below.

5.1. Replay Attack

If an attacker $\mathcal{A}$ eavesdrops on messages sent or received over a public channel, $\mathcal{A}$ can impersonate you by resending the intercepted messages from a previous session. The detailed process is as follows:

Step 1: $\mathcal{A}$ records a message M_0 from user A to the server, replaces b_{A0} in M_0 with his public key $b_{\mathcal{A}}$, and sends it to the server.

Step 2: The server cannot distinguish whether M_0 is from User A or $\mathcal{A}$. Therefore, $\mathcal{A}$ can start a session by impersonating User A.

Step 3: User B cannot tell whether it is User A or $\mathcal{A}$ who wants to generate the session key. Therefore, User B uses $b_{\mathcal{A}}$ to compute σ_2, k_2, and $Auth_1$.

Step 4: $\mathcal{A}$ can use his secret key $s_{\mathcal{A}}$ to compute the session key sk_1 with the server, and decrypt c_1 to get σ_2 and $Auth_1$. $\mathcal{A}$ can use σ_2, user B's public key b_{B0} and his own secret key $s_{\mathcal{A}}$ to generate v_2 and k_2, and can generate a session key sk_A with user B.

Guo et al. cannot verify the freshness of the message and cannot verify the identity of the sender of the message.

Therefore, Guo et al.'s scheme cannot resist replay attack.

5.2. Impersonation Attack

Attacker $\mathcal{A}$ can attempt user and server impersonation attacks. The process is as follows:

- User impersonation: The attacker $\mathcal{A}$ must generate a message M_0 to impersonate the user. $\mathcal{A}$ chooses a secret key s_{A0} and an error e_{A0} random number n_1 and computes a public key $b_{A0} = As_{A0} + e_{A0}$. Then $\mathcal{A}$ generates a message $M_0 = (TID_A, TID_B, n_1, b_{A0})$ and sends it to the server to initiate a session. The server does not know whether M_0 is a legitimate message or from an attacker.
- Server impersonation: The attacker $\mathcal{A}$ needs to generate message M_1 to impersonate the server. $\mathcal{A}$ chooses a secret key s_{S1}, an error e_{S2}, and a random number n_2, and computes a public key $b_{S1} = As_{S1} + e_{S2}$. Then $\mathcal{A}$ generates a message $M_1 = (TID_A, TID_B, n_1, n_2, b_{A0}, b_{S1})$ and sends it to user B, with whom user A wants to communicate. User B does not know whether M_1 is a legitimate message or a message from an attacker.

Therefore, Guo et al.'s scheme cannot resist impersonation attack.

5.3. Verification Table Leakage Attack

If the attacker $\mathcal{A}$ extracts the server's verification table, $\mathcal{A}$ can perform an impersonation attack. After receiving the message $M_1 = (TID_A, TID_B, n_1, n_2, b_{A0}, b_{S1})$, $\mathcal{A}$ sets the secret key $s_{\mathcal{A}}$ and error values $e_{\mathcal{A}0}$, $e_{\mathcal{A}1}$, and $e_{\mathcal{A}2}$ according to the protocol, then calculates k_1 and k_2 accordingly. Using k_1, $\mathcal{A}$ can generate $Auth_0$, the authentication information with the server, as well as the session key sk_0. Using k_2, $\mathcal{A}$ can generate $Auth_1$ and sk_I, the authentication information with User A. It then creates a message M_2 and sends it to the server. The server cannot distinguish whether the message M_2 it receives is legitimate or from an attacker.

Therefore, Guo et al.'s scheme cannot resist verification table leakage attacks.

5.4. Privileged Insider Attack

An attacker $\mathcal{A}$ can obtain sensitive values such as ID_i, PW_1 from the registration message with the help of a privileged insider. Based on these values, $\mathcal{A}$ can attempt a user impersonation attack.

$\mathcal{A}$ generates its own secret key $s_{\mathcal{A}}$ and error values $e_{\mathcal{A}0}$, $e_{\mathcal{A}1}$, and $e_{\mathcal{A}2}$, and computes k_1 and k_2. Using k_1, $\mathcal{A}$ can calculate the authentication information with the server, $Auth_0$, and the session key, sk_0. Using k_2, $\mathcal{A}$ can calculate the authentication information with user A, $Auth_1$, and the session key $sk_{\mathcal{A}}$. If $\mathcal{A}$ generates a message M_2 using the value generated using his secret key and sends it to the server, the server will not know whether M_2 is a legitimate message or a message from the attacker.

Therefore, Guo et al.'s scheme cannot resist privileged insider attack.

5.5. Lack of Mutual Authentication

Guo et al. asserted that their protocol offers mutual authentication. In the AKE phase, $Auth_0$ is used by the server to authenticate User B, while $Auth_1$ is used by User A to authenticate User B. However, the protocol lacks a mechanism for the server to authenticate User A and for User B to authenticate both the server and User A. Therefore, the scheme proposed by Guo et al. does not fully achieve mutual authentication.

5.6. Correctness

There was a problem with the structure proposed by Guo et al. that prevented two users from generating a session key in a normal way, which is detailed below.

- To generate MLWE-based secret and public keys, the equation $b_{S0} = A^T s_{S0} + e_{S0}$ needs to be modified to $b_{S0} = As_{S0} + e_{S0}$. This change is necessary to avoid inconsistency issues during the calculation process.
- When the session key sk_1 between the server and user A is generated, the hash function H_2 is used by the server, but when user A generates the session key sk_1, the hash function H_1 is used. This means that the session key between the server and user A is not generated properly, so the ciphertext c_1 cannot be decrypted normally. As a result, the session key between user A and user B cannot be generated normally.

5.7. DoS Attack

A malicious attacker $\mathcal{A}$ can attempt to send the message M_0 repeatedly. Since no login is required to generate M_0, and the server does not authenticate the user who sent it, the scheme proposed by Guo et al. is vulnerable to DoS attacks.

Therefore, Guo et al.'s scheme cannot effectively resist DoS attacks.

6. Proposed Scheme

In this section, we propose a three-party mutual authentication and key exchange protocol using MLWE. The proposed scheme consists of system initialization, registration, authentication, and key exchange phases. Figure 2 illustrates the flowchart of the proposed scheme.

6.1. System Initialization Phase

In our proposed protocol, we use the same open parameters and initialization process as those in the password-based AKE scheme proposed by Guo et al. Specifically, R and R_q denote $\mathbb{Z}[x]/(x^n + 1)$ and $\mathbb{Z}q[x]/(x^n + 1)$, respectively. Lowercase italics represent column vectors with elements in the ring R or R_q, while uppercase italics represent matrices. The open parameters (n, q, k) are defined, where n is a power of 2 and q is an odd prime number that satisfies $q \equiv 1 \mod n$. For instance, $A \leftarrow R_q^{k \times k}$ is a matrix sampled uniformly from $R_q^{k \times k}$, and $(s, e) \leftarrow \beta_\eta^k \times \beta_\eta^k$ represents a private key and noise sampled uniformly from β_η^k. The parameter A, which is used to generate the user's public key during the session, is a value that is prepublished by the server.

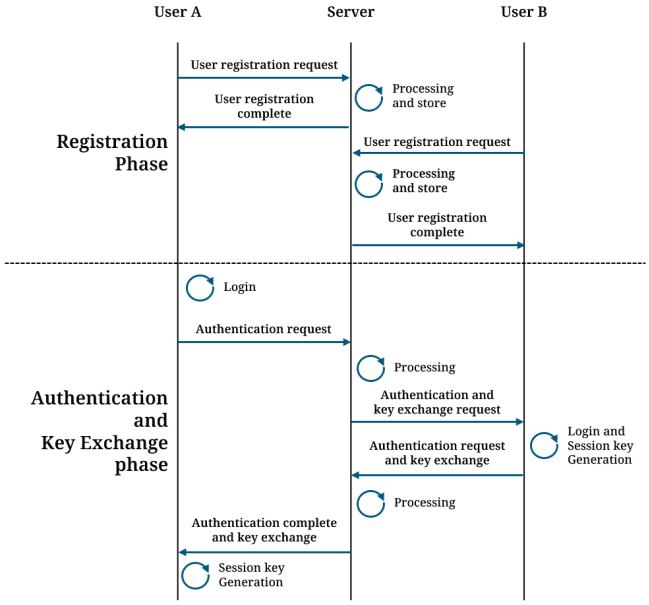


Figure 2. Flowchart of the proposed scheme.

6.2. Registration Phase

Before the authentication and session key generation process for communication, the user must complete a registration process with the server through a secure channel. During this process, the user registers information such as their identity, password, and biometric data. The detailed process is outlined below and illustrated in Figure 3.

Step 1: User i enters ID_i, PW_i, and Bio_i and selects a random number n_i. After that, i calculates $HID_i = h(ID_i||n_i)$, $HPW_i = h(PW_i||h_b(Bio_i))$, $GPW_i = HPW_i \oplus n_i$, and sends $\{HID_i, GPW_i\}$ to the server as a registration request message via a secure channel.

Step 2: After receiving the user's registration request, the server calculates $SID_i = h(HID_i||s)$, $d_i = SID_i \oplus GPW_i$, and sends $\{d_i\}$ to the user through a secure channel.

Step 3: After receiving $\{d_i\}$ from the server, the user calculates $RPW_i = h(ID_i||PW_i||h_b(Bio_i))$, $N_i = n_i \oplus RPW_i$, $V_i = h(HID_i||HPW_i||RPW_i||n_i)$, $Q_i = d_i \oplus n_i = SID_i \oplus HPW_i$, and stores $\{N_i, V_i, Q_i\}$ in the smartcard.

User i		**Server** S
Input ID_i, PW_i, Bio_i		
Random number n_i		
Computes		
$HID_i = h(ID_i\|\|n_i)$		
$HPW_i = h(PW_i\|\|h_b(Bio_i))$		
$GPW_i = HPW_i \oplus n_i$		
	$\{HID_i, GPW_i\} \longrightarrow$	
		$SID_i = h(HID_i\|\|s)$
		$d_i = SID_i \oplus GPW_i$
	$\longleftarrow \{d_i\}$	
$RPW_i = h(ID_i\|\|PW_i\|\|h_b(Bio_i))$		
$N_i = n_i \oplus RPW_i$		
$V_i = h(HID_i\|\|HPW_i\|\|RPW_i\|\|n_i)$		
$Q_i = d_i \oplus n_i = SID_i \oplus HPW_i$		
store $\{N_i, V_i, Q_i\}$ in smartcard		

Figure 3. Registration process of proposed scheme.

6.3. Authentication and Key Exchange Phase

In this process, the user logs in and uses the server to authenticate and generate a shared session key. The detailed process is as follows and is illustrated in Figure 4.

Step 1: U_A inputs ID_A, PW_A, Bio_A, and calculates $RPW_A = h(ID_A||PW_A||h_b(Bio_A))$, $n_A = N_A \oplus RPW_A$. $HID_A = h(ID_A||n_A)$, $HPW_A = h(PW_A||h_b(Bio_A))$, $V'_A = h(HID_A||HPW_A||RPW_A||n_A)$. $V_{AS} \stackrel{?}{=} V'_{AS}$ to authenticate the user, sample $(s_A, e_A) \leftarrow \beta_\eta^k \times \beta_\eta^k$ uniformly, and compute the public key $b_A = As_A + e_A$. A random number n_1 and timestamp T_1 are generated, then the following is computed: $SID_A = Q_A \oplus n_A \oplus HPW_A$, $V_{AS} = h(HID_A||SID_A||b_A||n_1||T_1)$, $N_1 = n_1 \oplus SID_A$. Finally, U_A sends the message $M_1 = (HID_A, HID_B, b_A, V_{AS}, N_1, T_1)$ to the server over an open channel.

Step 2: After receiving U_A's message, the server checks the timestamp $|T_1 - T_c| \leqq \Delta T$. If the timestamp is valid, the server computes $SID_A = h(HID_A||s)$, $n_1 = N_1 \oplus SID_A$, $V'_{AS} = h(HID_A||SID_A||b_A||n_1||T_1)$. And $V_{AS} \stackrel{?}{=} V'_{AS}$ to be certain that message M_1 came from user U_A. The server then chooses a random value n_2 and timestamp T_2. The server calculates $SID_B = h(HID_B||s)$, $V_{SB} = h(HID_B||SID_B||n_2||T_2)$, $N_2 = n_2 \oplus SID_B$, generates message $M_2 = (HID_A, b_A, V_{SB}, N_2, T_2)$, and sends it to user U_B over an open channel.

Step 3: U_B checks the timestamp $|T_2 - T_c| \leqq \Delta T$. If the timestamp is valid, U_B takes ID_B, PW_B, Bio_B, computes $RPW_B = h(ID_B||PW_B||h_b(Bio_B))$, $n_B = N_B \oplus RPW_B$, $HID_B = h(ID_B||n_B)$, $HPW_B = h(PW_B||h_b(Bio_B))$, $V'_B = h(HID_B||HPW_B||RPW_B||n_B)$, and authenticates the user via $V_B \stackrel{?}{=} V'_B$. After the login process, U_B computes $SID_B = Q_B \oplus n_B \oplus HPW_B$, $n_2 = N_2 \oplus SID_B$, $V'_{SB} = h(HID_B||SID_B||n_2||T_2)$ and verifies that message M_2 came from the server via $V_{SB} \stackrel{?}{=} V'_{SB}$. After verification, $(s_B, e_{B1}, e_{B2}) \leftarrow \beta_\eta^k \times \beta_\eta^k \times \beta_\eta$ is sampled uniformly and the public key $b_B = As_B + e_{B1}$ is computed. U_B computes $v_{AB} = b_A^T s_B + e_{B2}$, $\bar{v}_{AB} = dbl(v_{AB})$, $\sigma_{AB} = \langle \bar{v}_{AB} \rangle_{2q,2}$, $k_{AB} = \lfloor \bar{v}_{AB} \rceil_{2q,2}$ using a randomized doubling function, a cross-rounding function, and a modular rounding function. U_B generates random numbers n_3, n_4 and timestamp T_3, and calculates $V_{BS} = h(HID_B||SID_B||b_B||n_2||n_3||T_3)$, $N_3 = n_3 \oplus SID_B$, $V_{BA} = h(HID_A||HID_B||k_{AB}||n_4)$,$N_4 = n_4 \oplus k_{AB}$, $SK_{AB} = h(HID_A||HID_B||b_A||b_B||k_{AB})$. Here, SK_{AB} is the session key between U_A and U_B. U_B generates a message $M_3 = (HID_A, HID_B, b_A, \sigma_{AB}, V_{BS}, V_{BA}, N_3, N_4, T_3)$ and sends it to the server over the open channel.

Step 4: After receiving the message from U_B, the server checks the timestamp $|T_3 - T_c| \leqq \Delta T$. If the timestamp is valid, the server computes $n_3 = N_3 \oplus SID_B$, $V'_{BS} = h(HID_B||SID_B||b_B||n_2||n_3||T_3)$, and $V_{BS} \stackrel{?}{=} V'_{BS}$ to verify that message M_3 came from user U_B. The server then chooses a random value n_5 and timestamp T_4. The server calculates $V_{SA} = h(HID_A||SID_A||n_5||T_4)$, $N_5 = n_5 \oplus SID_A$, generates message $M_4 = (HID_A, b_B, \sigma_{AB}, V_{SA}, V_{BA}, N_4, N_5T_4)$, and sends it to user U_A over the open channel.

Step 5: After receiving message M_4 from the server, U_A checks the timestamp $|T_4 - T_c| \leqq \Delta T$. If the timestamp is valid, U_A computes $n_5 = N_5 \oplus SID_A$, $V'_{SA} = h(HID_A||SID_A||n_5||T_4)$ and uses $V_{SA} \stackrel{?}{=} V'_{SA}$ to verify that message M_4 came from the server. After that, U_A uses Peikert's error reconciliation function to compute $v'_{AB} = b_B^T s_A$, $k'_{AB} = Rec(2v'_{AB}, \sigma_{AB}) \in 0, 1^n$, computes $n_4 = N_4 \oplus k'_{AB}$, $V'_{BA} = h(HID_A||HID_B||k'_{AB}||n_4)$, and uses $V_{BA} \stackrel{?}{=} V'_{BA}$ to verify that U_B is the one with whom U_A wants to create a session key. Finally, U_A computes the session key $SK_{AB} = h(HID_A||HID_B||b_A||b_B||k'_{AB})$.

User A	**Server** S	**User** B
Inputs ID_A, PW_A, Bio_A		
Computes		
$RPW_A = h(ID_A \| \| PW_A \| \| h_b(Bio_A))$		
$n_A = N_A \oplus RPW_A$		
$HID_A = h(ID_A \| \| n_A)$		
$HPW_A = h(PW_A \| \| h_b(Bio_A))$		
$V'_A = h(HID_A \| \| HPW_A \| \| RPW_A \| \| n_A)$		
Verifies $V_A \stackrel{?}{=} V'_A$		
$(s_A, e_A) \leftarrow \beta_\eta^k \times \beta_\eta^k$		
$b_A = As_A + e_A$		
Random number n_1		
Timestamp T_1		
Computes		
$SID_A = Q_A \oplus n_A \oplus HPW_A$		
$V_{AS} = h(HID_A \| \| SID_A \| \| b_A \| \| n_1 \| \| T_1)$		
$N_1 = n_1 \oplus SID_A$		
	$M_1 = (HID_A, HID_B, b_A, V_{AS}, N_1, T_1) \longrightarrow$	
	Checks $\lvert T_1 - T_c \rvert \leqq \Delta T$	
	$SID_A = h(HID_A \| \| s)$	
	$n_1 = N_1 \oplus SID_A$	
	$V'_{AS} = h(HID_A \| \| SID_A \| \| b_A \| \| n_1 \| \| T_1)$	
	Verifies $V_{AS} \stackrel{?}{=} V'_{AS}$	
	Random number n_2	
	Timestamp T_2	
	Computes	
	$SID_B = h(HID_B \| \| s)$	
	$V_{SB} = h(HID_B \| \| SID_B \| \| n_2 \| \| T_2)$	
	$N_2 = n_2 \oplus SID_B$	
		$M_2 = (HID_A, b_A, V_{SB}, N_2, T_2) \longrightarrow$
		Checks $\lvert T_2 - T_c \rvert \leqq \Delta T$
		Inputs ID_B, PW_B, Bio_B
		Computes
		$RPW_B = h(ID_B \| \| PW_B \| \| h_b(Bio_B))$
		$n_B = N_B \oplus RPW_B$
		$HID_B = h(ID_B \| \| n_B)$
		$HPW_B = h(PW_B \| \| h_b(Bio_B))$
		$V'_B = h(HID_B \| \| HPW_B \| \| RPW_B \| \| n_B)$
		Verifies $V_B \stackrel{?}{=} V'_B$
		$SID_B = Q_B \oplus n_B \oplus HPW_B$
		$n_2 = N_2 \oplus SID_B$
		$V'_{SB} = h(HID_B \| \| SID_B \| \| n_2 \| \| T_2)$
		Verifies $V_{SB} \stackrel{?}{=} V'_{SB}$
		$(s_B, e_{B1}, e_{B2}) \leftarrow \beta_\eta^k \times \beta_\eta^k \times \beta_\eta$
		$b_B = As_B + e_{B1}$
		$v_{AB} = b_A^T s_B + e_{B2}$
		$\bar{v}_{AB} = dbl(v_{AB})$
		$\sigma_{AB} = \langle \bar{v}_{AB} \rangle_{2q,2}$
		$k_{AB} = \lfloor \bar{v}_{AB} \rceil_{2q,2}$
		Random number n_3, n_4
		Timestamp T_3
		Computes
		$V_{BS} = h(HID_B \| \| SID_B \| \| b_B \| \| n_2 \| \| n_3 \| \| T_3)$
		$N_3 = n_3 \oplus SID_B$
		$V_{BA} = h(HID_A \| \| HID_B \| \| k_{AB} \| \| n_4)$
		$N_4 = n_4 \oplus k_{AB}$
		$SK_{AB} = h(HID_A \| \| HID_B \| \| b_A \| \| b_B \| \| k_{AB})$
		$M_3 = (HID_A, HID_B, b_B, \sigma_{AB}, V_{BS}, V_{BA}, N_3, N_4, T_3) \longleftarrow$
	Checks $\lvert T_3 - T_c \rvert \leqq \Delta T$	
	$n_3 = N_3 \oplus SID_B$	
	$V'_{BS} = h(HID_B \| \| SID_B \| \| b_B \| \| n_2 \| \| n_3 \| \| T_3)$	
	Verifies $V_{BS} \stackrel{?}{=} V'_{BS}$	
	Random number n_5	
	Timestamp T_4	
	$V_{SA} = h(HID_A \| \| SID_A \| \| n_5 \| \| T_4)$	
	$N_5 = n_5 \oplus SID_A$	
	$M_4 = (HID_A, b_B, \sigma_{AB}, V_{SA}, V_{BA}, N_4, N_5 T_4) \longleftarrow$	
Checks $\lvert T_4 - T_c \rvert \leqq \Delta T$		
$n_5 = N_5 \oplus SID_A$		
$V'_{SA} = h(HID_A \| \| SID_A \| \| n_5 \| \| T_4)$		
Verifies $V_{SA} \stackrel{?}{=} V'_{SA}$		
$v'_{AB} = b_B^T s_A$		
$k'_{AB} = Rec(2v'_{AB}, \sigma_{AB}) \in 0, 1^n$		
$n_4 = N_4 \oplus k'_{AB}$		
$V'_{BA} = h(HID_A \| \| HID_B \| \| k'_{AB} \| \| n_4)$		
Verifies $V_{BA} \stackrel{?}{=} V'_{BA}$		
$SK_{AB} = h(HID_A \| \| HID_B \| \| b_A \| \| b_B \| \| k'_{AB})$		

Figure 4. AKE phase of proposed scheme.

7. Security Analysis

7.1. Formal Analysis Using the BAN Logic

In the proposed scheme, users and a server perform mutual authentication to prove that they are valid entities of the system. To demonstrate the mutual authentication of the proposed scheme, we utilize a widely recognized formal verification technique called BAN logic [27]. Many researchers have affirmed the mutual authentication of their approaches using BAN logic [28,29]. The BAN logic analysis is performed sequentially by setting protocol goals, idealized form of the protocol, establishing assumptions, and proof using BAN logic. To incorporate our approach with BAN logic, we provide the following notations and descriptions. Table 3 is the notation used in BAN logic.

Table 3. BAN logic notations.

Notation	Description
ρ_1, ρ_2	Two principals
s_1, s_2	Two statements
$\rho_1 \mid \equiv s_1$	ρ_1 believes s_1
$\rho_1 \mid \sim s_1$	ρ_1 once said s_1
$\rho_1 \Rightarrow s_1$	ρ_1 controls s_1
$\rho_1 \lhd \mu_1$	ρ_1 receives s_1
$\#s_1$	s_1 is fresh
$(s_1)_K$	s_1 is encrypted with K
$\rho_1 \overset{K}{\leftrightarrow} \rho_2$	ρ_1 and ρ_2 have shared key K
SK	The session key

7.1.1. BAN Logic Rules

1. Message meaning rule (MMR):

$$\frac{\rho_1 \mid \equiv \rho_1 \overset{K}{\leftrightarrow} \rho_2, \quad \rho_1 \lhd (s_1)_K}{\rho_1 \mid \equiv \rho_2 \mid \sim s_1}$$

2. Nonce verification rule (NVR):

$$\frac{\rho_1 \mid \equiv \#(s_1), \quad \rho_1 \mid \equiv \rho_2 \mid \sim s_1}{\rho_1 \mid \equiv \rho_2 \mid \equiv s_1}$$

3. Jurisdiction rule (JR):

$$\frac{\rho_1 \mid \equiv \rho_2 \mid \Longrightarrow s_1, \quad \rho_1 \mid \equiv \rho_2 \mid \equiv s_1}{\rho_1 \mid \equiv s_1}$$

4. Belief rule (BR):

$$\frac{\rho_1 \mid \equiv (s_1, s_2)}{\rho_1 \mid \equiv s_1}$$

5. Freshness rule (FR):

$$\frac{\rho_1 \mid \equiv \#(s_1)}{\rho_1 \mid \equiv \#(s_1, s_2)}$$

7.1.2. Goals

The goals are to show that U_A and U_B believe that they agreed upon the same session key.

Goal 1: $U_A| \equiv U_A \overset{SK}{\longleftrightarrow} U_B$

Goal 2: $U_A| \equiv U_B| \equiv U_A \overset{SK}{\longleftrightarrow} U_B$

Goal 3: $U_B| \equiv U_A \overset{SK}{\longleftrightarrow} U_B$

Goal 4: $U_B| \equiv U_A| \equiv U_A \overset{SK}{\longleftrightarrow} U_B$

7.1.3. Idealized Forms

The idealized forms of the messages exchanged during the authentication can be described as follows.

Msg_1 : $U_A \rightarrow S : (HID_A, HID_B, b_A, \{n_1\}_{SID_A}, T_1)$

Msg_2 : $S \rightarrow U_B : (HID_A, b_A, \{n_2\}_{SID_B}, T_2)$

Msg_3 : $U_B \rightarrow S : (HID_A, HID_B, b_B, \{n_4\}_{SID_B}, T_3)$

Msg_4 : $S \rightarrow U_A : (HID_B, b_B, \{n_5\}_{SID_A}, T_4)$

7.1.4. Assumptions

The basic assumptions for the BAN logic proof are as follows:

A_1: $S| \equiv \#(T_1)$

A_2: $U_B| \equiv \#(T_2)$

A_3: $S| \equiv \#(T_3)$

A_4: $U_A| \equiv \#(T_4)$

A_5: $U_A| \equiv U_A \overset{SID_A}{\longleftrightarrow} S$

A_6: $S| \equiv U_A \overset{SID_A}{\longleftrightarrow} S$

A_7: $U_B| \equiv U_B \overset{SID_B}{\longleftrightarrow} S$

A_8: $S| \equiv U_B \overset{SID_B}{\longleftrightarrow} S$

A_9: $U_A| \equiv U_B \Rightarrow (U_A \overset{SK}{\longleftrightarrow} U_B)$

A_{10}: $U_B| \equiv U_A \Rightarrow (U_A \overset{SK}{\longleftrightarrow} U_B)$

A_{11}: $U_A| \equiv U_A \overset{k_{AB}}{\longleftrightarrow} U_B$

A_{12}: $U_B| \equiv U_A \overset{k_{AB}}{\longleftrightarrow} U_B$

7.1.5. BAN Logic Proof

We perform the BAN logic proof of the proposed protocol as below:

Step 1: S receives Msg_1.

$$S_1 : S \lhd (HID_A, HID_B, b_A, \{n_1\}_{SID_A}, T_1)$$

Step 2: S can recognize that n_1 is from U_A by A_6.

$$S_2 : S| \equiv U_A| \sim (HID_A, HID_B, b_A, n_1, T_1)$$

Step 3: S can consider that the message is fresh by A_1 and the FR.

$$S_3 : S| \equiv \#(HID_A, HID_B, b_A, n_1, T_1)$$

Step 4: S can apply the NVR using S_2 and S_3.

$$S_4 : S| \equiv U_A| \equiv (HID_A, HID_B, b_A, n_1, T_1)$$

Step 5: U_B receives Msg_2.

$$S_5 : U_B \lhd (HID_A, b_A, \{n_2\}_{SID_B}, T_2)$$

Step 6: U_B can recognize that n_2 is from S by A_7.

$$S_6 : U_B| \equiv S| \sim (HID_A, b_A, n_2, T_2)$$

Step 7: U_B can consider that the message is fresh by A_2 and the FR.

$$S_7 : U_B| \equiv \#(HID_A, b_A, n_2, T_2)$$

Step 8: U_B can apply the NVR using S_6 and S_7.

$$S_8 : U_B| \equiv S| \equiv (HID_A, b_A, n_2, T_2)$$

Step 9: S receives Msg_3.

$$S_9 : S \lhd (HID_A, HID_B, b_B, \{n_4\}_{SID_B}, T_3)$$

Step 10: S can recognize that n_4 is from U_B by A_8.

$$S_{10} : S| \equiv U_B| \sim (HID_A, HID_B, b_B, \{n_4\}_{SID_B}, T_3)$$

Step 11: S can consider that the message is fresh by A_3 and the FR.

$$S_{11} : S| \equiv \#(HID_A, HID_B, b_B, \{n_4\}_{SID_B}, T_3)$$

Step 12: S can apply the NVR using S_{10} and S_{11}.

$$S_{12} : S| \equiv U_B| \equiv (HID_A, HID_B, b_B, \{n_4\}_{SID_B}, T_3)$$

Step 13: U_A receives Msg_4.

$$S_{13} : U_A \lhd (HID_B, b_B, \{n_5\}_{SID_A}, T_4)$$

Step 14: U_A can recognize that n_5 is from S by A_5.

$$S_{14} : U_A| \equiv S| \sim (HID_A, HID_B, b_A, n_1, T_1)$$

Step 15: S can consider that the message is fresh by A_1 and the FR.

$$S_{15} : U_A| \equiv \#(HID_B, b_B, \{n_5\}_{SID_A}, T_4)$$

Step 16: U_A can apply the NVR using S_{14} and S_{15}.

$$S_{16} : U_A| \equiv S| \equiv (HID_B, b_B, \{n_5\}_{SID_A}, T_4)$$

Steps 17 and 18: U_A considers that Msg_4 is from U_B and U_B considers that Msg_2 is from U_A by A_{11} and A_{12}, respectively. Therefore, U_A and U_B can believe that the other party can calculate SK

$$S_{17} : U_A| \equiv U_B| \equiv (U_A \overset{SK}{\longleftrightarrow} U_B) \quad \text{(Goal 2)}$$
$$S_{18} : U_B| \equiv U_A| \equiv (U_A \overset{SK}{\longleftrightarrow} U_B) \quad \text{(Goal 4)}$$

Steps 19 and 20: The JR can be applied to S_{17} and S_{18} using A_{10} and A_{11}, respectively.

$$S_{19} : U_A| \equiv (U_A \overset{SK}{\longleftrightarrow} U_B) \quad \text{(Goal 2)}$$
$$S_{20} : U_A| \equiv (U_A \overset{SK}{\longleftrightarrow} U_B) \quad \text{(Goal 2)}$$

Finally, the user and the server have mutually authenticated each other.

7.2. Formal Analysis Using the RoR Model

We prove the semantic security of the proposed protocol using the RoR model [30]. Under the RoR model, an adversary A can attempt passive and active queries (i.e., attacks) to obtain information about the session key SK. The notations and queries used in the RoR model are as follows. p^t is a participant with the t^{th} instance. For example, $p_{U_A}^{t_1}$, $p_S^{t_2}$, and $p_{U_B}^{t_3}$ represent participants U_A, S, U_B with t_1^{th}, t_2^{th}, and t_2^{th} instances, respectively. Table 4 presents queries executed by A under the RoR model.

Table 4. Queries of the RoR model.

Query	Description
$Execute(p_{U_A}^{t_1}, p_S^{t_2}, p_{U_B}^{t_3})$	This query represents passive attack (i.e., eavesdropping) executed by A. Using this query, A can obtain transmitted messages in the authentication phase.
$CorruptSC(p_{U_A}^{t_1})$	This query represents the smart card stolen attack executed by A. A steals the smart card of $p_{U_A}^{t_1}$ and extracts the stored values through the power analysis attack.
$Send(p^t, M)$	A can generate a message M and send M to a network participant and receive the response using this query.
$Test(A)$	At the end of each game, A executes this query. When *Test* is executed, an unbiased coin c is flipped. If c lands on heads, A returns SK, and if c lands on tails, A returns a random number. Otherwise, A returns $NULL$. Then, A must guess whether the returned value is SK or a random number. If A successfully guesses the returned value, A wins the game and the proposed protocol cannot guarantee semantic security.

A can obtain the output value of the hash function for the input by performing the $Hash$ query. When $Adv(A)$ is the probability of the advantage for A to break the session key, we prove the following equation:

$$Adv(A) \leq \frac{q_h^2}{|Hash|} + \frac{q_{send}}{|D_{id}||D_{pw}|} + Adv_A^{MLWE}, \tag{1}$$

where q_h represents the number of hash queries performed by A, $|Hash|$ represents the range space of cryptographic hash function, q_{send} represents the number of $Send$ queries performed by A, and $|D_{id}|$ and $|D_{pw}|$, respectively, represent the size of the password dictionary. Adv_A^{MLWE} represents the probability that A solves the MLWE problem in the polynomial time.

Proof. Under the RoR model, A plays games $G_i (i = 0, 1, 2, 3, 4)$. We denote that $Adv^{G_i}(A)$ is a advantage of A to calculate SK after the end of G_i.

- G_0: A selects the bit c at the start of the game. In the first game, A has no information about SK and no queries to perform. By the definition of semantic security, we can induce the following equation.

$$Adv(A) = |2Adv_{G_0}(A) - 1|. \tag{2}$$

- G_1: A attempts $Execute$ query (i.e., eavesdropping attack) in this game. Then, A can obtain the transmitted messages through wireless channels. After that, A performs $Test$ query and receives the return value. Finally, A tries to guess whether the return value is SK or a random number. To win the game, A should be able to calculate SK, which is computed by HID_A, HID_B, b_A, b_B, and k_{AB}. Among these values, A cannot obtain k_{AB}, which is a Diffie–Hellman key between HID_A and HID_B. Therefore, A has no advantage in this game, and the advantage function of A after at the end of G_1 has no change.

$$Adv_{G_0}(A) = Adv_{G_1}(A). \tag{3}$$

- G_2: In this game, A uses the $Send$ and $Hash$ queries to compromise SK. From the transmitted messages, A can obtain M_0, M_1, M_2, and M_3. A must obtain HID_A, HID_B, b_A, b_B, and k_{AB} to calculate the session key. However, A cannot obtain k_{AB} from the transmitted messages in the public channel; therefore, the only way for A to compromise the session key is to find the hash collision. The advantage function at the end of G_2 is as follows by the birthday paradox.

$$|Adv_{G_2}(A) - Adv_{G_1}(A)| \leq \frac{q_h^2}{2|Hash|}. \tag{4}$$

- G_3: A can conduct $CorruptSC$ query. For WLOS, we assume that A obtains the smart card of U_A and A extracts (N_i, V_i, Q_i) from the smart card. However, these values are masked with ID_A and PW_B, and A must succeed in guessing ID_A and PW_B simultaneously to login to the smart card. At the end of G_3, we can induce the following equation.

$$|Adv_{G_3}(A) - Adv_{G_2}(A)| \leq \frac{q_{send}}{|D_{id}||D_{pw}|}. \tag{5}$$

- G_4: A failed to compromise SK through the above queries. Finally, A can defeat the cryptosystem by solving the MLWE problem in the polynomial time. The probability can be represented as follows:

$$|Adv_{G_4}(A) - Adv_{G_3}(A)| \leq Adv_A^{MLWE}. \tag{6}$$

After all the games are over, A should guess the correct bit c from $Test$ query. It follows that

$$Adv_{G_4}(A) = \frac{1}{2}. \tag{7}$$

From Equations (2) and (3), we can obtain the following equation:

$$\begin{aligned} \frac{1}{2}Adv(A) &= |Adv_{G_1}(A) - \frac{1}{2}| \\ &= |Adv_{G_1}(A) - Adv_{G_4}(A)|. \end{aligned} \tag{8}$$

We can apply the triangular inequality to Equation (8).

$$\begin{aligned} |Adv_{G_4}(A) - Adv_{G_1}(A)| &\leq |Adv_{G_4}(A) - Adv_{G_3}(A)| + |Adv_{G_3}(A) - Adv_{G_2}(A)| \\ &\quad + |Adv_{G_2}(A) - Adv_{G_1}(A)| \\ &\leq \frac{q_h^2}{|Hash|} + \frac{q_{send}}{|D_{id}||D_{pw}|} + Adv_A^{MLWE}. \end{aligned} \tag{9}$$

We can obtain the following equation using Equations (8) and (9).

$$Adv(A) \leq \frac{q_h^2}{|Hash|} + \frac{q_{send}}{|D_{id}||D_{pw}|} + Adv_A^{MLWE}, \tag{10}$$

which is the same as Equation (1). Finally, we prove the semantic security of the proposed protocol using the RoR model. □

7.3. Formal Analysis Using AVISPA Tool

We conduct a formal verification of our proposed protocol using the AVISPA tool (v1.0) [31], employing code written in Higher Level Protocol Specification Language (HLPSL). AVISPA is a widely recognized simulation tool used for validating the security of Internet protocols and applications, particularly assessing protocols' resistance to man-in-the-middle (MITM) and replay attacks.

AVISPA evaluates protocol safety through four backends: On-the-Fly Model Checker (OFMC), Constraint Logic-based Attack Detector (CL-AtSe), SAT-based Model Checker (SATMC), and Automatic Approximation-based Tree Automata for Security Protocol Analysis (TA4SP). Among these, the OFMC and CL-AtSe backends are predominantly used. OFMC is responsible for verifying resistance to MITM attacks, while CL-AtSe checks for replay attack resilience. The simulation results using AVISPA for the proposed protocol are presented in Figure 5, confirming that the protocol is secure against both MITM and replay attacks.

```
% OFMC
% Version of 2006/02/13
SUMMARY
  SAFE
DETAILS
  BOUNDED_NUMBER_OF_SESSIONS
PROTOCOL
  /home/span/span/testsuite/results/protocol.if
GOAL
  as_specified
BACKEND
  OFMC
COMMENTS
STATISTICS
  parseTime: 0.00s
  searchTime: 0.09s
  visitedNodes: 64 nodes
  depth: 6 plies
```

```
SUMMARY
  SAFE

DETAILS
  BOUNDED_NUMBER_OF_SESSIONS
  TYPED_MODEL

PROTOCOL
  /home/span/span/testsuite/results/protocol.if

GOAL
  As Specified

BACKEND
  CL-AtSe

STATISTICS

  Analysed  : 2 states
  Reachable  : 0 states
  Translation: 0.04 seconds
  Computation: 0.00 seconds
```

Figure 5. Analysis of AVISPA simulation using OFMC and CL-AtSe.

7.4. Informal Security Analysis

7.4.1. Insider Attack

A malicious actor $\mathcal{A}$, who has gone through the registration phase as a legitimate user, might attempt to launch an insider attack using the obtained information. However, since the user's public and secret keys are generated anew for each session, the attacker cannot determine the user's secret key and, therefore, cannot compute k'_{AB}. Additionally, the attacker cannot know the values of the random numbers n_1, n_2, n_3, n_4, n_5, which are also generated anew for each session. Furthermore, the attacker cannot compute $SID_{\mathcal{A}}$ because they do not have access to the user's biometric information or the server's secret key. Therefore, the proposed scheme is secure against insider attacks.

7.4.2. Impersonation Attack

- User impersonation: For an attacker $\mathcal{A}$ to impersonate the legitimate user U_A, they must generate a valid message $M_1 = (HID_A, HID_B, b_A, V_{AS}, N_1, T_1)$. However, since $\mathcal{A}$ cannot obtain U_A's biometric information, they cannot derive SID_A, and, consequently, cannot compute V_{AS}. Without V_{AS}, the attacker cannot generate the message M_1 to send to the server. Therefore, the proposed scheme is secure against user impersonation attacks.
- Server impersonation: To execute this attack, $\mathcal{A}$ would need to send a message $M_2 = (HID_A, b_A, V_{SB}, N_2, T_2)$ to user U_B while pretending to be the server. However, even if $\mathcal{A}$ intercepts the transmission of M_1 over an open channel, they cannot generate the required V_{SB} and N_2 for the message without knowing SID_B. Therefore, the proposed scheme is secure against server impersonation attacks.

7.4.3. Replay Attack

All users and servers validate the received messages using T_1, T_2, T_3, and T_4. Additionally, different public and private keys are used for each session, and freshness is ensured through the use of random numbers. Therefore, the proposed scheme is secure against replay attacks.

7.4.4. Man-in-the-Middle Attack

When an attacker $\mathcal{A}$ intercepts a message and attempts to tamper with it, the user or server receiving the message uses V_{AS}, V_{SB}, V_{BS}, V_{BA}, and V_{SA} to verify that the message was sent by a legitimate user or server. As a result, it is impossible for an attacker to send a message while pretending to be a legitimate user or to modify it in transit. Therefore, the proposed protocol is secure against MITM attacks.

7.4.5. Privileged Insider Attack

In this attack scenario, attacker $\mathcal{A}$ is considered an authorized insider, meaning they have access to confidential values such as the user's enrollment request message HID_i, GPW_i. However, without the user's exact biometric information, ID, or password (PW) value, it is impossible to calculate RPW_i or $SID_i = A_2 \oplus (\sigma_i || PW_i)$. As a result, the proposed scheme is secure against privileged insider attacks.

7.4.6. Stolen Smartcard Attack

If a malicious attacker $\mathcal{A}$ obtains the user's smartcard, they can extract the values of N_i, V_i, Q_i. However, without the user's biometric information, the value of n_i cannot be determined from these parameters. Therefore, the proposed scheme is secure against stolen smartcard attacks.

7.4.7. DoS Attack

For two users to mutually authenticate with the server and generate a session key for communication, they must first complete the login process. Afterward, the messages

are generated and sent, and the receiver validates the received messages. Therefore, the proposed scheme is secure against DoS attacks.

7.4.8. Session Key Disclosure Attack

To compute the session key $SK_{AB} = h(HID_A||HID_B||b_A||b_B||k'_{AB})$, the attacker $\mathcal{A}$ would need access to the values of HID_A, HID_B, b_A, b_B, and k'_{AB}. However, the value of k'_{AB} is known only to the two users who are generating the session key, as it is used temporarily and exists only within a single session. Therefore, the proposed scheme is secure against session key disclosure attacks.

7.4.9. Mutual Authentication

In the AKE phase, the values V_{AS}, V_{SB}, V_{BS}, V_{BA}, and V_{SA} contained in messages M_1, M_2, M_3, and M_4 enable the verification of identity between the user and the server, as well as between the users themselves. The transparency of authentication is ensured because the open channel prevents an attacker from exploiting SID_i, k_{AB} values, or arbitrary values. Therefore, the proposed scheme provides mutual authentication.

8. Performance Analysis

8.1. Security Features Comparison

We visually compared the security features of our proposed scheme with those of related schemes [14,20–22,32], and the results are summarized in Table 5. The table includes various security aspects such as "Insider Attack", "Impersonation Attack", "Replay Attack", "MITM Attack", "Privileged Insider Attack", "Stolen Smartcard Attack", "DoS Attack", "Session Key Disclosure", and "Mutual Authentication". The comparison suggests that our proposed scheme not only offers more security features than Guo et al.'s scheme but also presents fewer instances of unidentified or unavailable features when compared with other related schemes.

Table 5. Security and functionality features (SFFs) comparison.

SFF	[32]	[21]	[22]	[20]	[14]	Proposed
SP1	✓	✓	✓	✓	✓	✓
SP2	×	×	×	×	×	✓
SP3	✓	×	✓	×	×	✓
SP4	✓	✓	✓	✓	✓	✓
SP5	×	×	×	×	×	✓
SP6	×	×	✓	×	△	✓
SP7	×	✓	✓	✓	×	✓
SP8	✓	✓	✓	✓	✓	✓
SP9	✓	✓	✓	×	×	✓

Note: SP1: insider attack; SP2: impersonation attack; SP3: replay attack; SP4: MITM attack; SP5: privileged insider attack; SP6: stolen smartcard attack; SP7: DoS attack; SP8: session key disclosure; SP9: mutual authentication; ✓: provides safety/functional features; ×: does not provides safety/functional features; △: not verified.

8.2. Communication Costs Comparison

We conducted a comparative analysis of the communication costs between the proposed scheme and other related schemes [14,20–22,32].

To represent the communication costs of existing authentication protocols, we assume that each binary string, including passwords, nonces, timestamps, and identities, is 256 bits in size. For the hashing process, we use the Secure Hash Algorithm (SHA-512) with a 512-bit output. Symmetric encryption and decryption are performed using a 256-bit key with the AES-256 algorithm. An elliptic curve point is assumed to be 320 bits, and each element from $\mathbb{Q}_q$ has a 4094-bit size. In the RLWE-based scheme, the public key is 7424 bits using the NowHope512 algorithm, whereas the MLWE-based scheme uses a 6400-bit public key with the Kyber512 algorithm.

Therefore, the message $M_1 = HID_A, HID_B, b_A, V_{AS}, N_1, T_1$ exchanged in the AKE phase of the proposed scheme contains $(512 + 512 + 4096 + 512 + 160 + 64 = 5856$ bits). The message $M_2 = HID_A, b_A, V_{SB}, N_2, T_2$ requires (512 + 4096 + 512 + 160 + 64 = 5344 bits), $M_3 = HID_A, HID_B, b_B, \sigma_{AB}, V_{BS}, V_{BA}, N_3, N_4, T_3$ requires (512 + 512 + 4096 + 512 + 512 + 512 + 160 + 160 + 64 = 7040 bits), and $M_4 = HID_A, b_B, \sigma_{AB}, V_{SA}, V_{BA}, N_4, N_5, T_4$ requires (512 + 4096 + 512 + 512 + 512 + 160 + 160 + 64 = 6528 bits).

Table 6 summarizes the communication costs of related schemes [14,20–22,32] and the proposed scheme.

Table 6. Comparison analysis of communication costs.

Scheme	Total Cost (Bits)	Number of Messages
Islam et al. [32]	21,122 bits	8
Chaudhary et al. (2023) [21]	21,507 bits	5
Rewal et al. [22]	20,226 bits	4
Chaudhary et al. (2024) [20]	16,899 bits	4
Guo et al. [14]	31,520 bits	4
Proposed	24,768 bits	4

8.3. Computation Costs Comparison

We performed a comparative analysis of computation costs during the AKA phase of the proposed scheme and other related schemes [14,20–22,32]. Experiments were conducted on an Ubuntu 22.04.3 system with a 12th Gen Intel(R) Core(TM) i5-12400F CPU and 3.8 GB of RAM, utilizing OpenSSL [33], NewHope [34], and Crystal-Kyber [35], all part of NIST's "Post-Quantum Cryptography Standardization" project [36]. Hash functions are denoted as T_h, and AES-256 encryption/decryption as T_{symenc}/T_{symdec}. For RLWE and MLWE protocols, $T_{rkeygen}$, T_{renc}, T_{rdec}, $T_{rsample}$ and $T_{mkeygen}$, T_{menc}, T_{mdec} represent key generation, encapsulation, decapsulation, and sampling times. Elliptic curve point multiplication is denoted as T_{ecpm}. The average execution time for each operation is detailed in Table 7, with protocol computation times shown in Table 8 and Figure 6.

As shown in Table 8, the total computation cost of the proposed scheme is more efficient than that of other schemes. Even though the cost of [22] is better than the proposed scheme, it is vulnerable to impersonation attacks and privileged insider attacks. To ensure robustness against these vulnerabilities, our proposed protocol incorporates additional hash functions, resulting in a marginal increase in computational cost. From the perspective of user B, the proposed protocol requires slightly more computations compared to those of [14,20], but, overall, the computational cost makes the proposed approach far more efficient.

Table 7. Average run time in nanoseconds.

Operation	Symbol	Time (Nanosecond)
Hash function	T_h	1253
AES-256 encryption/decryption	T_{symenc}/T_{symdec}	1146/408
ECMP	T_{ecmp}	36,183
NewHope512-keygen	$T_{rkeygen}$	48,114
NewHope512-encap	T_{renc}	67,841
NewHope512-decap	T_{rdec}	17,672
NewHope512-sampling	$T_{rsample}$	6413
Crystal-Kyber512-keygen	$T_{mkeygen}$	27,275
Crystal-Kyber512-encap	T_{menc}	32,811
Crystal-Kyber512-decap	T_{mdec}	39,565

Table 8. Comparison analysis of computation costs in nanoseconds.

Scheme	User *A*	User *B*	Server	Total Average (A + B + S)
[32]	$8T_h+T_{rkeygen} + T_{renc} +T_{rdec}$	$8T_h+T_{rkeygen} + T_{renc} +T_{rdec}$	$4T_h$	≈292,314
[21]	$8T_h+T_{ecpm} +T_{rkeygen} + T_{renc} +T_{rdec}$	$19T_h + 8T_h+T_{ecpm} +T_{rkeygen} + T_{renc} +T_{rdec}$	$10T_h + 2\ T_{rdec} + T_{rsample}$	≈413,955
[22]	$7T_h+T_{rkeygen} +T_{rdec}$	$3T_h+T_{rkeygen} + T_{renc}$	$6T_h$	≈201,789
[20]	$9T_h+T_{rkeygen} + T_{renc}$	$10T_h+T_{rkeygen} + T_{renc}$	$8T_h + 2\ T_{rdec}$	≈301,085
[14]	$4T_h+ T_{mkeygen} + 2T_{mdec} + T_{symdec}$	$4T_h+ T_{mkeygen} + 2T_{menc} + T_{symenc}$	$4T_h+ T_{mkeygen} +T_{menc} +T_{mdec} + T_{symenc} + T_{symdec}$	≈317,097
Proposed	$10T_h+ T_{mkeygen} + T_{mdec}$	$10T_h+ T_{rkeygen} + T_{menc}$	$6T_h$	≈215,373

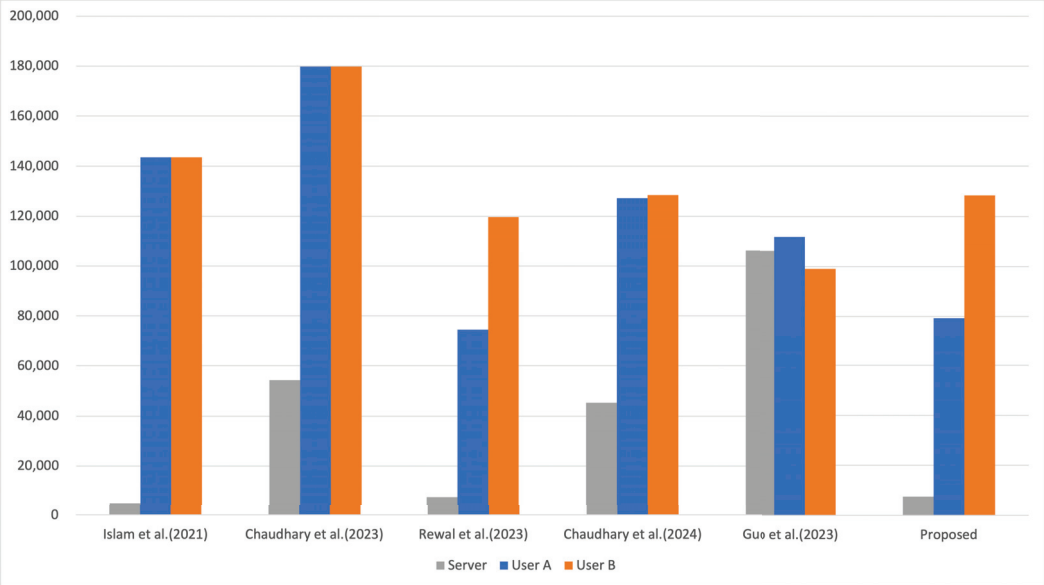


Figure 6. Computation costs comparison on the user's side [14,20–22,32].

9. Conclusions

In this study, we analyzed the three-party AKE protocol proposed by Guo et al. and identified several vulnerabilities, including susceptibility to impersonation attacks, replay attacks, privileged insider attacks, DoS attacks, verification table leakage attacks, and a lack of user mutual authentication. To address these vulnerabilities, we proposed a three-party secure AKE protocol based on MLWE.

The proposed protocol leverages the user's biometric information and the server's master key to prevent the exposure of sensitive parameters and to ensure security against various attacks. We verified the security of mutual authentication and the resulting session key using the BAN logic, RoR model, and AVISPA. Additionally, we assessed the protocol's resistance to insider attacks, impersonation attacks, replay attacks, man-in-the-middle attacks, privileged insider attacks, stolen smartcard attacks, DoS attacks, and session key disclosure attacks through informal analysis, confirming the effectiveness of mutual authentication. Comparative analysis with other protocols demonstrated that the proposed protocol offers reasonable security features, communication cost, and computational cost.

In conclusion, the proposed scheme offers strong security against quantum-computer-based attacks while maintaining computational efficiency, making it suitable for securely providing real-time services to users. Future work will focus on applying the proposed three-party AKE protocol to specific environments and developing a comprehensive security system that remains secure against quantum-computer-based attacks.

Author Contributions: Methodology, Y.P. (Youngho Park) and Y.P. (Yohan Park); validation, Y.P. (Yohan Park); formal analysis, S.S.; investigation, H.P. and Y.P. (Yohan Park); resources, H.P.; data curation, S.S.; writing—original draft preparation, H.P.; writing—review and editing, Y.P. (Yohan Park); supervision, Y.P. (Youngho Park) and Y.P. (Yohan Park); project administration, Y.P. (Yohan Park); funding acquisition, Y.P. (Yohan Park). All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Bisa Research Grant (No. 20230649) of Keimyung University in 2023.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

MDPI	Multidisciplinary Digital Publishing Institute
DOAJ	Directory of open access journals
TLA	Three-letter acronym
LD	Linear dichroism

References

1. Yan, S.Y. Primality testing and integer factorization in public-key cryptography. In *Advances in Information Security* ; Springer: Berlin/Heidelberg, Germany, 2009.
2. Adleman, L. A subexponential algorithm for the discrete logarithm problem with applications to cryptography. In Proceedings of the 20th Annual Symposium on Foundations of Computer Science (SFCS 1979), San Juan, PR, USA, 29–31 October 1979; IEEE Computer Society: Washington, DC, USA, 1979; pp. 55–60.
3. Shor, P.W. Algorithms for quantum computation: Discrete logarithms and factoring. In Proceedings of the 35th Annual Symposium on Foundations of Computer Science, Santa Fe, NM, USA, 20–22 November 1994; IEEE: Piscataway, NJ, USA, 1994; pp. 124–134.
4. Grover, L.K. A fast quantum mechanical algorithm for database search. In Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing, Philadelphia, PA, USA, 22–24 May 1996; pp. 212–219.
5. Ikematsu, Y.; Nakamura, S.; Takagi, T. Recent progress in the security evaluation of multivariate public-key cryptography. *IET Inf. Secur.* **2023**, *17*, 210–226. [CrossRef]
6. Basu, K.; Soni, D.; Nabeel, M.; Karri, R. Nist post-quantum cryptography—A hardware evaluation study. *Cryptol. ePrint Arch.* **2019**, *2019*, 47.
7. Baldi, M.; Bodrato, M.; Chiaraluce, F. A new analysis of the McEliece cryptosystem based on QC-LDPC codes. In Proceedings of the Security and Cryptography for Networks: 6th International Conference, SCN 2008, Amalfi, Italy, 10–12 September 2008; Proceedings 6; Springer: Berlin/Heidelberg, Germany, 2008; pp. 246–262.
8. Dam, D.T.; Tran, T.H.; Hoang, V.P.; Pham, C.K.; Hoang, T.T. A survey of post-quantum cryptography: Start of a new race. *Cryptography* **2023**, *7*, 40. [CrossRef]
9. Shin, K.; Cho, H. Post-Quantum Cryptography and Research Trend Analysis. In Proceedings of the 2023 Korean Institute of Information Scientists and Engineers Conference, Jeju, Republic of Korea, 17–20 June 2023; pp. 1244–1246.
10. Ajtai, M. Generating hard instances of lattice problems. In Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing, Philadelphia, PA, USA, 22–24 May 1996; pp. 99–108.
11. Regev, O. On lattices, learning with errors, random linear codes, and cryptography. *J. ACM (JACM)* **2009**, *56*, 1–40. [CrossRef]
12. Lyubashevsky, V.; Peikert, C.; Regev, O. On ideal lattices and learning with errors over rings. In Proceedings of the Advances in Cryptology–EUROCRYPT 2010: 29th Annual International Conference on the Theory and Applications of Cryptographic Techniques, French Riviera, France, 30 May–3 June 2010; Proceedings 29; Springer: Berlin/Heidelberg, Germany, 2010; pp. 1–23.
13. Langlois, A.; Stehlé, D. Worst-case to average-case reductions for module lattices. *Des. Codes Cryptogr.* **2015**, *75*, 565–599. [CrossRef]
14. Guo, S.; Song, Y.; Guo, S.; Yang, Y.; Song, S. Three-Party Password Authentication and Key Exchange Protocol Based on MLWE. *Symmetry* **2023**, *15*, 1750. [CrossRef]
15. Ding, J.; Xie, X.; Lin, X. A simple provably secure key exchange scheme based on the learning with errors problem. *Cryptol. ePrint Arch.* **2012**, *2012*, 1–15.
16. Peikert, C. Lattice cryptography for the internet. In Proceedings of the International Workshop on Post-Quantum Cryptography, Waterloo, ON, Canada, 1–3 October 2014; Springer: Berlin/Heidelberg, Germany, 2014; pp. 197–219.
17. Wang, Q.; Wang, D.; Cheng, C.; He, D. Quantum2FA: Efficient quantum-resistant two-factor authentication scheme for mobile devices. *IEEE Trans. Dependable Secur. Comput.* **2021**, *20*, 193–208. [CrossRef]
18. Dharminder, D.; Reddy, C.B.; Das, A.K.; Park, Y.; Jamal, S.S. Post-quantum lattice-based secure reconciliation enabled key agreement protocol for IoT. *IEEE Internet Things J.* **2022**, *10*, 2680–2692. [CrossRef]

19. Kumar, U.; Garg, M.; Kumari, S.; Dharminder, D. A construction of post quantum secure and signal leakage resistant authenticated key agreement protocol for mobile communication. *Trans. Emerg. Telecommun. Technol.* **2023**, *34*, e4660. [CrossRef]
20. Chaudhary, D.; Dadsena, P.K.; Padmavathi, A.; Hassan, M.M.; Alkhamees, B.F.; Kumar, U. Anonymous Quantum Safe Construction of Three Party Authentication and Key Agreement Protocol for Mobile Devices. *IEEE Access* **2024**, 12, 74572–74585. [CrossRef]
21. Chaudhary, D.; Kumar, U.; Saleem, K. A construction of three party post quantum secure authenticated key exchange using ring learning with errors and ecc cryptography. *IEEE Access* **2023**, *11*, 136947–136957. [CrossRef]
22. Rewal, P.; Singh, M.; Mishra, D.; Pursharthi, K.; Mishra, A. Quantum-safe three-party lattice based authenticated key agreement protocol for mobile devices. *J. Inf. Secur. Appl.* **2023**, *75*, 103505. [CrossRef]
23. Jin, A.T.B.; Ling, D.N.C.; Goh, A. Biohashing: Two factor authentication featuring fingerprint data and tokenised random number. *Pattern Recognit.* **2004**, *37*, 2245–2255. [CrossRef]
24. Dolev, D.; Yao, A. On the security of public key protocols. *IEEE Trans. Inf. Theory* **1983**, *29*, 198–208. [CrossRef]
25. Son, S.; Lee, J.; Park, Y.; Park, Y.; Das, A.K. Design of blockchain-based lightweight V2I handover authentication protocol for VANET. *IEEE Trans. Netw. Sci. Eng.* **2022**, *9*, 1346–1358. [CrossRef]
26. Ju, S.; Park, Y. Provably Secure Lightweight Mutual Authentication and Key Agreement Scheme for Cloud-Based IoT Environments. *Sensors* **2023**, *23*, 9766. [CrossRef]
27. Burrows, M.; Abadi, M.; Needham, R. A logic of authentication. *ACM Trans. Comput. Syst. (TOCS)* **1990**, *8*, 18–36. [CrossRef]
28. Lee, S.; Kim, S.; Yu, S.; Jho, N.; Park, Y. Provably secure PUF-based lightweight mutual authentication scheme for wireless body area networks. *Electronics* **2022**, *11*, 3868. [CrossRef]
29. Park, Y.; Ryu, D.; Kwon, D.; Park, Y. Provably secure mutual authentication and key agreement scheme using PUF in internet of drones deployments. *Sensors* **2023**, *23*, 2034. [CrossRef]
30. Abdalla, M.; Fouque, P.A.; Pointcheval, D. Password-based authenticated key exchange in the three-party setting. In Proceedings of the Public Key Cryptography-PKC 2005: 8th International Workshop on Theory and Practice in Public Key Cryptography, Les Diablerets, Switzerland, 23–26 January 2005; Proceedings 8; Springer: Berlin/Heidelberg, Germany, 2005; pp. 65–84.
31. Armando, A.; Basin, D.; Boichut, Y.; Chevalier, Y.; Compagna, L.; Cuéllar, J.; Drielsma, P.H.; Héam, P.C.; Kouchnarenko, O.; Mantovani, J.; et al. The AVISPA tool for the automated validation of internet security protocols and applications. In Proceedings of the Computer Aided Verification: 17th International Conference, CAV 2005, Edinburgh, UK, 6–10 July 2005; Proceedings 17; Springer: Berlin/Heidelberg, Germany, 2005; pp. 281–285.
32. Islam, S.H.; Basu, S. PB-3PAKA: Password-based three-party authenticated key agreement protocol for mobile devices in post-quantum environments. *J. Inf. Secur. Appl.* **2021**, *63*, 103026. [CrossRef]
33. Project, OpenSSL OpenSSL Releases. Available online: https://github.com/openssl/openssl/releases (accessed on 13 August 2024).
34. Team, NewHope Project NewHope Post-Quantum Cryptography Project. Available online: https://newhopecrypto.org/ (accessed on 13 August 2024).
35. Team, PQC Crystals PQC Crystals Project. Available online: https://pq-crystals.org/ (accessed on 13 August 2024).
36. National Institute of Standards and Technology (NIST). Post-Quantum Cryptography Standardization. Available online: https://csrc.nist.gov/pqc-standardization (accessed on 13 August 2024).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

electronics

MDPI

Article

IoT-GChain: Internet of Things-Assisted Secure and Tractable Grain Supply Chain Framework Leveraging Blockchain

Karan Singh Thakur [1,†], Rohit Ahuja [1,†] and Raman Singh [2,*]

[1] Department of Computer Science & Engineering, Thapar Institute of Engineering & Technology, Patiala 147004, India; kthakur1_be20@thapar.edu (K.S.T.); rohit.ahuja@thapar.edu (R.A.)
[2] School of Computing, Engineering and Physical Sciences, University of the West of Scotland, Glasgow G72 0LH, UK
* Correspondence: raman.singh@uws.ac.uk
† These authors contributed equally to this work.

Abstract: The grain supply chain is crucial for any nation's self-sustainability due to its huge impact on food security, economic stability, and the livelihoods of several people. The path grain takes from farmers to consumers is opaque and complicated, due to which consumers cannot trust grain quality and its origin. Although blockchain is widely used for fair and secure transactions between farmers and buyers, issues related to transparency and traceability in the grain supply chain, such as counterfeiting and middlemen involvement, have not been adequately addressed. To tackle these issues, a blockchain-based solution is proposed that unites farmers, warehouses, government central and state agencies, transporters, and food corporations on a single platform to enhance transparency, traceability, and trust among all parties. This system involves minting a non-fungible token (NFT) corresponding to each lot of grain approved by government officials. The NFT comprises grain quality, type, temperature data from sensors, weight, and ownership information, which updates as the grain lot moves across the supply chain from central agencies to state agencies and so on. NFTs enable stakeholders to track the grain lot from cultivation to end-users, providing insights into grain conditions and quality. An Internet of Things-based circuit is designed using a Digital-output relative humidity & temperature (DHT22) sensor, which offers real-time temperature and humidity readings, and geolocation coordinates are gathered from the GPS module across the supply chain. Farmers can directly interact with warehouses to sell grains, eliminating the need for middlemen and fostering trust among all parties. The proposed four-tier framework is implemented and deployed on the Ethereum network, with smart contracts interacting with React-based web pages. Analysis and results of the proposed model illustrate that it is viable, secure, and superior to the existing grain supply chain system.

Keywords: blockchain; grain supply chain; Internet of Things; non-fungible tokens (NFTs)

1. Introduction

In developing countries, the agricultural sector is a primary source of livelihood for many citizens. For sustainable agriculture, high-quality grain is essential, as it significantly influences the effectiveness of other farming inputs. Estimates suggest that the direct contribution of high-quality grain to overall production ranges from 15% to 20%, depending on the crop, and can reach up to 45% with effective management of other inputs. In recent decades, the grain sector has seen substantial advancements [1]. Government bodies have restructured the grain industry to strengthen grain infrastructure. To remain competitive and contribute to national food and nutritional security, grain corporations need to transform their infrastructure, technologies, approach, and management culture in line with industry advancements. Producers procure grain and get it validated through government entities such as food corporations, which then store this grain in warehouses.

Farmers receive this grain based on demand. The food supply chain plays a crucial role in a country's economy and the sustainability of its agricultural sector. However, the exchange of goods within this supply chain is often marred by complex, opaque settlement processes, increasing the risk of fraud and the involvement of middlemen, which raises costs and complicates access to quality grain for farmers [2].

Existing grain supply chain solutions aim to enhance transparency between producers (farmers) and consumers. However, these solutions often overlook the specific challenges related to grain distribution, like grain condition across the chain, traceability of grains, continuous tracking, and accounting [3]. Furthermore, in existing grain chains, there is an intermediary who plays maliciously for illicit profit. Thus, there is a need for a grain distribution framework that enables farmers to track grain throughout the supply chain from producers and identify its sources.

Blockchain technology aims to revolutionize the existing public distribution system using a distributed and immutable ledger. Blockchain technology, along with non-fungible tokens (NFTs), enhances traceability, transparency, and accountability throughout the grain supply chain; the proposed scheme seeks to streamline operations, reduce fraud, and ensure the delivery of safe and high-quality grains to consumers. The primary objective is to create a decentralized and transparent grain supply chain system that addresses challenges such as counterfeiting, contamination, and lack of data sharing among stakeholders. The vision is to transform the current public distribution system into a technologically advanced ecosystem that ensures the availability of genuine and safe grains to every consumer. The proposed framework suggests a grain distribution system, i.e., with key entities, as discussed below:

1. Governing authority: As the central authority overseeing food distribution, it plays a pivotal role in the grain supply chain traceability system. It orchestrates the entire process by registering each level. It ensures adherence to regulations, assigns blockchain addresses, and manages permissions to ensure transparency and accountability in grain distribution.
2. Level 1 tier: This level collaborates closely with the governing authority to ensure effective implementation of the traceability process. It provides overarching guidance and regulatory framework and, through the governing authority, ensures that stakeholders are registered and operate within established protocols.
3. Level 2 tier: Level 2 works in tandem with the governing authority to ensure region-specific grain distribution. It facilitates the registration of level 2 and level 3 tiers within its jurisdictions, enhancing localized management and coordination.
4. Level 3 tier: Level 3, under the oversight of the governing authority, manages the distribution of grains within its respective sub-area. It is responsible for registering and supervising the level 4 tier, ensuring efficient and equitable access to essential commodities.
5. Level 4 tier: This level acts as the final point of contact with consumers. It provides grains to end-users, adhering to the regulations and maintaining grain quality and authenticity. Through the traceability system, the level 4 tier upholds the integrity of distributed grains.

Contributions of this article: The contributions of our proposed framework are discussed below:

- An Internet of Things-enabled circuit is designed to capture the real-time humidity and temperature of grain across the supply chain and continuously update it on blockchain so it can be monitored by all stakeholders.
- A non-fungible token-based system is created to verify grain quality and its ownership, which enhances the transparency of grains to consumers.
- We ensure grain traceability through each level using blockchain technology and an NFT-QR-based traceable mechanism for real-time monitoring and discrepancy detection.
- The proposed framework is implemented, Remix Ide is used for building smart contracts, and EtherScan is used for transaction monitoring.

- Security and threat analysis of smart contracts is conducted using the Solidity Scan tool.

2. Existing Systems

The existing grain supply chain in agriculture relies on minimal technology and predominantly manual record-keeping, making it susceptible to malpractices such as misrepresentation of product availability and black marketing of crucial and high-demand grains. This lack of transparency and traceability hinders the efficiency and integrity of the supply chain, ultimately impacting farmers, consumers, and the overall economy, as depicted in Figure 1. This situation allows for corruption in government-controlled product pricing, sales, and purchases due to the lack of transparency and efficient record-keeping [4,5]. In India, corruption at Food Corporation of India (FCI) warehouses is frequently reported. Officials often input incorrect information into government databases. For example, if a farmer purchases 150 kg of grain, FCI officials might record it as 225 kg sold. The remaining 75 kg is then back-channelled and sold later at inflated prices, creating an expensive market [6]. This corruption increases costs and makes it difficult for poor farmers to access good-quality grain, posing a significant issue at the initial stages of the food supply chain.

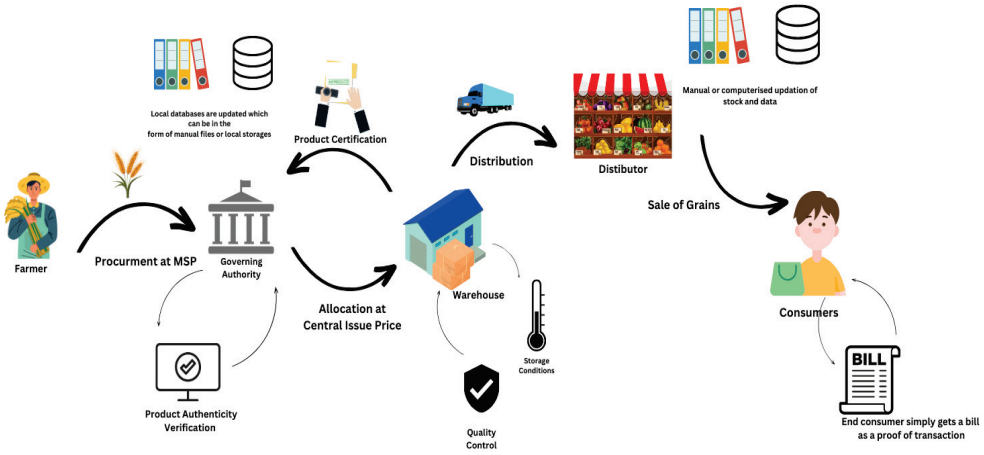


Figure 1. Existing grain distribution.

A detailed survey of the agri-food sector and a framework corresponding to the Vietnamese cashew nut business have been discussed to address such concerns [7–10]. In Iraqi agriculture, blockchain technology has been incorporated to tackle these issues by improving data management, accountability, and intelligent contracts, thus increasing productivity and competitiveness. Blockchain's efficiency and transparency make it a wise investment for Iraq's agricultural industry, with potential applications in precision agriculture, food supply chains, crop insurance, and agricultural product transactions.

A blockchain-based system called "AgriOnBlock" is designed to enhance transparency among various stakeholders, such as bankers, retailers, customers, farmers, and wholesalers [11]. It aims to handle industry-specific issues and connect stakeholders efficiently using Ethereum smart contracts and Internet of Things (IoT) devices. Additionally, a blockchain framework built on an IoT that incorporates artificial intelligence is intended to oversee and control smart water management. An IoT-based smart water irrigation system is recommended to address the water crisis efficiently, considering factors like fertilizer quality [12]. Most existing techniques focus primarily on food supply and food security, benefiting all stakeholders by increasing transparency and trust. However, grain, being a fundamental element in the farming sector, has not received sufficient attention, and its distribution suffers from the following issues:

1. Identification of beneficiaries: Governments or relevant authorities identify individuals or households eligible for subsidized food based on socio-economic indicators, income levels, or specific criteria.

2. Allocation of food resources: Governments allocate essential food resources, including grains like rice and wheat, to regions or states based on population size, socioeconomic factors, and nutritional needs. The quantity and types of allocated food resources may vary to align with regional preferences and dietary requirements.
3. Distribution points: Authorized distribution points, which may include government-operated stores, local markets, or community centers, facilitate the distribution of subsidized food items.
4. Distribution process: Beneficiaries visit designated distribution points to collect their allocated food resources. Transactions may involve manual or digital record-keeping, depending on the technological infrastructure of the region.
5. Monitoring and evaluation: Governments and relevant agencies monitor the distribution process to ensure fair and effective delivery of food resources to beneficiaries. Regular evaluations may be conducted to assess the performance of the distribution system and identify opportunities for improvement.

The existing public distribution system faces challenges such as leakages, corruption, and inefficiencies. There have been instances of grains not reaching the intended beneficiaries, leading to concerns about the system's effectiveness.

Limitations of Existing Systems

Before delving into the reasons for employing blockchain technology in grain supply chain management, it is essential to understand the limitations of previous systems. These traditional systems, which often rely on centralized databases and manual processes, are fraught with several challenges that hinder the effective management of grain distribution. Below are some generic limitations that transcend geographical boundaries and apply to grain supply chains worldwide:

1. Lack of traceability: Traditional systems often lack robust mechanisms for tracing the journey of grains from farm to fork. This absence of traceability makes it difficult to identify the origin of grains, track their movement through the supply chain, or ascertain their quality and safety status accurately.
2. Limited transparency: Centralized systems typically offer limited transparency regarding grain transactions and movement. Stakeholders may not have access to real-time information about inventory levels, pricing, or quality checks, leading to opacity in the supply chain and potential opportunities for fraud or malpractice.
3. Vulnerability to fraud and counterfeiting: Manual record-keeping and paper-based documentation systems are susceptible to errors, tampering, and fraudulent activities. Without robust authentication mechanisms, such as digital signatures or unique identifiers, there is a heightened risk of counterfeit products entering the supply chain, compromising food safety and integrity.
4. Inefficiencies in record-keeping: Paperwork-intensive processes are inherently inefficient and prone to delays, inaccuracies, and data discrepancies. Manual record-keeping not only consumes time and resources but also increases the likelihood of human errors, leading to inefficiencies in inventory management, order processing, and regulatory compliance.
5. Limited accountability and compliance: Traditional systems may lack mechanisms to hold stakeholders accountable for their actions or ensure compliance with regulatory standards and industry best practices. Without transparent audit trails or mechanisms for verifying the authenticity of data, it becomes challenging to enforce accountability or address non-compliance issues effectively.
6. Fragmented information silos: Grain supply chains often involve multiple stakeholders, including farmers, distributors, retailers, and regulatory authorities. In traditional systems, information may be siloed within individual organizations or departments, leading to fragmentation and duplication of efforts. This lack of data interoperability and collaboration impedes the seamless flow of information across the supply chain, hindering decision-making and coordination efforts.

7. Susceptibility to contamination and quality issues: Inadequate monitoring and oversight in traditional systems can exacerbate the risk of grain contamination or quality degradation during storage, handling, or transportation. Without real-time visibility into environmental conditions or quality control measures, it is challenging to prevent or mitigate contamination incidents effectively.
8. High operational costs: The manual nature of traditional systems often results in high operational costs associated with labor, paperwork, and administrative overheads. Moreover, inefficiencies in inventory management, order fulfilment, and compliance management can further inflate operational expenses, reducing profitability and competitiveness in the market.

Addressing these limitations requires a paradigm shift towards more transparent, efficient, and secure approaches to grain supply chain management. Blockchain technology offers a promising solution to overcome these challenges by providing a decentralized, transparent, and tamper-resistant platform for tracking and tracing grain transactions throughout the supply chain.

3. Preliminaries

This section discusses blockchain technology, smart contracts, and the various types of blockchain solutions, which are essential for grasping the effectiveness of the proposed scheme.

3.1. Blockchain

Traditional systems are primarily based on the client–server architecture and centralized database. Blockchain is a decentralised ledger with a peer-to-peer system of networks. It is a cryptographically secure, immutable, append-only ledger, which can replace the current system [13,14]. Blockchain comprises blocks that are append-only, and a particular block may store data such as a list of transactions, as depicted in Figure 2. Next, a block is attached to the pertaining block, forming a chain, giving it the name blockchain. Each block except the first block, called the genesis block, stores a hash of the previous and next block, which is a cryptographic code. The hash must match in order to form a link between blocks. If anything in the block is changed, the hash is updated to a new value, thus breaking the link and the chain. The blockchain is a decentralized network and works on a mutual consensus paradigm to make any change. The consensus algorithms verify and authenticate the transactions so the public ledgers remain in harmony and ensure only valid transactions are further added to the blockchain. Proof of work [15,16] is the most widespread consensus algorithm today, which requires the exertion of the computational power of a node to solve an arbitrary puzzle to validate transactions.

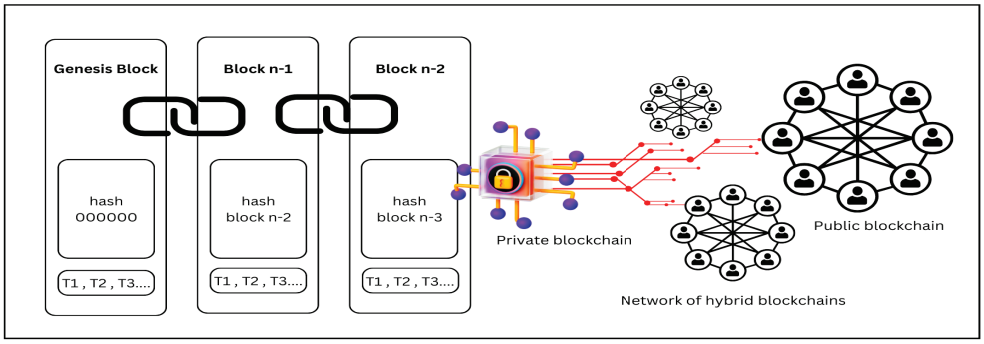


Figure 2. Blockchain.

The data stored in a blockchain are considered practically unalterable because if a single block needs to be altered, the whole network needs to consent to the change, and that would require computation power greater than the whole network combined. This renders the human intervention aspect of tampering entries moot and makes blockchain

a plus point over the centralized system where a hack may be possible and thus lead to tampered data [17].

3.2. Smart Contracts

A smart contract is a programme/script that is triggered to execute on a blockchain whenever a specified condition is met, as depicted in Figure 3. It uses blockchain technology to verify, validate, capture, and enforce agreed-upon terms between multiple parties. In addition, it allows transactions and agreements to be carried out among anonymous parties without the need for a central entity, external enforcement, or legal system. Smart contracts define rules as well as enforce them automatically [13]. The user interacts with the contract, and each interaction is a transaction, which incurs some transactional cost called gas fees to record the transaction on the blockchain. This enables the secure, instantaneous, and efficient transfer of land titles. This also ensures trust that the ownership is immune to external factors and is conclusive in nature.

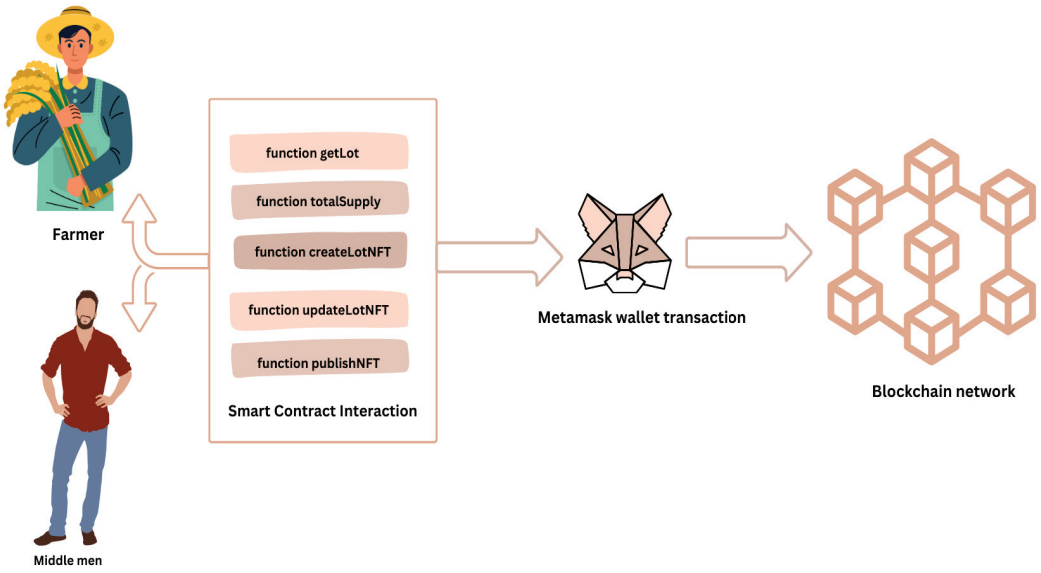


Figure 3. Smart contract.

3.3. Types of Blockchain-Based Solutions

A decentralized app that uses technologies like blockchain and smart contracts can be a viable solution for a new land registration system [14]. There are four primary types of blockchain solutions that can be employed:

1. Public blockchain: A public blockchain is an open network that allows anyone to set up a node for the blockchain. There is no restriction on joining the blockchain, and every transaction is visible on the ledger (the users are anonymous), thus making it transparent. It is decentralized and easy to set up [18]. An Ethereum blockchain is a commonly used public blockchain. Currently, the Ethereum blockchain is the most prevalent and established smart contract provider. It uses the Ether cryptocurrency *ETH* and works on the proof of stake consensus mechanism. Every time a transaction occurs, a reward (in terms of Ether) is given to the validator of the transaction. This provides an incentive for people to set up nodes and has greatly increased outreach. Public blockchains are easy to set up and open for all, due to which they are difficult to regulate and often slower.
2. Private blockchain: A private blockchain can be considered a blockchain network that is owned by a particular entity or authority and is directly managed by it. It can be considered a special type of private networking solution that is decentralized [14]. A private blockchain network requires an invitation and must be validated by either the network starter or by a set of rules by the network starter. It is not fully transparent,

as only the nodes on the network may view the transactions, and it is not open to everyone. The managing authority may perform relevant changes, albeit not that easily. This can be considered a partially decentralized system.
Several companies offer private blockchain solutions to make businesses more secure. Private blockchains usually require low costs to set up and are more efficient due to less network congestion. A private blockchain is easier to manage but may not necessarily be immutable, as the whole network may be updated, which may be possible depending on the size of the network and is an inescapable drawback to private blockchains. The issuing authority has to ensure trust; thus, it makes it similar to traditional centralized systems but with distributed nodes acting as servers.

3. Hybrid blockchain: This blockchain aims to combine both public and private blockchains by offering a solution that is open for all but reserves some permissions and activities for a selected group of nodes in the network. This is made possible by using a permission-based system within the permission-less system of a public blockchain. This may largely help to reduce fraud, as only verified users may be allowed to use the blockchain. This makes it a permissioned blockchain, which is also open to the public, ensuring better decentralization of the system.
4. Consortium blockchain: Consortium and hybrid blockchains are similar as they both have a public and private blockchain component. The difference lies in the structure of the organization's management of the permissioned part. For a consortium blockchain, multiple entities govern the platform. The entities also define the roles of nodes and the permissions granted to each.

4. The Proposed Framework

The proposed framework integrates various steps, utilizing blockchain technology, Firebase, and additional components to establish transparency and traceability within the public distribution system for grains [19–21]. The framework involves the collection of real-time sensor data, including temperature, humidity, and geolocation coordinates, which are stored in the Firebase real-time database for immediate accessibility. The Firebase integration acts as a dynamic link between sensor inputs and the web application, ensuring instantaneous reflection of any changes. The blockchain interaction encompasses the creation of non-fungible tokens (NFTs) for grain lots and updating and transferring ownership through transactions on the Sepolia blockchain. IPFS integration further contributes to the decentralized and tamper-resistant storage of NFT metadata. The web application serves as the user interface, featuring a dashboard with a QR code scanner, status log, location map, and quality log for stakeholders to retrieve detailed information about grain lots and enhance traceability. Quality verification involves a check on the grain's temperature against defined thresholds, with results displayed in the quality log. User interaction is facilitated through MetaMask integration, offering features like the connect button, notification icon, dark mode, and profile icon for a seamless and secure experience. Continuous monitoring ensures system health, detects anomalies, and plans regular updates based on user feedback, technological advancements, and industry standards. The procedural workflow, depicted in Figure 4, underscores the seamless integration of sensor data, blockchain technology, IPFS, and Firebase, contributing to the proposed scheme's overarching goals of improving traceability [22], accountability, and real-time monitoring in the grain supply chain.

The implementation of non-fungible tokens (NFTs) within the proposed scheme is designed to facilitate comprehensive traceability of grains, offering intricate insights into the origin, transportation, and handling of each grain lot. The decentralized nature of the blockchain technology employed ensures a transparent environment, granting stakeholders real-time access to critical information regarding the state, location, and condition of both grain lots and orders. The permanence of data recorded on the blockchain guarantees integrity, as it cannot be altered or deleted, thereby ensuring the accuracy and reliability of information across the entire supply chain. Leveraging NFTs and smart contracts, the

proposed scheme establishes a robust system of accountability by meticulously recording interactions and actions taken by each stakeholder. This not only encourages responsible behavior but also significantly reduces the potential for fraudulent activities. The incorporation of smart contracts introduces automation to streamline processes, eliminating manual paperwork and reducing administrative overhead, ultimately leading to enhanced operational efficiency. As a result, the optimized processes, coupled with the reduction in fraudulent activities, are anticipated to contribute substantially to cost savings within the grain supply chain.

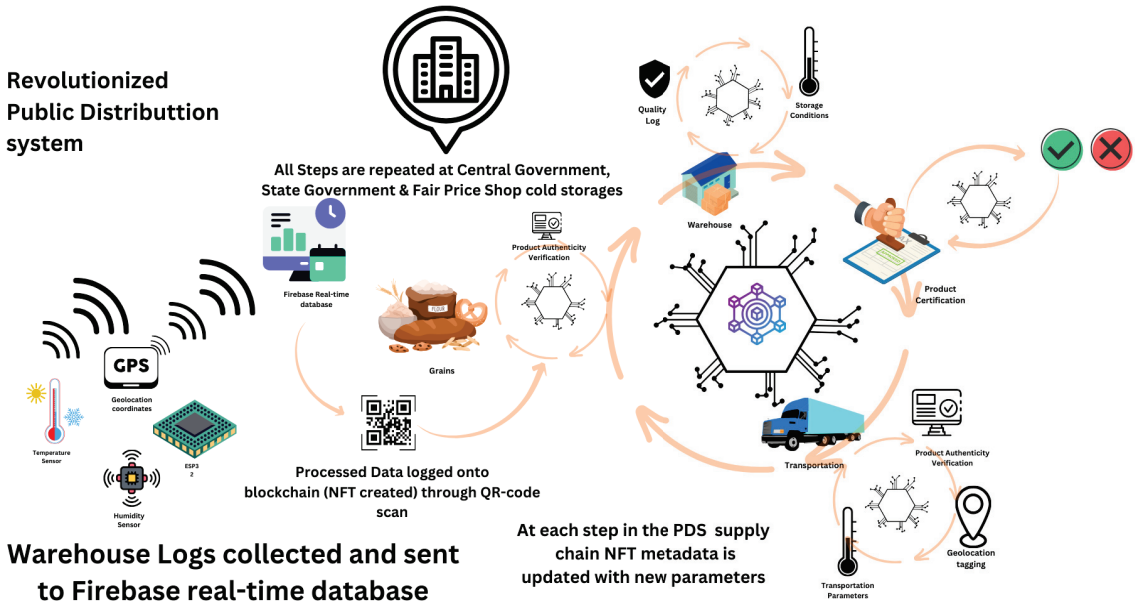


Figure 4. Proposed framework.

Farmer registration and verification: To register and verify a farmer based on their public Ethereum address, Algorithm 1 is employed.

Algorithm 1 Farmer registration and verification

Input: Farmer's Ethereum address F_{Addr}.
Output: Farmer successfully registered and verified.

$F_{Addr} \xleftarrow{Input} (Farmer)$
if $(\exists F_{Addr} == 0)$ **then**
  $F_{Details} \xleftarrow{Extract} getDetailsFromBlockchain(F_{Addr})$
  $F_{Verified} \xleftarrow{Verify} verifyDetails(F_{Details})$
  if $(F_{Verified} == 1)$ **then**
    $F_{Reg} \leftarrow registerFarmer(F_{Details})$
    if $(F_{Reg} == 1)$ **then**
      return Farmer successfully registered and verified
    else
      return Farmer registration failed
    end if
  else
    return Farmer details verification failed
  end if
else
  return Farmer already registered
end if

Mint NFT for grain sack: To mint an NFT corresponding to a grain sack, Algorithm 2 is employed. The grain type is taken as input, i.e., G_{type} and G_Q grain quality, the weight of the grain sack G_{w_S}, its location $G_{S_loc} : (G_{S_latitute}, G_{S_longitude})$, the grain sack temperature, and its humidity G_{S_teamp} and G_{S_humid}.

Algorithm 2 Mint NFT of grain sack G_{Sack}

Input: G_{type}, G_Q, G_{w_S}, G_{S_loc}, G_{S_temp}, G_{S_humid}
Output: NFT $\mathbb{N}$ for G_{Sack}

$FCI \xleftarrow[Submitted]{G_{type}, G_Q, G_{w_S}, G_{S_loc}, G_{S_temp}, G_{S_humid}} G_{Sack}$
if $(G_{Sack}(Source) == 1)$ **then**
  $G_{Sack} \cdot append(Id) \leftarrow tokenIdCounter$
  $G_{Sack} \cdot append(Manuf) \leftarrow msg.sender$
  $G_{Sack} \cdot append(Ts) \leftarrow timestamp$
  $G_{Sack} \cdot append(state) \xleftarrow{Input} State : State \in \{Central, District, Retail\}$
  $\mathbb{N} \xleftarrow[mint()]{G_{type}, G_Q, G_{w_S}, G_{S_loc}, G_{S_temp}, G_{S_humid}, G_{Sack.Id}, G_{Sack.Manu}, G_{Sack.Ts}, G_{Sack.state}} G_{Sack}$
  $\mathbb{N}$ created successfully
  $tokenCounter++$
else $\mathbb{N}$ creation aborted
end if

Grain sack NFT update: To update G_{Sack} NFT, Algorithm 3 is employed. This algorithm updates the information associated with G_{Sack}, including the grain description $G^{new}_{S_loc}$, $G^{new}_{W_S}$, $G^{new}_{S_temp}$, $G^{new}_{S_humid}$, and grain quality $G^{new}_{S_Q}$.

Algorithm 3 Grain sack NFT update

Input: $G_{Sack.Id}$, $G^{new}_{S_Q}$, $G^{new}_{S_loc}$, $G^{new}_{S_temp}$, $G^{new}_{S_humid}$, $G^{new}_{S_Q}$
Output: Returns true if update is successful.

$G_{Sack}[] \leftarrow G_{Sack.Id} : \forall Id \in \{1, ..., n\}$
for i $\in G_{Sack}[]$ **do**
  $W_{prev} \leftarrow G_{Sack}[i].weight$
  $G_Q \leftarrow G_{Sack}[i].Quality$
  $G_{Sack}\left(\text{push}\left(G^{new}_{S_loc}, G^{new}_{W_S}, G^{new}_{S_temp}, G^{new}_{S_humid}, G^{new}_{S_Q}\right)\right)$
  $\triangle \text{T} = \left(G^{new}_{S_temp} - G_{S_temp}[i]\right)$
  if $\left(\triangle T <= \delta\right)$ **then** ▷ $\triangle$T is change in temperature, δ is threshold
    if $\left(W_{prev} == G^{new}_{W_S}[i]\right)$ **then**
      $\mathbb{N}$ updated successfully
    end if
  end if
end for
Generate signal for G_{Sack} tampering

Buying grains: Algorithm 4 is employed to buy grains. This algorithm handles updating all necessary data structures and ensures the proper handling of transactions, as depicted in Figure 5.

Algorithm 4 Buying grains

Input: Grain Sack Id $G_{Sack.Id}$ and buyer address B_{Addr}.
Output: Grain purchased successfully.

$G_{Sack}[] \leftarrow G_{Sack.Id} : \forall Id \in \{1, \ldots, p\}$
for i $\in G_{Sack}[]$ **do**
  if $(\exists\ G_{Sack}[i])$ **then**
    if (Authenticate(Buyer)==1) **then** ▷ Change State and Ownership Status
      $G_{Sack}\Big(\text{push}\Big(\ G_{Sack_Id}\ \cdot\ (State = Sold), G_{Sack_Id} \cdot (Ownership = B_{Addr})\Big)\Big)$
      Transaction T_{Id} ▷ Update Weight Changes
      Emit event for a successful purchase
      $G_{Sack_Id}\Big(\text{push}\Big(G_{W_S_Id} = G_{W_S_Id} - G_{Sack_Id} \cdot (\text{W}(\text{B}_{Addr}))\Big)\Big)$
    else
      Authentication fails
    end if
  else
    $G_{Sack}[i]$ is sold
  end if
  Required grain does not exist
end for

The proposed system is depicted in Figure 4, and the flow between all the discussed algorithms is depicted in Figure 6.

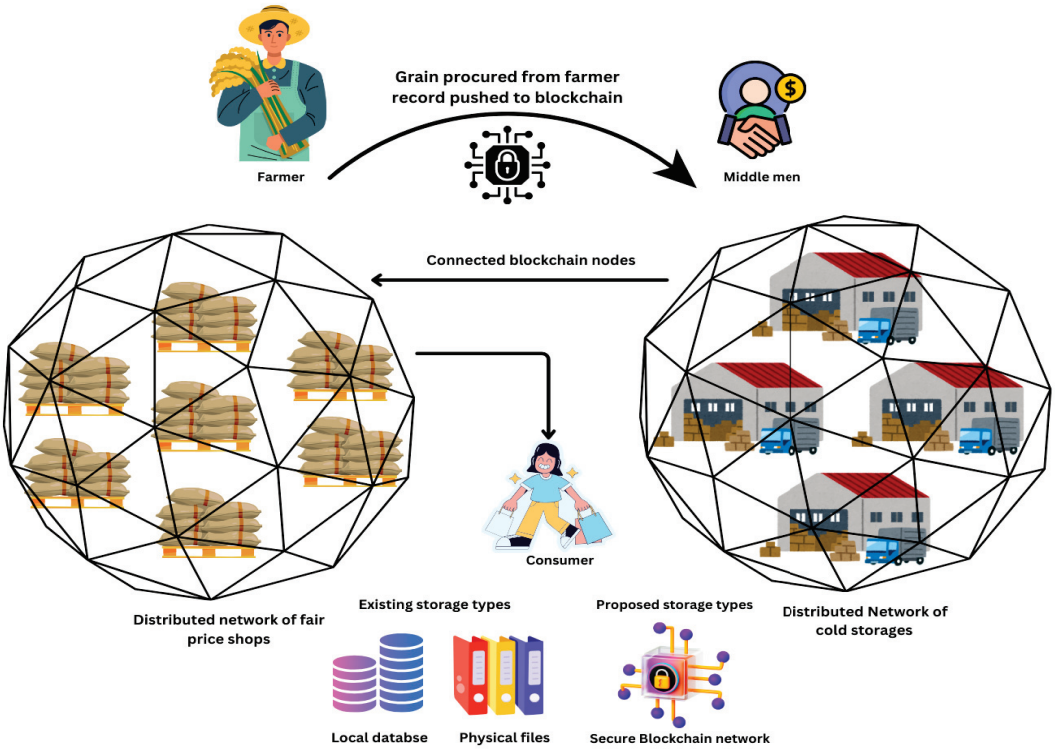


Figure 5. Blockchain nodes interaction.

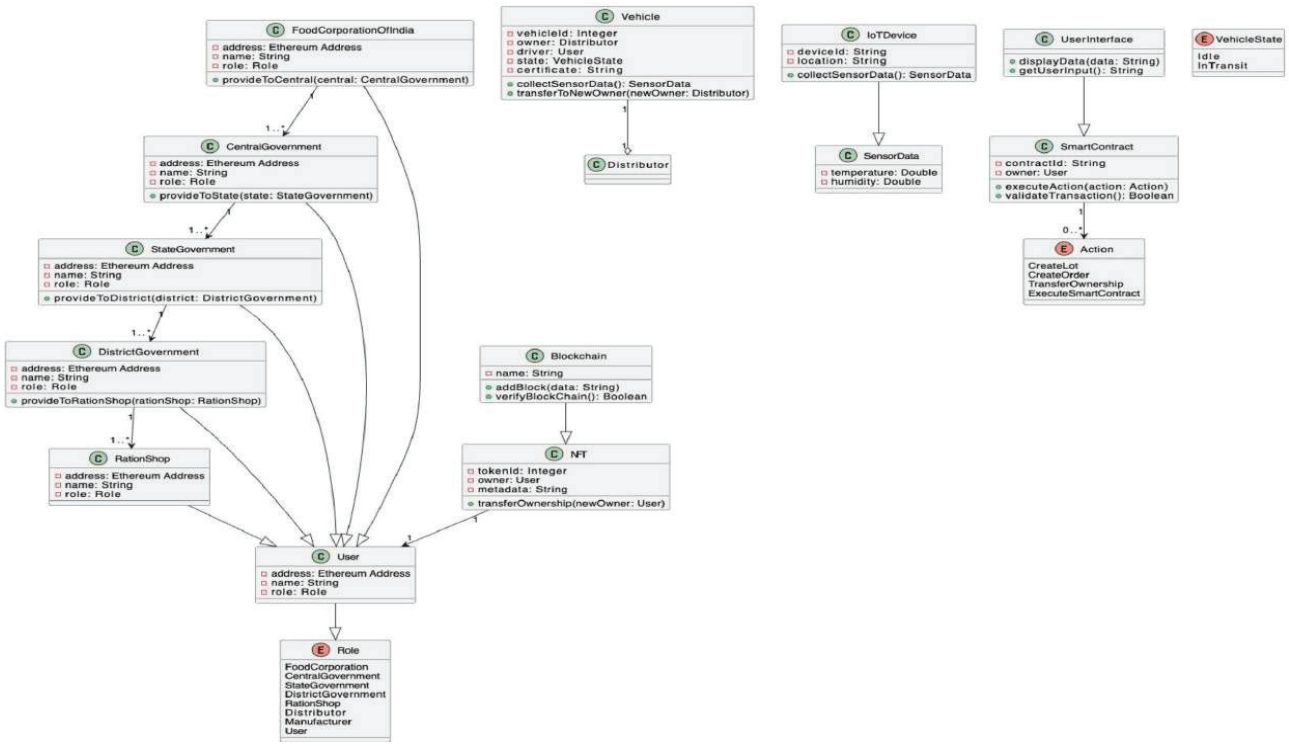


Figure 6. Algorithm flow.

5. Implementation

This section discusses the experimental setup, architecture, and smart contract design employed in the proposed scheme, as well as the integration and feature analysis utilized.

5.1. Experimental Setup

This subsection describes the tools and technology utilized in the implementation of our proposed architecture.

5.1.1. InterPlanetary File System (IPFS)

IPFS will be used to securely store sensitive data off-chain. It provides decentralized, distributed, and verifiable storage, with only the hash values stored on the blockchain.

5.1.2. Remix IDE

The development and initial testing of the smart contract were conducted using Remix IDE, a powerful web-based integrated development environment tailored toward Solidity smart contracts. Remix IDE offers a range of features that streamline the process of writing, debugging, and deploying smart contracts.

5.1.3. EtherScan for Transaction Monitoring

EtherScan is employed to monitor and verify transactions of smart contract "GrainChain". EtherScan is a widely used block explorer for the Ethereum blockchain, providing detailed information on transactions, blocks, addresses, and smart contracts. Monitoring through EtherScan comprises of the following steps:

1. Contract deployment: After deploying the 'GrainChain' smart contract to the Ethereum testnet, the contract address is obtained and entered into EtherScan.
2. Transaction tracking: EtherScan is used to monitor all transactions related to the smart contract, including NFT creation, state updates, and event emissions.

3. Verification: The details of each transaction are verified on EtherScan to ensure that the contract functions behaves as expected and that state changes are accurately recorded on the blockchain.
4. Event logs: EtherScan provides access to event logs, which are crucial for tracking the 'NFTCreated' event and other significant state changes within the contract.

5.1.4. SolidityScan

SolidityScan is employed to ensure the security and reliability of the 'GrainChain' smart contract, which is a cloud-based smart contract vulnerability scanner [23]. SolidityScan is designed to uncover potential vulnerabilities and anti-patterns in smart contract code. It integrates seamlessly into the development pipeline, scanning the contract's code and flagging known vulnerabilities. SolidityScan generates a detailed audit report that highlights which parts of the code might pose security risks. The process of SolidityScan is defined below:

1. Code upload: The 'GrainChain' smart contract code is uploaded to SolidityScan's cloud-based platform.
2. Automated scanning: SolidityScan performs a comprehensive scan of the code, using both automated detection methods and predefined rules for identifying vulnerabilities.
3. Report generation: The scanner generates a detailed audit report outlining the identified issues, their severities, and recommendations for remediation.

Here, the SolidityScan tool provided a security score of 84.56 and a threat score of 87.50, based on the lines of code and the severity of detected issues.

5.2. Architecture of the Proposed Scheme

Figure 7 depicts the architecture of the proposed system. A blockchain-based grain chain management solution can be designed using a layered architecture that incorporates various components to ensure transparency, security, and efficiency [24]. The following is an overview of the architecture for a blockchain-based grain chain solution:

1. Data collection: Information on grain transactions can be held at this state, i.e., warehouse logs, a description of the grain, data from its sensors, temperature and humidity from the DHT22 sensor, geolocation coordinates from the GPS module, grain lot, among others, as depicted in Figure 8. This phase uses distributed ledger technology to safeguard information in a decentralized manner, thus ensuring transparency and security.
2. Smart contract: Next, intelligent contracts are used to facilitate the various transactions in the grain chain. Smart contracts refer to self-executing contracts in which the agreement terms between two parties are directly written into the software. For example, in the grain chain, smart contracts are used to automate the transfer of ownership of a grain lot from the level 1 to level 2 tier.
3. Consensus mechanism: Afterwards, a consensus mechanism between the nodes of the network is employed. Consensus is vital to ensure all network nodes agree on the state of the blockchain. This layer might be materialized with a consensus algorithm, such as proof-of-work or proof-of-stake.
4. User interaction: The Farmer/retailer/agency interacts with the blockchain environment at this stage, along with the user interface. For instance, the applications could be a create lot portal, a transfer lot portal, an update lot portal, and a sell portal. It uses RESTful APIs to interact with the blockchain.

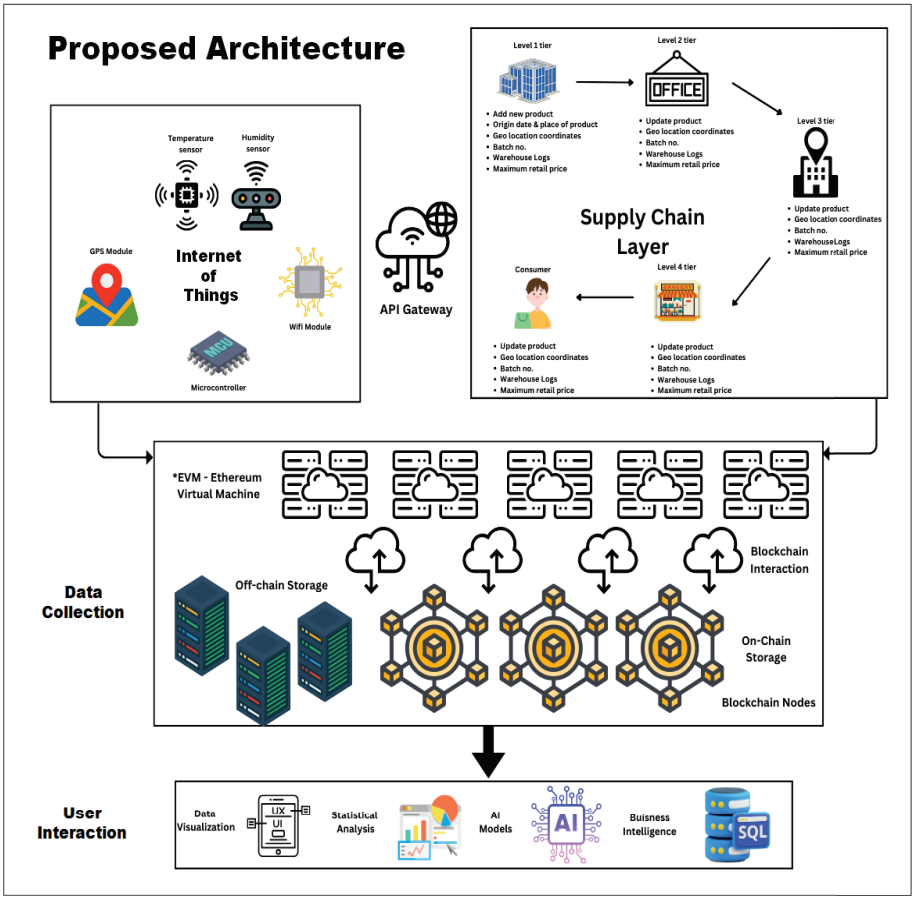


Figure 7. Proposed architecture.

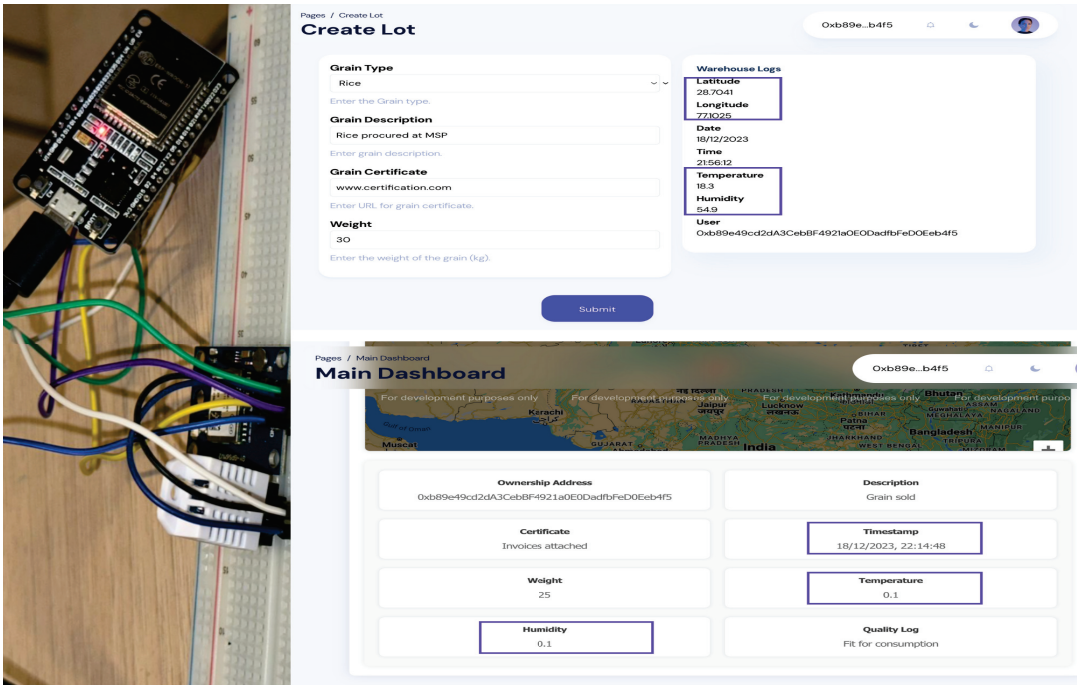


Figure 8. Temperature monitoring using sensors.

5.3. Designing the Smart Contract

The smart contract provided outlines the functionality of a decentralized grain supply chain tracking system using blockchain and non-fungible tokens (NFTs). Let us delve into the design aspects of this contract:

1. Inheritance and Imports
   - The contract is inherited from ERC-721 URIStorage, which is part of the OpenZeppelin library for ERC-721 token standards [18]
   - It also implements the ERC-721 Supply interface, indicating that it adheres to the ERC-721 token standard with additional supply-related functions.
2. State Variables
   - 'tokenIdCounter' tracks the unique identifier for each NFT token.
   - 'allLots' uses mapping to store information about each grain lot, indexed by their token IDs.
   - 'ownerOfToken' uses mapping to track the owner of each token.
3. Structs
   - Locate represents the geographical location with latitude and longitude coordinates.
   - State defines the state of a grain lot, including ownership, description, timestamp, weight, certificate URL, temperature, humidity, location, and exceeded temperature.
4. Constructor
   - The initializes the contract with the name "GrainNFT" and symbol "GNFT" upon deployment.
5. Functions:
   - 'getLot' retrieves information about a specific grain lot based on its token ID.
   - 'totalSupply' returns the total number of NFT tokens minted.
   - 'createLotNFT' creates a new grain lot NFT, mints it to the caller, and stores relevant information about the lot.
   - 'updateLotNFT' updates information about an existing grain lot based on its token ID.
   - 'publishNFT' sets the token URI for a specific NFT, essentially linking it to off-chain metadata such as descriptions or images.
6. Helper Function
   - 'uintToString' converts a uint256 value to its string representation.
7. Design Considerations
   - Efficiency: The contract design appears efficient, with minimal gas usage for common operations like creating and updating grain lots.
   - Security: The contract should be audited for potential vulnerabilities, especially regarding access control and input validation to prevent unauthorized operations or data manipulation.
   - Scalability: While the current design can handle a moderate number of grain lots, considerations for scaling, such as gas optimization and data storage efficiency, should be taken into account for large-scale adoption.
   - Interoperability: The contract should conform to relevant standards like ERC721 to ensure compatibility with existing NFT marketplaces and infrastructure.
   - User Experience: User interaction with the contract should be intuitive and well-documented to facilitate adoption by stakeholders in the grain supply chain.

By addressing these design considerations, the smart contract can serve as a robust foundation for implementing a decentralized grain supply chain tracking system, leveraging blockchain and NFTs to enhance transparency, traceability, and efficiency in grain distribution.

5.4. Integration

Integrating the smart contract with a MetaMask wallet and deploying it on the Ethereum Sepolia test chain enables users to interact with the grain supply chain tracking system in a decentralized manner. Here is how the integration process would typically unfold, along with minting procedures:

1. Setting up MetaMask
    - Users install the MetaMask browser extension or mobile app and create or import their Ethereum wallet.
    - They connect MetaMask to the Sepolia test network to interact with smart contracts deployed on this test chain.
2. Deploying the Smart Contract
    - The developer compiles the smart contract code using a Solidity compiler like Remix or Truffle.
    - Using MetaMask version 11.0.0 the developer selects the Sepolia test network and deploys the compiled contract to this network.
    - Upon deployment, MetaMask prompts the developer to confirm the transaction and pay the gas fee using test Ether (ETH) from their MetaMask wallet.
3. Interacting with the Contract
    - After deployment, users can interact with the smart contract through a decentralized application (dApp) interface connected to MetaMask.
    - They can view existing grain lots, create new grain lots, update information about existing lots, and perform other permitted actions defined by the contract's functions.
4. Minting Procedures
    - To mint a new grain lot NFT, users initiate the minting process through the dApp interface.
    - They specify details such as grain type, description, weight, certificate URL, location coordinates, temperature, humidity, and any other relevant information required by the contract.
    - Upon confirmation, MetaMask prompts the user to sign the transaction and pay the gas fee for minting the NFT.
    - Once the transaction is confirmed and included in a block on the Sepolia test chain, the new grain lot NFT is minted, and its details are recorded on the blockchain.
5. Transaction Confirmation
    - Users can monitor the status of their transactions on MetaMask, which provides real-time updates on transaction confirmations, gas fees, and network activity.
    - Once the transaction is confirmed, users can view the newly minted grain lot NFT in their MetaMask wallet or explore its details through the dApp interface.
6. Testing and Debugging
    - Developers and users can test the functionality of the smart contract on the Sepolia test chain without incurring any actual Ether transactions or gas fees.
    - They can simulate various scenarios, such as creating multiple grain lots, updating information, and transferring ownership, to ensure the contract behaves as expected.

By integrating the smart contract with the MetaMask wallet and deploying it on the Ethereum Sepolia test chain, users can seamlessly interact with the grain supply chain tracking system, mint new grain lot NFTs, and contribute to the development and testing of decentralized applications for enhancing transparency and traceability in grain distribution. Figure 9 depicts the gas fees used for our implementation.

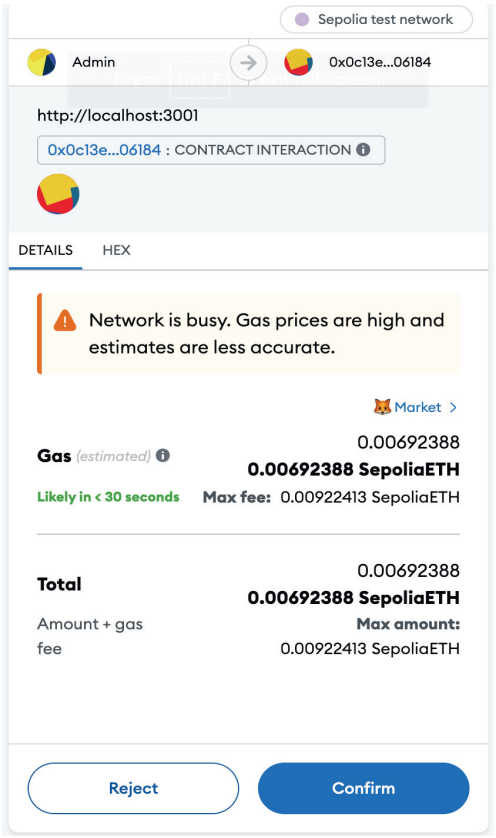


Figure 9. Example of gas fees paid.

5.5. Feature Analysis

1. Decentralization: We have created a system that significantly enhances transparency and decentralization compared to conventional practices. This decentralization not only provides clarity for end-users but also streamlines processing tasks during the product's production life cycle.
2. Reliable metadata handling: The proposed scheme has effectively published NFT metadata on IPFS and retrieved grain status information, showcasing its robust integration into the InterPlanetary File System (IPFS). The system's ability to store and retrieve decentralized and tamper-resistant metadata on IPFS has been validated.
3. Accurate real-time data handling: The logging and real-time exchange of sensor data with Firebase have been thoroughly tested and proven accurate. This ensures that the system precisely captures sensor inputs and updates the Firebase real-time database, establishing a reliable mechanism for data collection. This is critical for maintaining transparency and traceability in the grain distribution process.
4. Efficient grain status retrieval: The get lot function, designed to retrieve grain status information based on the grain ID, has successfully returned the expected grain info object. This functionality demonstrates the system's precision in retrieving and presenting detailed information about specific grain lots, contributing significantly to transparency and traceability.
5. User-friendly transaction facilitation: The generation of QR-CODEs during lot NFT creation has passed testing, highlighting the user-friendly nature of the system. The successful generation of QR codes is crucial for facilitating seamless transactions and interactions within the platform, ultimately enhancing the overall user experience.

5.6. Security and Threat Analysis

The "GrainChain" smart contract is designed to create and manage NFTs representing lots of grain. Each lot tracks various states, including ownership, description, timestamp,

weight, certificate, temperature, humidity, location, and exceeded temperature. This subsection presents an in-depth analysis of the security aspects of the contract, highlighting identified issues, explaining the security and threat scores, and recommending improvements as depicted in Figure 10.

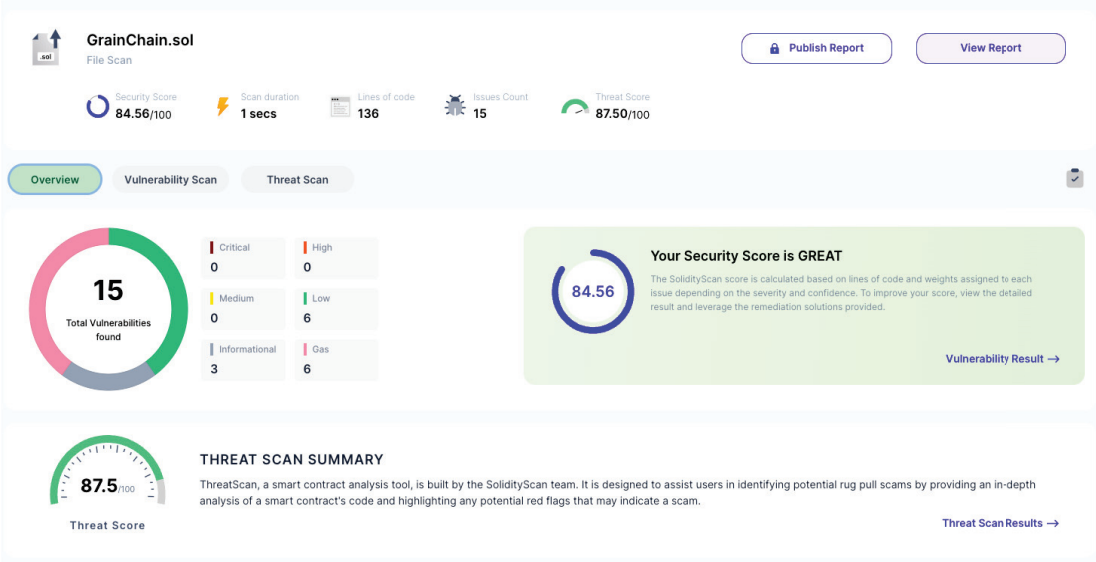


Figure 10. SolidityScan security analysis.

5.6.1. Security and Threat Scores Explanation

The security score is calculated based on the lines of code and the weights assigned to each issue depending on the severity and confidence of detection. The score reflects the overall security posture of the contract, with a higher score indicating fewer and less severe issues. In this case, a score of 84.56 suggests that the contract is generally secure but has some areas for improvement.

1. Lines of code (136): A relatively moderate codebase size, which typically helps in maintaining readability and manageability.
2. Issues count (15): The number of identified issues impacts the score, with each issue's severity contributing to the overall reduction of the score.
3. Severity and confidence: Issues with higher severity and detection confidence weigh more heavily against the security score.

5.6.2. Threat Score

The threat score is a measure of the potential risk posed by the identified issues. A higher threat score indicates a greater risk, taking into account the likelihood and impact of potential vulnerabilities being exploited. The score of 87.50 suggests a significant level of risk, warranting attention to the highlighted issues to mitigate potential threats.

6. Conclusions and Future Work

In conclusion, the proposed scheme demonstrates the transformational potential of integrating blockchain technology with IoT sensors to revolutionize grain supply chain operations. Milestones have presented significant improvements toward attaining more transparency, traceability, and security of the grain supply chain. The evidence of the scheme's viability can create additional value for stakeholders regarding the state of grain; the proposed solution will not only motivate but emphasize an essential contribution of such solutions in industry problem-solving areas. Thus, a blockchain-based grain chain management solution is envisaged with an architectural style that uses a layered approach that includes data storage, execution of smart contracts, consensus algorithms, development of user interface, and integration of the system. The architecture can ensure

transparency, security, and efficiency in supply chain transactions. Such conclusive evidence places solutions based on blockchain in the category of a viable and impactful approach towards grain supply chain management with an enhancement of its overall effectiveness and reliability. The results add to this knowledge bank as it pertains to supporting the adoption of innovative technologies in a fluid and increasingly complex grain supply chain landscape.

Author Contributions: Conceptualization, R.A., K.S.T. and R.S.; methodology, K.S.T.; software, R.A.; validation, R.A. and R.S.; formal analysis, R.A., K.S.T. and R.S.; investigation, R.A.; resources, R.A.; data curation, K.S.T. and R.A.; writing—original draft preparation, K.S.T.; writing—review and editing, R.A. and R.S.; visualization, K.S.T.; supervision, R.A.; project administration, R.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data can be shared up on request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Zhang, X.; Sun, P.; Xu, J.; Wang, X.; Yu, J.; Zhao, Z.; Dong, Y. Blockchain-based safety management system for the grain supply chain. *IEEE Access* **2020**, *8*, 36398–36410. [CrossRef]
2. Zhang, X.; Li, Y.; Peng, X.; Zhao, Z.; Han, J.; Xu, J. Information traceability model for the grain and oil food supply chain based on trusted identification and trusted blockchain. *Int. J. Environ. Res. Public Health* **2022**, *19*, 6594. [CrossRef] [PubMed]
3. Lucena, P.; Binotto, A.P.; Momo, F.D.S.; Kim, H. A case study for grain quality assurance tracking based on a Blockchain business network. *arXiv* **2018**, arXiv:1803.07877.
4. Zhang, Y.; Wu, X.; Ge, H.; Jiang, Y.; Sun, Z.; Ji, X.; Jia, Z.; Cui, G. A blockchain-based traceability model for grain and oil food supply chain. *Foods* **2023**, *12*, 3235. [CrossRef] [PubMed]
5. Thakur, K.S.; Ahuja, R. Empowering Fast-Moving Consumer Goods Supply Chain Leveraging Blockchain. In *International Conference on Recent Trends in Machine Learning, IOT, Smart Cities & Applications*; Springer Nature Singapore: Singapore, 2023; pp. 729–740.
6. Albaaji, G.F.; Chandra, S.V. Blockchain technology in agriculture: Digitizing the Iraqi agricultural environment. *Environ. Dev. Sustain.* **2024**, 1–12. [CrossRef]
7. Farooq, M.S.; Ansari, Z.K.; Alvi, A.; Rustam, F.; Díez, I.D.L.T.; Mazón, J.L.V.; Rodríguez, C.L.; Ashraf, I. Blockchain based transparent and reliable framework for wheat crop supply chain. *PLoS ONE* **2024**, *19*, e0295036. [CrossRef] [PubMed]
8. Balamurugan, S.; Ayyasamy, A.; Joseph, K.S. IoT-Blockchain driven traceability techniques for improved safety measures in food supply chain. *Int. J. Inf. Technol.* **2022**, *14*, 1087–1098. [CrossRef]
9. Caro, M.P.; Ali, M.S.; Vecchio, M.; Giaffreda, R. Blockchain-based traceability in Agri-Food supply chain management: A practical implementation. In Proceedings of the 2018 IoT Vertical and Topical Summit on Agriculture-Tuscany (IOT Tuscany), Tuscany, Italy, 8–9 May 2018; pp. 1–4.
10. Chang, A.; El-Rayes, N.; Shi, J. Blockchain Technology for Supply Chain Management: A Comprehensive Review. *J. Supply Chain Manag.* **2022**, *1*, 191–205. [CrossRef]
11. Patel, H.; Shrimali, B. AgriOnBlock: Secured data harvesting for agriculture sector using blockchain technology. *ICT Express* **2023**, *9*, 150–159. [CrossRef]
12. Zeng, H.; Dhiman, G.; Sharma, A.; Sharma, A.; Tselykh, A. An IoT and Blockchain-based approach for the smart water management system in agriculture. *Expert Syst.* **2023**, *40*, e12892. [CrossRef]
13. Ahuja, R.; Chugh, S.; Singh, R. SeedChain: A secure and transparent blockchain-driven framework to revolutionize the seed supply chain. *Future Internet* **2024**, *16*, 132. [CrossRef]
14. Devrani, S.; Ahuja, R.; Goel, A.; Kharbanda, S.S. A blockchain-driven framework for issuance of NFT-based warranty to customers on E-commerce. In *International Conference on Multi-Disciplinary Trends in Artificial Intelligence*; Springer Nature Switzerland: Cham, Switzerland, 2023; pp. 265–276.
15. Gervais, A.; Karame, G.O.; Wüst, K.; Glykantzis, V.; Ritzdorf, H.; Capkun, S. On the security and performance of proof of work blockchains. In Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, Vienna, Austria, 24–28 October 2016; pp. 3–16.
16. Lim, M.K.; Li, Y.; Wang, C.; Tseng, M.L. A literature review of blockchain technology applications in supply chains: A comprehensive analysis of themes, methodologies and industries. *Comput. Ind. Eng.* **2021**, *154*, 107133. [CrossRef]
17. Yan, Z.; Gan, G.; Riad, K. BC-PDS: Protecting Privacy and Self-Sovereignty through BlockChains for OpenPDS. In Proceedings of the 2017 IEEE Symposium on Service-Oriented System Engineering (SOSE), San Francisco, CA, USA, 6–9 April 2017; pp. 138–144.
18. Lai, R.; Chuen, D.L.K. Blockchain–from public to private. In *Handbook of Blockchain, Digital Finance, and Inclusion*; Academic Press: Cambridge, MA, USA, 2018; Volume 2, pp. 145–177.

19. Patidar, A.; Sharma, M.; Agrawal, R.; Sangwan, K.S. A Smart Contracts and Tokenization Enabled Permissioned Blockchain Framework for the Food Supply Chain. In *IFIP International Conference on Advances in Production Management Systems*; Springer International Publishing: Cham, Switzerland, 2021; pp. 228–235.
20. Mishra, H.; Maheshwari, P. Blockchain in Indian Public Distribution System: A conceptual framework to prevent leakage of the supplies and its enablers and disablers. *J. Glob. Oper. Strateg. Sourc.* **2021**, *14*, 312–335. [CrossRef]
21. Menon, S.; Jain, K. Blockchain Technology for Transparency in Agri-Food Supply Chain: Use Cases, Limitations, and Future Directions. *J. Distrib. Ledger Traceability Transpar.* **2024**, *71*, 106–120. [CrossRef]
22. Ellahi, R.M.; Wood, L.C.; Bekhit, A.E.-D.A. Blockchain-Based Frameworks for Food Traceability: A Systematic Review. *Foods* **2023**, *12*, 3026. [CrossRef] [PubMed]
23. Mekkouri, M.; Hennebert, C. Practices for Assessing the Security Level of Solidity Smart Contracts. In *International Symposium on Foundations and Practice of Security*; Springer Nature Switzerland: Cham, Switzerland, 2023; pp. 71–86.
24. Bhat, S.A.; Huang, N.-F.; Sofi, I.B.; Sultan, M. Agriculture-Food Supply Chain Management Based on Blockchain and IoT: A Narrative on Enterprise Blockchain Interoperability. *Agriculture* **2022**, *12*, 40. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

electronics

MDPI

Article

Enhancing Scalability of C-V2X and DSRC Vehicular Communication Protocols with LoRa 2.4 GHz in the Scenario of Urban Traffic Systems

Eduard Zadobrischi [1,*] and Ștefan Havriliuc [1,2]

[1] Department of Computers, Electronics and Automation, Faculty of Electrical Engineering and Computer Science, "Stefan cel Mare" University, No. 13. Str. Universitatii, 720229 Suceava, Romania; stefan.havriliuc@usm.ro

[2] Integrated Center for Research, Development and Innovation in Advanced Materials, Nanotechnologies and Distributed Systems for Fabrication and Control, "Stefan cel Mare" University, 720229 Suceava, Romania

* Correspondence: eduard.zadobrischi@usm.ro

Abstract: In the realm of Intelligent Transportation Systems (ITS), vehicular communication technologies such as Dedicated Short-Range Communications (DSRC), Cellular Vehicle-to-Everything (C-V2X), and LoRa 2.4 GHz play crucial roles in enhancing road safety, reducing traffic congestion, and improving transport efficiency. This article explores the integration of these communication protocols within smart intersections, emphasizing their capabilities and synergies. DSRC, based on IEEE 802.11p, provides reliable short-range communication with data rates up to 27 Mbps and latencies below 50 ms, ideal for real-time safety applications. C-V2X leverages LTE and 5G networks, offering broader coverage up to 10 km and supporting data rates up to 100 Mbps, with latencies as low as 20 ms in direct communication mode (PC5). LoRa 2.4 GHz, known for its long-range (up to 15 km in rural areas, 1–2 km in urban settings) and low-power characteristics, offers data rates between 0.3 and 37.5 kbps, suitable for non-critical data exchange and infrastructure monitoring. The study evaluates the performance and interoperability of these technologies in urban environments, focusing on data latency, transmission reliability, and scalability. Experimental results from simulated and real-world scenarios show that DSRC maintains reliable communication within 1 km with minimal interference. C-V2X demonstrates superior scalability and coverage, maintaining robust communication over several kilometers in high-density urban settings. LoRa 2.4 GHz exhibits excellent penetration through urban obstacles, maintaining connectivity and efficient data transmission with packet error rates below 10%.

Keywords: vehicle-to-vehicle; vehicle-to-infrastructure; vehicle-to-everything; dedicated short-range communication; cellular-V2X; LoRa; intelligent system transportation; vehicular communication

1. Introduction

According to global reports, but also according to many other organizations, including the World Health Organization, it is estimated that more than 50 million people are injured every year as a result of road crashes, and more than 1 million lose their lives as a result of human error or poor road infrastructure without management systems [1]. These reports also show that many more people lose their lives as a result of road accidents, especially in the 15 to 35 age group, which is a matter of great concern to both the academic and administrative worlds. In addition to the loss of human lives, these reports also highlight the financial effort to which the gross domestic products of the affected countries are subjected, and the percentage is around 3%.

Several congresses and assemblies have even discussed resolutions on improving global road safety, which aims to reduce the loss of human life by as much as 50% by 2030, including the reduction of carbon emissions by the same percentage by the end of 2050 [2].

A series of measures implemented by the International Transport Forum has begun to show through in the form of a reduction of around 10% in the number of fatalities, reinforcing and highlighting just how important these issues are. According to a report issued by the U.S. Department of Transportation's Innovative Technology and Research Administration, the deployment of road safety applications that use inter-vehicle communications and co-communications between vehicles and infrastructure could reduce road crashes by up to 81% [3]. This reduction is distributed as follows: 83% for crashes involving light vehicles and 72% for those involving heavy vehicles. Publications supporting the major transformations in transport, part of the Europe on the Move initiative and the EU Road Safety Policy Framework 2021–2030, focus on mobility [4].

It can be said that from all the current stages analyzed, globally, several directions and developments can be identified, including implementations of cooperative intelligent transportation systems and also dedicated systems (C-ITS). Therefore, vehicular communications are an essential part of these solutions, aiming to include traffic participants, but also pedestrians, in a single network. These communications are found in several forms: communications between one or more vehicles, vehicle-to-vehicle (V2V), between vehicles and road infrastructure (V2I), between pedestrians and infrastructure (I2P), or between vehicles and pedestrians (V2P).

Cellular Vehicle-to-Everything (C-V2X) and dedicated short range communications (DSRC) are two key technologies for vehicle communications. C-V2X, based on 4th Generation (4G) and 5th Generation (5G) cellular technologies, enables vehicles to communicate directly with each other, as well as road infrastructure and pedestrians, offering low latency and high bandwidth capacity. C-V2X operates in direct mode (PC5 models) and network mode (Uu models), providing flexibility and reliability in communications.

DSRC, on the other hand, utilizes dedicated frequencies in the 5.9 GHz band for vehicular communications, with very low latency and high reliability, and is deployed in various road safety and traffic management systems. Although DSRC was initially the predominant standard for vehicle-to-everything V2X, C-V2X is gaining ground due to its integration with existing cellular networks and support for advanced 5G-based applications. LoRa (Long Range) technology at 2.4 GHz is another emerging solution for vehicular communications, in particular for applications requiring long-distance data transfers but at lower data rates. The 2.4 GHz LoRa offers significant advantages in terms of low power consumption and the ability to cover long distances, making it ideal for vehicle-to-infrastructure (V2I) and vehicle-to-vehicle (V2P) communications in urban and rural environments. This technology is particularly useful for infrastructure condition monitoring and real-time data collection applications.

The initial development of these communications aimed to facilitate the transmission of information in traffic control systems, increase road safety, and prevent accidents. Subsequently, with technological advances and the development of new communication infrastructures and new types of cars, they have also been deployed in the field of autonomous vehicles to facilitate the transfer of data in larger quantities [5]. These developments are also accompanied by significant challenges, in particular those related to high costs, the need to develop new large-scale infrastructures, the harmonization of technologies developed by vehicle manufacturers, and the different standards existing worldwide. On the other hand, the development of technologies that may prove disruptive to existing vehicular communication technologies, such as new 5th generation (5G) cellular communication technologies and 6th generation prototypes, bring a number of new challenges to vehicular communication researchers while also offering potential solutions to various open problems.

By analyzing the challenges of the field, I have tried to develop, both simululatively and practically, a modular, scalable, and adaptive solution that can perform communications and computational analysis useful in road safety applications as well as in driver assistance. The solutions are based both on existing hardware components and on modules designed and implemented to streamline certain operations within the applications. The

research has led to the development of a modular hardware-software architecture and a hybrid communications and vehicular computing applications platform.

In the domain of Intelligent Transportation Systems (ITS), vehicular communication technologies like dedicated short-range communications (DSRC), cellular vehicle-to-everything (C-V2X), and LoRa 2.4 GHz are crucial for enhancing road safety, reducing traffic congestion, and improving transport efficiency. This study examines the integration of these communication protocols within smart intersections, highlighting their capabilities and synergies. This study evaluates the performance and interoperability of these technologies in urban environments, focusing on data latency, transmission reliability, and scalability. Experimental results from both simulated and real-world scenarios show that DSRC maintains reliable communication within 1 km, with minimal interference in less congested bands. C-V2X demonstrates superior scalability and coverage, maintaining robust communication over several kilometers even in high-density urban settings, with data rates that support high-throughput applications. LoRa 2.4 GHz exhibits excellent penetration through urban obstacles, maintaining connectivity and efficient data transmission with packet error rates below 10% under typical urban interference conditions. The findings highlight the complementary nature of these technologies: DSRC excels in latency-sensitive safety applications, C-V2X provides extensive coverage and high data rates suitable for real-time traffic management and autonomous vehicle coordination, while LoRa 2.4 GHz offers a reliable backbone for low-power, long-range communication in smart infrastructure monitoring. Furthermore, this research underscores the potential of these technologies to support the transition towards fully autonomous vehicles and smart mobility infrastructure. A hybrid communication approach, integrating DSRC, C-V2X, and LoRa 2.4 GHz, is essential for creating resilient and adaptive ITS solutions. This approach ensures robust and efficient data transmission across different urban scenarios, significantly contributing to the development of future smart cities [6].

The article begins with a brief introduction to the field and the importance of the communications presented and studied, and in Section 2, information is presented to conclude and discuss what has been obtained in the article and ideas for the development of V2X, DSRC, and LoRa 2.4 GHz communications, as well as the current state of the art in traffic safety and road safety systems. Section 3 exposes elements and aspects that summarize the actual progress of these communication protocols, as well as simulating the impact that a mixed architecture including LoRa technology will have in order to have a much better scalability and extended communication areas. In Section 4, information is presented to conclude and discuss what has been obtained in the article and how this direction will be approached in the future.

2. State of Developments of Vehicular Communications Based on C-V2X, DSRC and LoRa

We can say that the interconnection of vehicles can be realized by the automotive industry, but when the ideal protocol for such communications will be chosen and how it will be perfectly scalable regardless of the way information is transmitted depends on the density of vehicles or the degree of urban congestion.

For these connections, wireless protocols are required, with 802.11 [7] and cellular being the main options for these systems. There is still debate within the industry as to which protocol will prevail or whether all will migrate to cellular with the advent of 5G. If 802.11p is chosen, then a Dedicated Short-range Communication (DSRC) transceiver is required for wireless communication. There are strong arguments to favor DSRC in this debate, and some components are already available and compatible with this protocol. According to the analysis, it can be stated that in October 1999, the United States Federal Communications Commission (FCC) succeeded in allocating 75 Mhy within the 5.9 GHY band for use in intelligent transportation systems (ITS) operating on the DSRC protocol. However, the objective was to be able to provide a dedicated channel to realize vehicle-to-vehicle communications. Then the European Union realized and adopted a similar measure

by 2008, and they succeeded in allocating 30 MHz in the same band for ITS systems. This was conducted for countries such as Japan, Singapore, and European countries that use DSRC to collect tolls on certain road sectors. DSRC is a wireless protocol, derived from WiFi technology, operating at the upper limit of the WiFi 5 band. Specifically, this protocol operates in the frequency range 5.850–5.925 GHz, using modulations such as binary phase-shift keying (BPSK), quadrature phase shift keying (QPSK), and quadrature modulation (16QAM) and 64QAM, in combination with orthogonal frequency-division multiplexing (OFDM) and CSMA.

The protocol complies with the specifications defined in the IEEE 802.11p standard [8]. Compared to DSRC, C-V2X (Cellular Vehicle-to-Everything) vehicular communication technology offers additional functionalities such as direct communication with pedestrians (called V2P), thus marking a considerable step forward. Both protocols have the ability to connect to cellular networks for non-safety applications, making DSRC a hybrid system for V2X communications. On a printed circuit board (PCB), the layout methods used for WiFi can also be applied to DSRC, facilitating rapid integration of these systems for many designers, see Table 1 in in which the characteristics and particularities of DSRC and C-V2X are presented.

Table 1. Characteristics, features, and parameters of DSRC and C-V2X communication protocols.

Parameters	DSRC	C-V2X
Protocol	IEEE 802.11p	C-V2X Rel. 14/15 (for 4G-LTE) and Rel 16 (for 5G NR)
Modulation	BPSK/QPSK/16QAM/64QAM, Orthogonal Frequency Division Multiplexing (OFDM)	QPSK/16QAM/64QAM for 4G-LTE, OFDM (QAM + FDM) for 5G, Single Carrier—Frequency Division Multiplexing (SC-FDM)
Transmission time	0.4 ms	1 ms
Symbol duration	8 µS	71 µS
Security	Public-key cryptography requires the isolation of secure functions from insecure ones.	Same
Transfer rate	27 Mbps	4–12 Mbps, 50 Mbps peak download
Time sync.	Asynchronous	Synchronous
Maximum latency	~150 ms	Less than 50 ms in 4G-LTE tests

2.1. Current Status of C-V2X Developments

Cellular vehicle-to-everything (C-V2X) is a type of technology that enables wireless communications, allowing vehicles to communicate with each other, pedestrians, and the surrounding infrastructure, all over a cellular network. Therefore, C-V2X is being developed to improve road safety, reduce traffic congestion, and significantly increase the efficiency of intelligent transportation systems. However, the basis of this C-V2X technology is founded on heterogeneous network environment; it utilizes (LTE) and 5G cellular communication technologies [9], which are extremely popular and highly praised for realizing high-speed data transfer [10,11].

Considering this direction to be extremely important, the development and standardization of this C-V2X technology has been achieved by several organizations and in direct partnership with 3GPP—3rd Generation Partnership Project, the European Telecommunications Standards Institute (ETSI), but also with the help of the Institute of Electrical and Electronics Engineers (IEEE), as well as the 5G Automotive Associations (5GAA).

According to all speculations, it is expected that C-V2X will play an essential role in the development of intelligent transportation systems (ITS) [12] and in vehicle networks, which are interconnected to the infrastructure and to each other. Urban areas, including roads, are an extremely important element of road infrastructures and an important factor that supports the mobility or logistics of vehicles. We can say that it is at the basis of social and economic growth, but also at the interconnection of various routes, such as direct links between cities, ports, airports, or even links between neighboring countries. With the

growth of the last decade, the migration of populations, urban progress, industrialization, and also the large number of vehicles in active use, we are realizing the negative side of these developments [13].

According to the United Nations, in 1950, 30% of the world's population lived in cities; today, this percentage has risen to 54%. The United Nations estimates suggest that by 2050, 66% of the global population will live in urban areas [6,14,15]. Specific societal trends that will influence the future of transportation include:

- ✓ A new wave of urbanization will put pressure on existing transport infrastructure.
- ✓ Increasingly stringent emission policies and regulations imposed by governments.
- ✓ Increased pressure on public transport and logistics/delivery services due to the emergence of different business prospects.
- ✓ Effective use of real-time traffic data and open data for traffic management.

In order to manage in time all these problems, a first and main solution, which can be addressed in the directions presented and in order to reduce all this vicious circle, is the improvement of road infrastructures and highlighting the various methods to achieve an intelligent transport able to manage in real time every negative event, and this approach to be sustainable from all points of view, especially the economic and environmental for a sustained sustainability. Looking ahead over the next 2 to 5 years, various key transformations in the transportation environment are anticipated, including:

- ✓ Elimination of road accidents,
- ✓ Implementation of an adaptive support system to manage high traffic flows,
- ✓ Increasing vehicle utilization, and
- ✓ Promoting more efficient and greener transportation, including zero emission vehicles [16,17].

The rapid evolution of information technology (IT), sensors, geolocation, and communication technologies has facilitated the development of remarkable applications that enable the implementation of strategies in an intelligent way, through a better understanding of the environment. This advancement led to the development of intelligent vehicle-highway systems (IVHS), which were later officially renamed to intelligent transportation systems (ITS) in 1994 by the United States Department of Transportation (USDOT) to reflect a broadened mission that includes all aspects of public transportation and intermodal connections [18].

Intelligent Transportation Systems (ITS) are defined as "a system designed to solve the various problems facing road traffic by exchanging information between roads, people, vehicles, and cities" [4]. The main objectives of ITS include the evaluation, development, analysis, and integration of advanced technologies to achieve efficiency, energy conservation, time savings, improved safety measures, enhanced sustainability, and the satisfaction of drivers, pedestrians, and other traffic participants [19–21]. By implementing ITS, it can transform an unorganized transport environment into an intelligent and connected system, thus addressing current road traffic challenges and promoting safer, more efficient, and more sustainable transport.

To achieve intelligent and adaptive management of transportation systems, the International Society of Automotive Engineers (SAE) [22] has developed a detailed roadmap with six levels of evolution, shown in Table 2. We are currently witnessing significant developments in autonomous vehicles. For example, various commercial companies such as Volvo, MG, Tesla, and other car manufacturers are developing and launching autonomous vehicles. However, reaching level 4 and 5 vehicle automation remains an ambitious goal for many automotive organizations and researchers.

An essential aspect for achieving full automation in transportation is the efficient exchange of information between all entities involved in traffic. In this context, cooperative—vehicle to everything (C-V2X) communication adds a vital dimension to the mission of intelligent transportation systems (ITS). It enables vehicle-to-vehicle (V2V), vehicle-to-vehicle (V2P), vehicle-to-infrastructure (V2I), and vehicle-to-network (V2N) communica-

tions, thus contributing to safer and more connected mobility in both line-of-sight (LOS) and line-of-sight overlay (NLOS) situations [23].

Table 2. LoRa features and characteristics. Experimental standard configurations used.

Parameters	LoRa Sub-GHz	LoRa 2.4 GHz
Spreading Factor (SF)	From 6 to 12	From 5 to 12
Bandwidth (kHz) (BW)	125, 250, 500	203, 406, 812, 1625
Coding Rate (CR)	4/5, 4/6, 4/7, 4/8	
Data Rates	183 bps–62.5 kbps	297 bps–202 kbps
Link Budget	168 dB (SX1276 chip)	144.5 dB (SX1280 chip)
Consumption	Tx: 28 mA at 13 dBm Rx: 11.5 mA SX1276 chip	Tx: 24 mA at 12.5 dBm Rx: 7 mA for BW = 812 SX1280 chip
Time sync.	Asynchronous	Synchronous
Maximum latency	~150 ms	Less than 50 ms in 4G-LTE tests

2.2. Current Status of DSRC Developments

When we discuss dedicated short-range communications (DSRC), we can say that it is a wireless communications protocol designed for the purpose of connecting vehicles (CV) within multiple applications or systems of type (ITS). This protocol is based on the IEEE 802.11p standard as part of the Wireless Access within Vehicular Environments (WAVE) system in the United States, using a 75 MHz spectrum with a frequency range between 5.850 and 5.925 GHz [24,25]. However, unlike the IEEE 802.11a standard [26], IEE 802.11p can optimize the mobility of connected vehicles by reducing the bandwidth by about 10 MHz, thus realizing a doubling of temporal resolution [27,28]. Knowing that it has a different broadcast nature, DSRC, which is based on IEEE 802.11p, does not require an acknowledgment of the frames being transmitted. Thus, to avoid the uncontrolled extension of contention frames and to increase latency, the MAC protocol that IEEE 802.11p uses strictly for contention adopts a fixed contention frame, replacing an exponential retreat [24,29].

The proven effectiveness of the DSRC protocol in safety applications and the widespread availability of DSRC devices have led major automobile manufacturers to adopt this technology in their fleets [30,31]. To address new technological requirements and challenges, a study group was formed in 2018 to develop next-generation standards such as IEEE 802.11bd [32]. These standards aim to optimize and improve the processes and operational modes of DSRC technology. According to the IEEE 802.11bd Working Group (TGbd), backward compatibility will enable 802.11p and 802.11bd compliant devices to communicate efficiently over the same operational channel, ensuring interoperability and continuity of functionality [33,34]. In direct accordance with the architecture of the implemented physical layer, also referred to as (PHY), the IEEE 802.11p standard is derived from IEEE 802.11a more precisely by reducing the sub-carrier spacing by a factor of two [33,34]. In view of the fast mobility characteristics of vehicular network, the reduced IEEE 802.11p sub-carrier spacing makes the DSRC communication link vulnerable to errors caused by Doppler effect, while the double symbol duration improves the performance in the presence of channel multipath effects.

Therefore, everything becomes paradoxical and we have a trade-off that we identify between the robustness of the link between the interference between symbols and the way the multipath spreading is relayerized, and the error is relative in the case of Doppler spreading [35]. There is also an architecture of the medium access control (MAC) layer of the 802.11p standard, which is based on the out-of-context mode of operation given by an underlying service set (OCB), where nodes can transmit and receive data as well as control frames without the need to form or identify a service set (SS) [36,37]. Another extremely notable aspect is that IEEE 802.11p incorporates the distributed coordination function (DCF) into the radio resource allocation procedure. DCF, we can say, is a MAC technique, which utilizes the Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) protocol, which, at its core, has an advanced binary fallback algorithm.

It can be said that IEEE 802.11p incorporates the Distributed Coordination Function (DCF) in the radio resource allocation process. DCF is a MAC strategy, which utilizes the Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) protocol together with a binary backoff algorithm. In a direct comparison with IEEE 802.11a, in an IEEE 802.11p based network, radio units compete for channel access using the CSMA/CA protocol with a fixed backoff time and contention window parameters [38]. According to previous expositions, exponential backoff time in IEEE 802.11a can lead to high message transmission latency, which is not suitable for a vehicular network; thus, a fixed backoff time in CSMA/CA of DSRC allows control of data transmission latency at the cost of reduced network scalability. Enhanced Distribution Channel Access (EDCA) is a DCF-based MAC technique that aims to solve the scalability problem in DSRC. In the EDCA approach, high priority messages are more likely to be transmitted than low priority messages [39,40].

2.3. Current State of LoRa 2.4 GHz Developments

We now turn our attention to the chapter focusing on emerging technologies like the Internet of Things (IoT), which have become integral to our daily lives. Despite their limited application in the automotive industry, these technologies are gaining traction in road safety and Intelligent Transportation Systems (ITS). Notably, LoRa is proving to be a highly scalable solution for IoT data collection, offering low power consumption and extensive coverage, making it particularly effective for wide-area monitoring and communication in ITS applications. We can say that LoRa usually operates on the sub-GHz in-band principle; this characteristic is specific to certain regions, and therefore applications are limited worldwide. However, the improved version of LoRa for the 2.4 GHz ISM band has realized a turning point and has stimulated the research community, as it allows the use of frequency channels used globally.

Therefore, LoRa is required to coexist with many other competing standards and systems in the 2.4 GHz bandwidth category, especially through the evolution of MAC and network protocols dedicated to Lo-RAWAN connectivity. Low Power Wide Area Networking (LPWAN) or LoRa, NB-IoT, has received a global impetus from the telecommunication industry and operators worldwide, who are deploying different systems based on these protocols [41].

Thus, LoRa [42] can be considered as one of the most representative LPWANs, especially because it uses spread-spectrum technologies in sub-GHz bands, which are unlicensed (e.g., 868 MHz for Europe, 915 MHz in North America). However, the parameters of the frequency channel or maximum duty cycle, the maximum transmission power, or the dedicated medium access mechanism may differ from one geographical area to another, and the design or hardware components have to be taken into account in the realization of new communication systems or protocols. Alternatively, in order to achieve substantial progress and to avoid such problems, a new LoRa version in ISM on 2.4 GHz frequency has been proposed, this band having similar technical specifications, only that they allow a worldwide coverage. In addition to these aspects, the advantages it offers must be highlighted, namely the robustness due to the use of its own LoRa modulation, the use of common frequency bands, and the lack of duty cycle limitations [43].

Knowing all the aspects presented above about the standards involved in the discussion and having the experience of previous implementations, we now focus on LoRa 2.4 GHz, considering that a significant contribution can be made to automotive, road safety, and intelligent road transportation systems by using such an approach. Even if the impact of a change in frequency has been studied for other modulations, the LoRa behaviors are not 100% known, especially in highly agglomerated areas for 2.4 GHz, as it is also used in protocols such as Wi-Fi, Bluetooth, ZigBee, in home appliances, or mobile telephony for children. Perhaps the biggest challenge in the current case is still, for the LoRa 2.4 GHz protocol, the mitigation of interference from a large number of devices; here we are talking about developed and intensely industrialized countries, but also the existence of other devices using equivalent technologies. The study focuses very much on the introduction of

LoRa 2.4 GHz for outdoor testing to cover heavily trafficked and congested areas, which are already tested and analyzed by DSRC—C-V2X based systems. In the case of LoRa 2.4 GHz, it can be said that it could have a coverage range of more than 2 km outdoors, with a packet delivery rate of at least 1 packet per 40–50 m indoors with the presence of obstacles [44].

Here we introduce the features that 2.4 GHz LoRa has, but also what sub-GHz LoRa offers. If we are to discuss the physical layer, both technologies seem to use the same type of chirp spread spectrum modulation, where the signal is encoded using a chirp sequence. The chirp can be said to be a linear frequency over a given bandwidth and time interval. Each chirp can be positive up if the frequency parsing increases or positive down if it decreases. If an encoding procedure is performed for this instant then there are some fluctuations [45].

When the chirp start frequency is encoded, its value is also encoded. In the case of LoRa 2.4 GHz standard parameter values are used as in the case of sub-GHz LoRa in the idea that there is a trade-off between reliability and how to optimize the power consumption and hence the transfer rates; these are also a spreading factor (SF) bandwidth, which we denote by (BW) or coding rate (CR). When increasing the SF value, the duration for each chirp may increase, but also the energy consumption, in this case being necessary to transmit the same message in the loop. We can say that by increasing the SF, it automatically scales the receiver sensitivity, which in turn increases the communication range. When the bandwidth (BW) decreases with it, the chirp duration is also more energy efficient, but the disadvantage is that it can reduce reliability. In all these cases, the coding rate (CR) can control how many redundancy bits can be added to the data in the total payloads. For this purpose, Table 2 has been created, summarizing the values for each protocol and the parameters for sub-GHz LoRa and 2.4 GHz LoRa. It can be said that LoRa 2.5 GHz can also realize the introduction of SF 5 for much larger bandwidths; these facilitate an exponential increase in the transfer rate [46,47].

Thus, the LoRa modulation offers several advantages, regardless of the frequency bands for a good operation, and here it must be emphasized that it offers an extremely robust oriented design and can allow extremely good and oriented links even around 168 dB or 145 dB with different adjustments, even using the Semtech SX1276 or 1280 chip (Semtech, Camarillo, CA, USA). In addition to these aspects, we also have in discussion the change of the communication intervals with the corresponding transfer rates and low power consumption, the presence of quasi-orthogonality of the spreading factor, which allows simultaneous communications on different SFs, and all this in conjunction with low power consumption, more precisely 11.5 mA for transmit and approximately 7 mA for receive [48,49].

3. Previous Results C-V2X and DSRC vs. LoRa 2.4 GHz Approach

We know from previous studies that when the experimental testing was carried out to achieve communication modalities via DSRC and a hybrid C-V2X architecture, direct communication distances with signal loss were achieved over a distance of more than 300 m under ideal conditions, and subsequently that the degree of complexity increased, reaching more than 800 m in a straight line and 1200 m on the taxiway. At that time, it was observed that if the measurements reach the performance limit for the mixed DSRC system, adaptive behavior under high mobility, or intense flow dynamics, problems with network scalability start to appear. Therefore, as the signal quality starts to decrease, channel coding can be compensated by multipath fading procedures, and in the respective cases the signal to noise ratio can be influenced by any other source identified in the area, even by the thermal factor. Measurements performed in [50] provide a clear exposition in terms of connectivity, reliability, scalability, and behavior for LoS or stationary and dynamic point measurements between nodes in the network [51,52].

It was notable that in tests of the engineered system that at around 9 ms, the connection was blocked, which prevented the exchange of information and caused the communication to drop. Subsequently, the device was able to receive GPS coordinates and messages, reestablishing communication for a short period. Analysis of this system behavior suggests

that in urban areas with heavy traffic and high dynamics, loss of connection and data transmission are affected. Data flow was interrupted at a distance of approximately 300–350 m since the last display. This distance was covered on a stretch of road with a level difference of 50–60 m and an inclination of more than 30°. The resumption of communication occurred at the maxim point, calculated over a straight-line distance of about 800 m and with a total level difference of 75 m from the transmit point. Under these conditions, the measurements do not indicate a notable feature, as the transfer rate is low and the number of packets received relative to those transmitted decreases exponentially. It is important to note that these scenarios were realized using only two OBU-Mk5 modules, without an RSU [53] or other type of ITS control center, which could have retrieved the information from node to node, optimizing the reception. According to Figure 1, we can expose the whole scenario tested in the presented case and using DSRC.

Figure 1. Representative area tested to expose the measurements performed by DSRC—C-V2X communications.

By directly analyzing the information that we obtain as a result of the measurements, we can say that decreasing speed is a measure to protect the communication link. When the signal quality decreases significantly, there is also a direct loss of access points created in the vehicular network. Although that absolute average error that we can exhibit is around 5 to 7 m, using a Gaussian distribution we can say that the standard deviation is 0.5 m, or is even negligible in conditions of extremely high data volume. When the degree of mobility is high and there are measurements between vehicles or between stationary points and vehicles through the level difference of the coordinates, deviations of about 7 m are also present in Figure 2. Each initialization function returns a descriptor (integer variable) file, allowing access to the updated interfaces. In the absence of information available on the CAN interface, default values are loaded to avoid locking the whole system. Query and event limiting functions, as well as the definition of descriptors, are used to filter information. The essential data that we target for timely extraction and communication within this system include: Min-Max of data packets, transfer rate, longitude, latitude, distance, transmit/receive signal, and transmit/receive noise. When we talk about dynamics in an outdoor environment, we refer to elements and objects that can appear and distort or disturb the test area in various ways. Under current conditions, the test environment is expanding such that the parking lot opens onto a heavily traveled section of road. Thus, in this context, the analysis process becomes much more complex and the GPS identification in the Rx—Tx direction may suffer message or data packet distortions. Next, we denote with C the car under study. GPS observations extracted by BSMs describe the location as $xi = (Li, li)$, $i \in \{0, \ldots Nd\ GPPS\}$ recorded at time ti, where Li and li represent the longitude and latitude of location i, respectively.

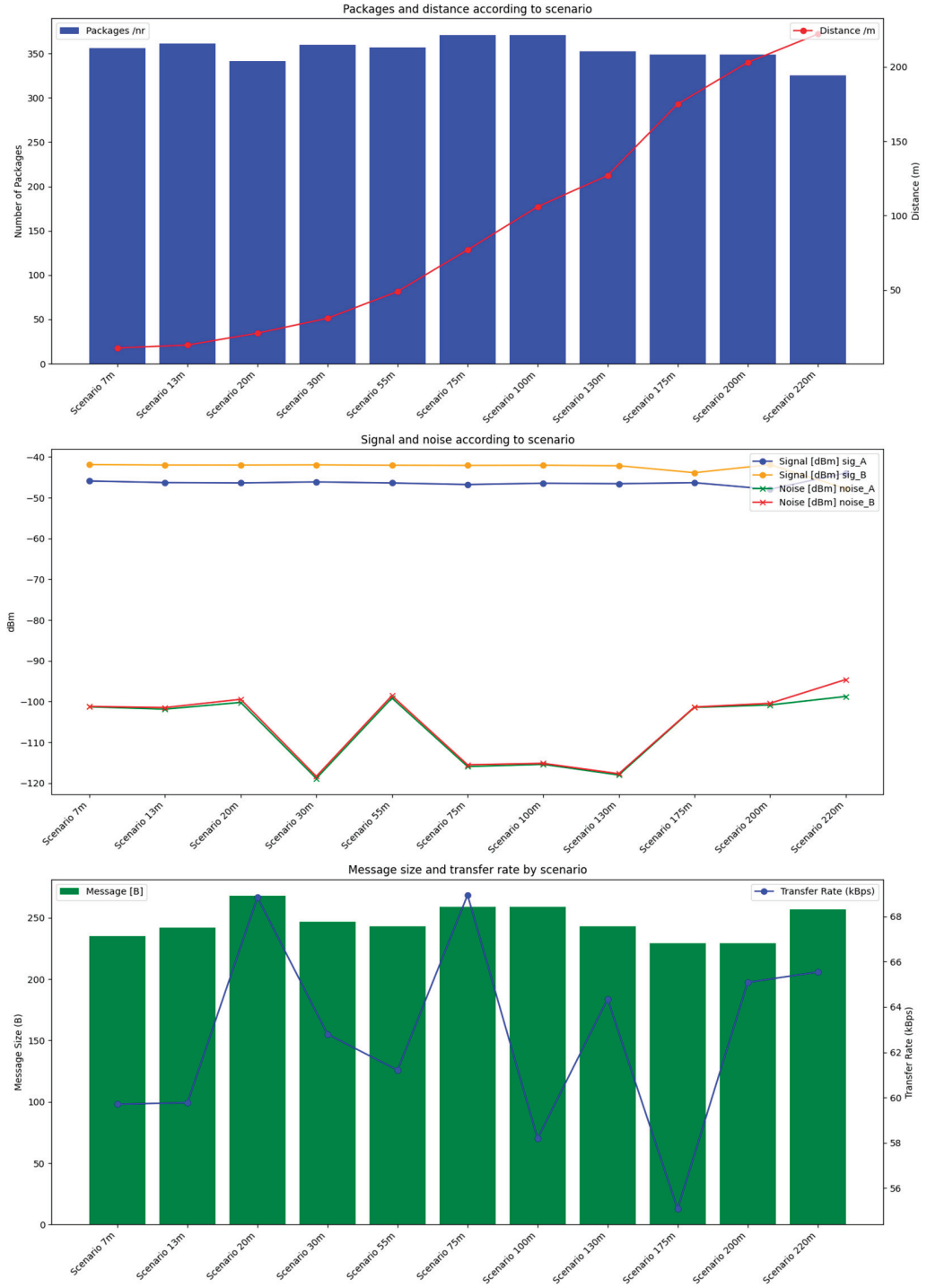


Figure 2. Data cohesion at simulated scenario level over more than 220 m distances from reception and data interpretation.

Therefore, the mean absolute error may vary depending on the speed at the time of the measurements; if in this case there is an upward trend, this is caused by the directly proportional relationship between speed and travel distance. Even in these cases, we are talking about covering a heavily trafficked area with a vehicle density exceeding 300–500/min at peak hours on a stretch of road where the maximum capacity is approximately 2–3 times lower; it is about aspects related to the existing infrastructure and how many vehicles can transition through the analyzed area from the perspective of the way in which the respective road sector is thought out. Under the conditions of the assimilation of a new

dynamic approach from an architectural point of view and the ITS type system, everything must be corroborated with autonomous vehicles. According to [54], it is possible to reach a transition of over 2000 vehicles, i.e., over 25–30 of vehicles per minute, in ideal conditions.

Knowing that the format and structure of the standard in the message transmission procedure must contain a header with all necessary fields, algo-criteria filtering is applied. When the accuracy of the GPS signal decreases, system-wide corrections are applied to correct the uncertainty, with the caveat that the variation should be around 2–3 for each coordinate. According to the tests, it has been observed that in the case of the routes with variations, the communication is affected and we do not have LoS. The communication can be disturbed even at 50 m, but to correct this we increase the effective radiated power (EIRP) [55], reaching about 30 dBm. The results of the whole scenario can be seen in Figure 3. We note that these results have been obtained by a system already tested and presented in the article referred to at the beginning of the current chapter, without the need to reiterate the information.

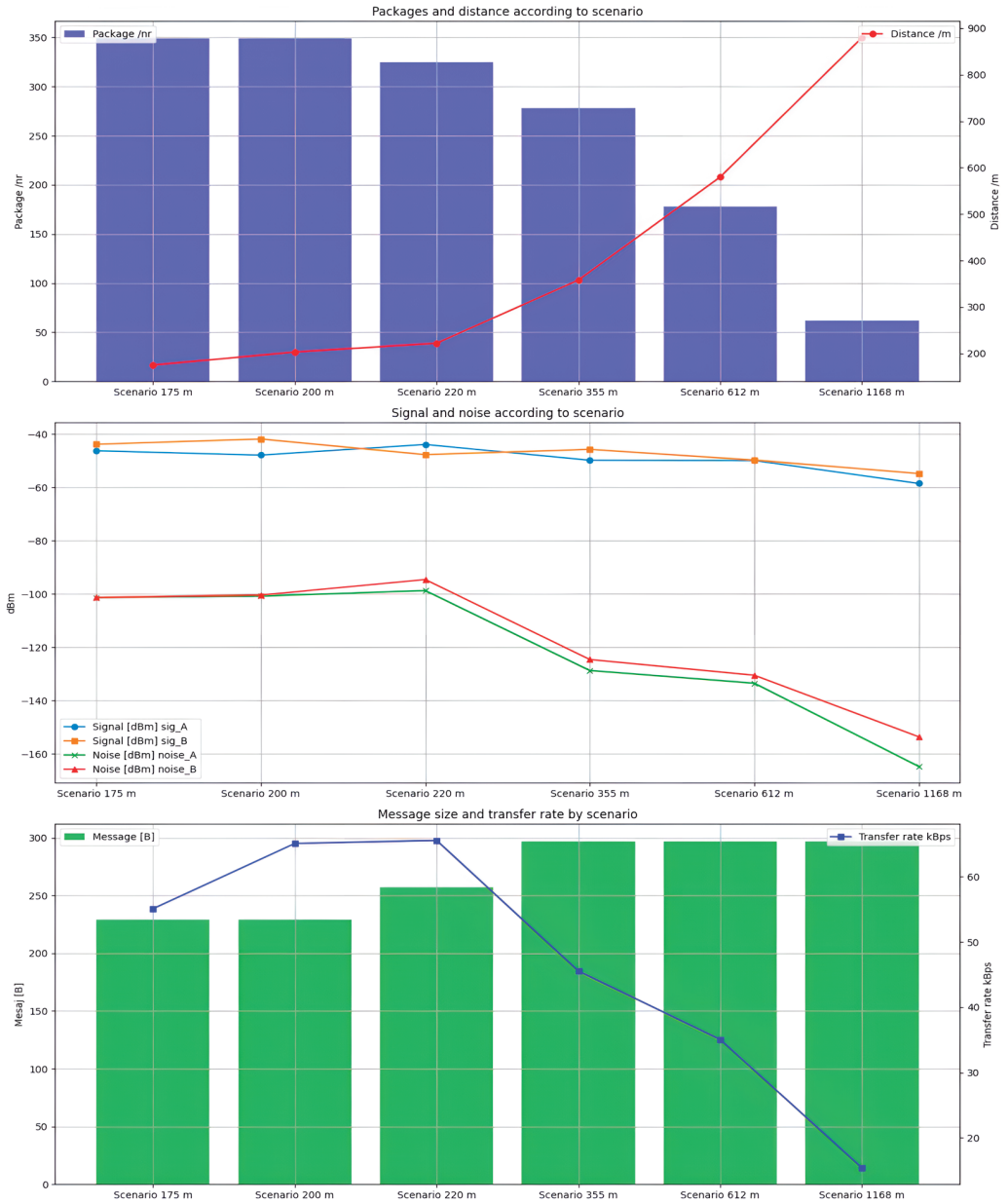


Figure 3. Data cohesion at simulated scenario level over distances of more than 600 m from reception and data interpretation.

3.1. Handling Particular Cases and Re-Testing Scenarios at Infrastructure Level Using LoRa 2.4 GHz

The architecture presented in the paper illustrates a communication system based on LoRa technology at 2.4 GHz frequency, used to connect various types of nodes (V2V, V2I, V2P) to an end application via the Internet. This architecture integrates multiple components and technologies to ensure efficient and reliable communication in a dynamic urban environment.

LoRa 2.4 GHz nodes are devices that communicate using LoRa technology at the 2.4 GHz frequency and are divided into three main categories: vehicles, people, and infrastructure. Vehicles are equipped with LoRa nodes to transmit data on position, vehicle status, fuel consumption, speed, direction, and other information relevant to safety and traffic management. Wearable nodes can transmit data on location, heart rate, physical activity level, or other personal parameters essential for health monitoring. Infrastructure, such as traffic lights, digital billboards, and other smart devices, is equipped with LoRa nodes to communicate traffic information, equipment status, weather conditions, and other data relevant for efficient urban traffic management [56].

The LoRa 2.4 GHz Gateway is a central point of the reception and transmission of data collected from LoRa nodes [57]. It communicates bidirectionally with the nodes and transmits the data to the Internet. The Gateway serves as a hub that aggregates data from the nodes and sends them to servers or analysis applications. The gateway can also perform data pre-processing to filter out irrelevant information and reduce the amount of data transmitted, thus optimizing the use of network resources. The Internet is the means of transporting data from the gateway to the end applications. Data are transmitted over Internet networks to analysis servers or platforms where they are stored, processed, and analyzed. These platforms can use machine learning and artificial intelligence algorithms to extract valuable insights from the collected data and generate real-time predictions or alerts. The app is the final interface where data are collected, analyzed, and presented to users. Users can access the app to visualize real-time data, generate detailed reports, and receive alerts based on predefined parameters. The application can also provide control and monitoring functions for infrastructure equipment, allowing rapid intervention when needed.

Figure 4 illustrates the integration of LoRa 2.4 GHz technology into Intelligent Traffic Systems (ITS), showing two main aspects: the architectural proposal for the infrastructure and the architecture of a LoRa transceiver. LoRa 2.4 GHz nodes are distributed in vehicles, pedestrians, and road infrastructure (traffic lights, traffic signs). These nodes enable vehicle-to-vehicle (V2V), vehicle-to-pedestrian (V2P), and vehicle-to-infrastructure (V2I) communication, facilitating the exchange of information critical for road safety and efficiency. The LoRa 2.4 GHz gateway Figure 4a collects data from all nodes and relays them to central applications over the internet. It acts as a central hub, aggregating and processing data for real-time analysis and monitoring. The LoRa transceiver Figure 4b is composed of analog and digital sensors that capture data from the environment. These data are processed by a microprocessor, which applies specific algorithms to extract useful information. The processed data are then transmitted through a radio transmitter using LoRa technology. The power supply provides continuous power to all components of the transceiver, ensuring reliable operation. The integration of the transceiver with the gateway and the Internet network enables efficient transmission of data to end applications, where it is collected, analyzed, and presented to users.

The communication flow starts with the transmission of data from LoRa nodes on vehicles, people, and infrastructure to the LoRa gateway at 2.4 GHz. The transmitted data may include information about location, operational status, environmental parameters, traffic conditions, and more. The gateway collects these data and forwards them over the Internet, a process that may include pre-processing the data to reduce the transmitted volume and filter out irrelevant information. The data then reach the end application, where they are stored, processed, and analyzed. Users can view the data in real time,

generate reports, and receive alerts based on predefined parameters, helping to improve the safety and efficiency of urban traffic.

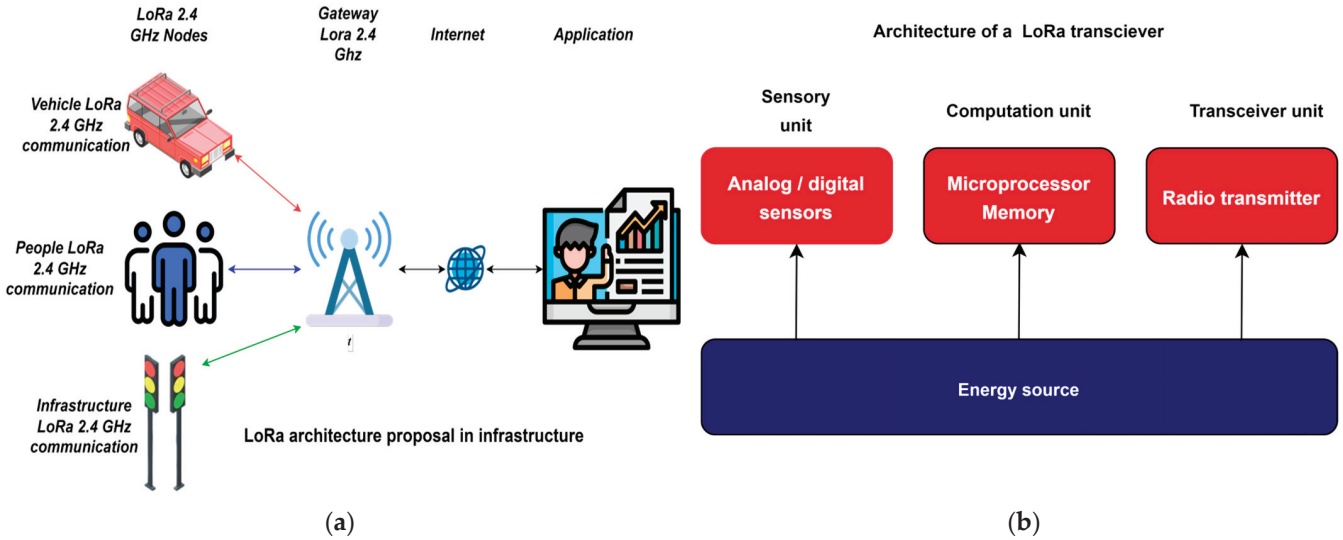


Figure 4. Generic architectures for the presentation of LoRa technology and its introduction in the structure of intelligent traffic systems: (**a**) LoTa architecture proposal in infrastructure; (**b**) Architecture of a LoRa Transceiver.

The general operation of the transceiver starts with the sensors in the sensor unit capturing data from the environment and sending these data to the computing unit. The microprocessor in the computing unit processes the data, applying specific algorithms to extract useful information. The processed data are then sent to the transceiver unit, where they are transmitted via radio transmitter using LoRa technology. The whole process is powered by an energy source, which ensures continuous and reliable operation of the transceiver. The sensor unit is responsible for capturing data from the environment. This may include analog or digital sensors that measure parameters such as location, motion, and environmental conditions. The sensors collect the raw data and send them to the computing unit for processing. The computing unit consists of a microprocessor and memory. The microprocessor processes data received from sensors, running algorithms for analysis and filtering. Memory is used to temporarily store data and programs running on the microprocessor. This unit is essential for pre-processing data before they are transmitted. The Transceiver Unit includes a radio transmitter that transmits the processed data to a gateway or other network devices. It uses 2.4 GHz LoRa technology to provide long-distance communications with low power consumption. The transceiver unit is responsible for transmitting data and receiving commands or updates from other devices on the network.

The efficient integration of these units ensures a continuous and reliable flow of information between the environmental sensors and the central data collection points, thus facilitating real-time monitoring and management of various parameters. LoRa technology, with its advantages of low power consumption and long-distance transmission capability, is ideal for IoT applications in urban and rural environments, where wide coverage and long battery life of the LoRA devices is essential. Figure 5 shows the interconnectivity offered by LoRa for the other types of communication protocols, such as V2V, V2I, and V2P, and how LoRa re-realizes these procedures for road and pedestrian infrastructures.

The LoRa 2.4 GHz network consists of multiple nodes, which can be vehicles, infrastructure (traffic lights, traffic signs), and devices carried by pedestrians. The LoRa Gateway acts as a central point to collect data from all nodes in the network and relays these data to a central application over the Internet.

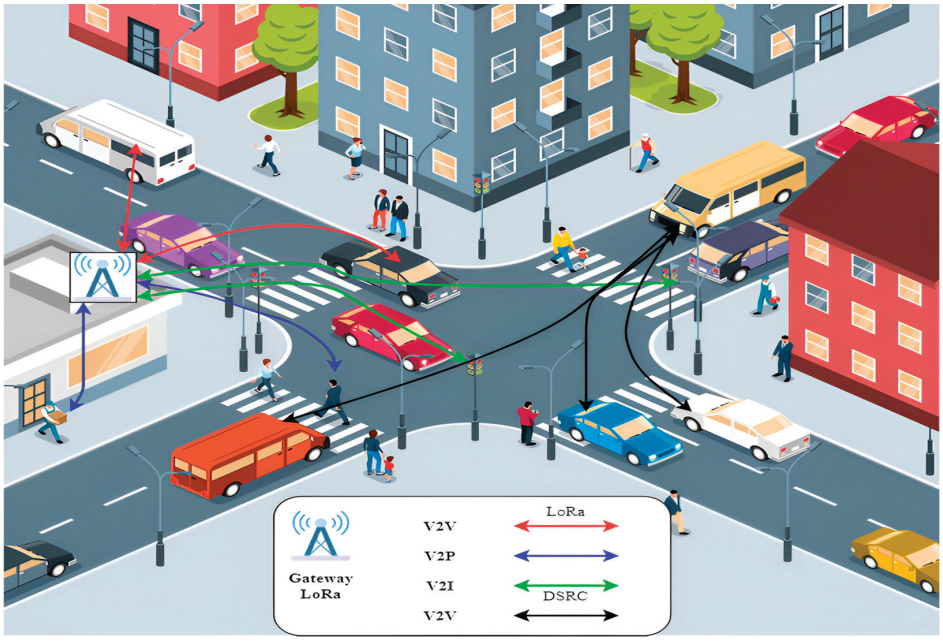


Figure 5. The representative diagram in which all the discussed protocols are introduced and how they are interconnected in the ITS systems dedicated to traffic management in urban environments.

3.2. Practical and Deterministic Analysis of LoRa Capabilities within Intelligent Transport Systems and Smart Cities

In the first step we evaluated the signal spread computation modalities, but also how LoRa performs, and for this aspect we have to validate each parameter, including signal strength, bit error rate (BER), and latency. We can say that one of the most essential formulas used to calculate the signal loss for distances is the Friis equation for the attenuation of a signal in free space:

$$P_r = P_r + G_t + G_r - L_r, \tag{1}$$

P_r is the received signal strength (dBm), P_t represent the power of the transmitted signal, G_t is the transmitter antenna gain (dBi), G_r is the gain of the receiver antenna (dBi), and L_p represents loss of track (d). Band can be calculated as follows:

$$L_p = 20\, log_{10}(d) + 20\, log_{10}(f) + 20\, log_{10}\left(\frac{4\pi}{c}\right), \tag{2}$$

d is the distance between transmitter and receiver (m), f represents the signal frequency (Hz), and c is the speed of light (3×10^8 m/s). The distance from the term d, when $20\, log_{10}(d)$, represents the path loss due to the distance the signal travels. As distance increases, path loss increases logarithmically. When we discuss the term frequency, $20\, log_{10}(f)$ accounts for the frequency dependence of the path loss. Higher frequencies experience higher losses

However, we also consider the calculation of the Bit Error Rate (BER), being a key performance indicator for LoRa networks, and in this case, we also have to take into account the Chirp Spread Spectrum; we treat this case as follows:

$$BER = \frac{1}{2}\, e^{-\frac{SNR}{2}}, \tag{3}$$

SNR is the signal-to-noise ratio and is calculated as follows:

$$SNR = \frac{P_r}{N_0}, \tag{4}$$

where N_0 represents the spectral denoising of noise ($\frac{\text{dBm}}{\text{Hz}}$).

Latency calculation aspects were also taken into account, and the latency is related to the Time on Air (ToA) for a packet, which is also influenced by the spreading factor (SF), bandwidth (BW), or coding rate (CR):

$$ToA = \frac{P_r}{N_0}\left(\frac{n_p + n_s}{R_b}\right) \times 1000 \tag{5}$$

Each element being noted in the following form: n_p, number of preamble bytes, n_S number of payload bytes, and R_b, bit rate, which we calculate as follows:

$$R_b = \frac{SF \times BW}{2^{SF}}, \tag{6}$$

where SF stands for Spreading Factor and BW for Bandwidth. In all these cases, research has also been extended to the way in which the path can be lost, namely Path Loss in dB:

$$L_p = 20\, log_{10}(d) + 20\, log_{10}(f) + 20\, log_{10} - 147.55 \tag{7}$$

For the connection given at the level of controlled measurements in dynamic environments, we have channeled our attention to the LoRa connection to highlight the reliability, consumption, and robustness, but also the low costs compared to other technologies. Table 3 shows the practical parameters and data used in the implementation of the LoRa 2.4 GHz communication prototype. These are essential for evaluating system performance under real-world conditions. The power of the transmitted signal (Pt) is 14 dBm. This is the effective signal strength emitted by the transmitter. The transmitter antenna gain (Gt) is 2 dBi. This measures how effectively the antenna directs the emitted power. The antenna gain of the receiver (Gr) is 8 dBi, indicating the efficiency of the antenna in receiving the signal. The distance (d) between transmitter and receiver is 500 m, and the signal must cover this distance.

Table 3. Practical parameters and data for LoRa 2.4 GHz prototype for implementation.

Properties	Technical Features
Power transmitted (P_t)	14 dBm
Transmitter antenna gain (G_t)	2 dBi
Receiver antenna gain (G_r)	8 dBi
Distance (d)	500 m
Frequency (f)	2.4 GHz
Noise spectral density (N_0)	−174 dBm/Hz

The frequency (f) used for transmission is 2.4 GHz, characteristic of the LoRa technology used. The noise spectral density (N_0) is −174 dBm/Hz. This is the noise level present in the transmission channel. Based on this information, the received signal strength can be calculated using the parameters in the table. These parameters are crucial for determining the performance of the communication system and for optimizing the settings needed to ensure reliable and efficient data transmission.

We calculate the received signal strength based on the following information:

$$L_p = 20\, log_{10}(500) + 20\, log_{10}\left(2.4 \times 10^9\right) - 147.55 \tag{8}$$

$$L_p = 53.98 + 193.98 - 147.55 = 100.41 \text{ dB}$$

$$P_r = 14 + 2 + 8 - 100 = -76.41 \text{ dBm}$$

To find the SNR we have the following relation:

$$SNR = P_r - N_0 = -76.41 - (-174) = 97.59 \text{ dB}, \tag{9}$$

and for BER, we have the following calculated relation:

$$BER = \frac{1}{2} e^{-\frac{97.59}{2}} \approx 2.82 \times 10^{-22} \tag{10}$$

Therefore, we end with a final calculation, that of ToA, when we assume SF = 10, BW = 125 kHz, and payload = 20 bytes:

$$R_b = \frac{10 \times 125 \times 10^3}{2^{10}} = 1220 \text{ bps} \tag{11}$$

$$ToA = \left(\frac{8.20 + (4.25 + 1 + 8 + 16)}{1220} \right) \times 1000 \approx 39.34 \text{ ms}$$

We can say that these calculations expose the way in which the performance of the proposed LoRa network has been evaluated, highlighting different conditions or settings, as this aspect is also based on an analysis for performance optimization and ensuring stable communication between vehicular applications and their connection to ITS systems. In the context of developing sensor networks for smart city applications, analyzing the propagation of radio signals is essential to ensure optimal coverage and high performance. In this chapter, we present the use of a ray tracing model to evaluate the propagation of LoRa signals at 2.4 GHz frequency in a complex urban area. This approach is comparable to previous evaluations carried out for DSRC and C-V2X technologies, providing a comprehensive insight into the efficiency of vehicular communication networks. The modeling of radio signal propagation was performed by a 3D simulation using ray tracing techniques [58]. These techniques allow detailed analysis of signal interaction with obstacles in the environment, such as buildings, vegetation, and other urban structures. The simulation was carried out on a digital map of an urban area, providing a realistic context for evaluating the performance of the LoRa network. Similarly, simulations for DSRC and C-V2X involved detailed analysis of signal reflections and refractions in densely built urban environments (see Figure 6).

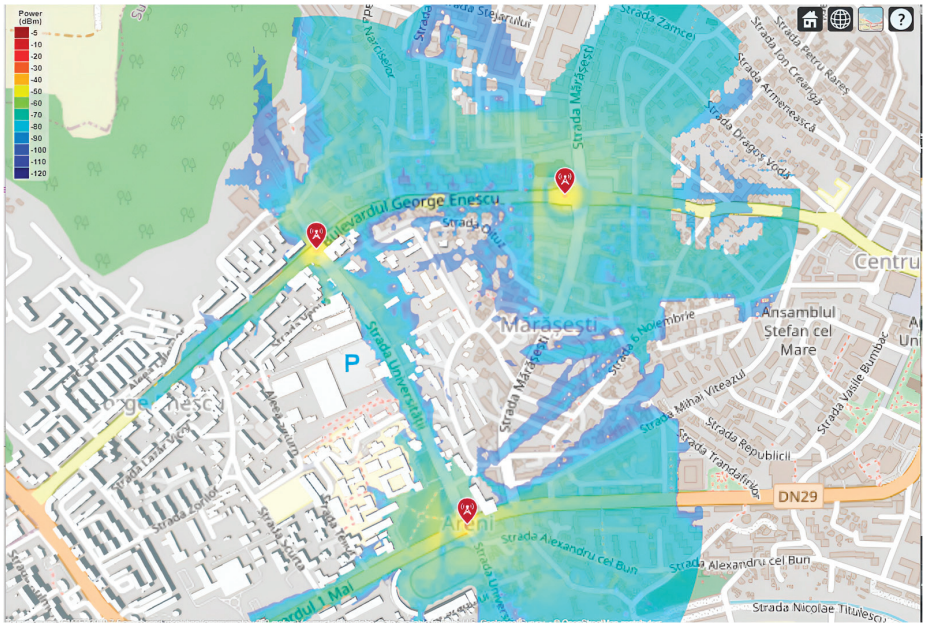


Figure 6. Representative area tested to expose the measurements performed by LoRa 2.4 GHz.

For this testing, LoRa modules, strategically placed in various locations in the study area, were used. The main parameters of the simulation include: 2.4 GHz signal frequency, transmission power according to LoRa standard specifications, propagation environments such as buildings, open spaces, green areas, and a ray tracing model that calculates the reflections, refractions, and diffractions of the radio signal. The results of this analysis provide a direct comparison with the performance observed in DSRC and C-V2X networks. While DSRC and C-V2X have demonstrated the ability to support critical vehicular communications with low latencies and high transmission rates, LoRa at 2.4 GHz has the advantage of extended coverage and low power consumption, albeit with lower throughput and higher latency. These differences emphasize the importance of choosing the appropriate technology depending on the specific application requirements.

Thus, by applying ray tracing modeling and 3D simulations, this chapter contributes to the understanding of how 2.4 GHz LoRa can be effectively integrated into vehicular and smart city communication infrastructure, complementing DSRC and C-V2X technologies to create a smart and interconnected transportation ecosystem.

The testing results are presented in the form of a three-dimensional visualization, which illustrates signal propagation within the complex urban environment. The generated image highlights the following key aspects:

- ✓ Areas of signal intensity are represented by a color range from blue to green. Regions colored in green indicate optimal signal coverage, while areas colored in blue signal low signal strength. This visual coding allows a quick and accurate assessment of network performance at different locations in the urban area.
- ✓ Buildings and other structures in the urban environment create shadows and reflections that adversely affect radio signal propagation. These phenomena are illustrated in the 3D visualization by chromatic variations in the vicinity of obstacles, highlighting how these structures disturb and reflect the signal. The analysis of these effects allows the identification of critical points where the signal is attenuated, thus helping to optimize the positioning of network nodes to ensure the most uniform and robust coverage.

This advanced visual approach provides a detailed insight into the performance of the LoRa network at 2.4 GHz, facilitating the understanding of the complexity of interactions between radio signals and the urban environment. By integrating these observations with previous theoretical and empirical analyses, we can develop effective strategies for deploying vehicular communication and sensor networks in the context of smart cities.

4. Discussion and Conclusions

In conclusion, this research has demonstrated the potential of LoRa technology at 2.4 GHz for smart city applications and vehicular communications. The study included a detailed signal propagation analysis using advanced ray tracing techniques, as well as performance evaluation in terms of bit error rate (BER), latency, and the ability to sustain packet handover.

While the paper highlights the many advantages of integrating DSRC, C-V2X, and LoRa 2.4 GHz technologies in intelligent transportation systems (ITS), it is important to recognize the limitations and prospects for further research. One of the main limitations is the interference in the 2.4 GHz band used by LoRa, which may affect the reliability of data transmission in dense urban environments. Also, variable latencies and transmission rates between different protocols may require further optimizations to ensure efficient interoperability.

Simulations and empirical measurements have shown that LoRa 2.4 GHz is capable of covering distances comparable to DSRC and C-V2X systems. The signal strength distribution showed robust coverage with strong signal areas in most of the tested locations. In complex urban areas, LoRa technology demonstrated excellent penetration capability and resistance to interference from obstacles such as buildings and vegetation.

A major advantage of using LoRa 2.4 GHz is the significantly lower cost compared to DSRC and C-V2X technologies. LoRa devices are more affordable and have much lower power consumption, making them ideal for large-scale deployments in urban and vehicular sensor networks. This cost-effectiveness enables the deployment of large networks without compromising performance. Performance analysis, including BER and latency evaluation, showed that LoRa 2.4 GHz could transmit messages quickly and efficiently while maintaining a high level of reliability. Although error rates increase at longer distances and in dense obstacle environments, the technology demonstrated robust packet recovery and retransmission capabilities, thus ensuring communication continuity. A promising direction for future research is the integration of LoRa technologies with DSRC and C-V2X to create hybrid vehicular communication networks. This integration could combine the advantages of each technology, thus providing extended coverage, low latency, and increased reliability. Future studies should focus on developing interoperability protocols and optimizing network resources to maximize the benefits of this synergy.

In order to validate and extend the findings of this research, it is necessary to carry out large-scale case studies and scalability tests in various urban and rural environments. These studies should include a wide range of usage scenarios and analyze the performance of the network under various conditions, including heavy traffic situations and the presence of significant obstacles. LoRa 2.4 GHz has enormous potential for integration into IoT and smart city infrastructures. Future research should focus on developing innovative solutions that leverage LoRa's extended coverage capacity and low power consumption to improve various aspects of urban management, such as traffic monitoring, waste management, and energy resource optimization.

Given the importance of vehicular communications and smart city applications, security and data protection are essential. Future research must address the development of advanced security mechanisms to protect the integrity and confidentiality of data transmitted over LoRa networks, thus ensuring a safe and reliable implementation.

By addressing these research directions, LoRa technology at 2.4 GHz can become a central pillar in the development of vehicular communication networks and smart city infrastructures, thus contributing to the creation of more efficient, safe, and connected urban environments.

In terms of future perspectives, future research should focus on improving resource management algorithms to optimize the combined use of DSRC, C-V2X, and LoRa. It would also be beneficial to develop mechanisms to dynamically adjust transmission parameters in real time to respond adaptively to changing network conditions and application requirements. Further studies are needed to assess the performance of these technologies in different traffic scenarios and to explore their potential in supporting fully autonomous vehicles and smart mobility infrastructures. These directions will help create more resilient and adaptable ITS solutions for the smart cities of the future.

Author Contributions: Conceptualization, E.Z. and Ș.H.; methodology, E.Z.; software, E.Z.; validation, E.Z. and Ș.H.; formal analysis, E.Z.; investigation, E.Z.; resources, E.Z.; data curation, Ș.H.; writing—original draft preparation, E.Z.; writing—review and editing, E.Z.; visualization, Ș.H.; supervision, E.Z.; project administration, E.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data are contained within the article.

Acknowledgments: This work is supported by NetZeRoCities Competence Center, funded by European Union—NextGenerationEU and Romanian Government, under National Recovery and Resilience Plan for Romania, contract no. 760007/30.12.2022 with Romanian Ministry Of Research, Innovation and Digitalisation, through specific research project P4—Smart Mobility and Infrastructure.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Road Traffic Injuries. Available online: https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries (accessed on 8 July 2024).
2. Decade of Action for Road Safety 2021–2030. Available online: https://www.who.int/teams/social-determinants-of-health/safety-and-mobility/decade-of-action-for-road-safety-2021-2030 (accessed on 8 July 2024).
3. Safety | US Department of Transportation. Available online: https://www.transportation.gov/priorities/safety (accessed on 8 July 2024).
4. Texts Adopted—EU Road Safety Policy Framework 2021–2030—Recommendations on Next Steps towards "Vision Zero"—Wednesday. 6 October 2021. Available online: https://www.europarl.europa.eu/doceo/document/TA-9-2021-0407_EN.html (accessed on 8 July 2024).
5. Wiseman, Y. Autonomous Vehicles. In *Encyclopedia of Information Science and Technology*, 5th ed.; IGI Global: Hershey, PA, USA, 2020; pp. 1–11; ISBN 9781799834809.
6. Valean, A.I. *EU Road Safety Policy Framework 2021–2030-Next Steps towards "Vision Zero"*; European Commission: Brussels, Belgium, 2019.
7. *IEEE Std 802.11-2020 (Revision of IEEE Std 802.11-2016)*; IEEE Standard for Information Technology--Telecommunications and Information Exchange between Systems-Local and Metropolitan Area Networks--Specific Requirements-Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications-Redline. 26 February 2021; pp. 1–7524.
8. *IEEE Std 802.11p-2010 (Amendment to IEEE Std 802.11-2007 as amended by IEEE Std 802.11k-2008, IEEE Std 802.11r-2008, IEEE Std 802.11y-2008, IEEE Std 802.11n-2009, and IEEE Std 802.11w-2009)*; IEEE Standard for Information Technology--Local and Metropolitan Area Networks--Specific Requirements--Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 6: Wireless Access in Vehicular Environments. 15 July 2010; pp. 1–51. [CrossRef]
9. Pérez, D.V.; Linnartz, J.P.M. Modelling message flows in Networked Lighting Control Systems using Trickle and CSMA. In Proceedings of the 2014 IEEE 21st Symposium on Communications and Vehicular Technology in the BeNeLux, Delft, The Netherlands, 10 November 2014; ISBN 9781479980307.
10. Miao, Y.; Dong, X.; Bourassa, M.A.; Zhu, D. Effects of Different Wave Spectra on Wind-Wave Induced Doppler Shift Estimates. In Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS), Waikoloa, HI, USA, 26 September 2020; Institute of Electrical and Electronics Engineers Inc.: New York City, NY, USA, 2020; pp. 5705–5708.
11. Rajput, N.S. Measurement of IEEE 802.11p Performance for Basic Safety Messages in Vehicular Communications. In Proceedings of the 2018 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS), Indore, India, 16–19 December 2018.
12. Yang, Z.; Wei, L. Logic Circuit and Memory Design for In-Memory Computing Applications using Bipolar RRAMs. In Proceedings of the 2019 IEEE International Symposium on Circuits and Systems (ISCAS), Sapporo, Japan, 26–29 May 2019; ISBN 9781728103976.
13. Cao, M.; Shuai, Q.; Li, V.O. Emergency Vehicle-Centered Traffic Signal Control in Intelligent Transportation Systems. In Proceedings of the 2019 IEEE Intelligent Transportation Systems Conference—ITSC, Auckland, New Zealand, 27–30 October 2019; ISBN 9781538670248.
14. Najm, W.G.; Koopmann, J.; Smith, J.D.; Brewer, J. *Frequency of Target Crashes for IntelliDrive Safety Systems*; National Highway Traffic Safety Administration: Washington, DC, USA, 2010.
15. European Transport Policy. Europe on the Move—Volume 1 | EESC. Available online: https://www.eesc.europa.eu/en/our-work/publications-other-work/publications/european-transport-policy-europe-move-volume-1 (accessed on 8 July 2024).
16. Kang, Y.; Tian, C. Study on the Emission of Connected Autonomous Vehicle Considering the Control of Electronic Throttle Opening. *World Electr. Veh. J.* **2024**, *15*, 233. [CrossRef]
17. Neugebauer, M.; Żebrowski, A.; Esmer, O. Cumulative Emissions of CO_2 for Electric and Combustion Cars: A Case Study on Specific Models. *Energies* **2022**, *15*, 2703. [CrossRef]
18. Wittmann, M.; Neuner, L.; Lienkamp, M. A Predictive Fleet Management Strategy for on-Demand Mobility Services: A Case Study in Munich. *Electronics* **2020**, *9*, 1021. [CrossRef]
19. Mertens, J.C.; Knies, C.; Diermeyer, F.; Escherle, S.; Kraus, S. The Need for Cooperative Automated Driving. *Electronics* **2020**, *9*, 754. [CrossRef]
20. Wang, Q.; Zhang, R.; Wang, Y.; Lv, S. Machine Learning-Based Driving Style Identification of Truck Drivers in Open-Pit Mines. *Electronics* **2020**, *9*, 19. [CrossRef]
21. Zhang, T.; Zhou, W.; Meng, F.; Li, Z. Efficiency Analysis and Improvement of an Intelligent Transportation System for the Application in Greenhouse. *Electronics* **2019**, *8*, 946. [CrossRef]
22. Wu, Y.; Lee, W.W.; Gong, X.; Wang, H. A Hybrid Intrusion Detection Model Combining SAE with Kernel Approximation in Internet of Things. *Sensors* **2020**, *20*, 5710. [CrossRef] [PubMed]
23. Li, Y.; Liu, W.; Liu, Q.; Zheng, X.; Sun, K.; Huang, C. Complying with ISO 26262 and ISO/SAE 21434: A Safety and Security Co-Analysis Method for Intelligent Connected Vehicle. *Sensors* **2024**, *24*, 1848. [CrossRef] [PubMed]
24. Arena, F.; Pau, G.; Severino, A. A Review on IEEE 802.11p for Intelligent Transportation Systems. *J. Sens. Actuator Netw.* **2020**, *9*, 22. [CrossRef]
25. Iriqat, S.; Yenikaya, S.; Secmen, M. Dual-Band 2 × 1 Monopole Antenna Array and Its MIMO Configuration for WiMAX, Sub-6 GHz, and Sub-7 GHz Applications. *Electronics* **2024**, *13*, 1502. [CrossRef]

26. *IEEE Std 802.11a-1999*; IEEE Standard for Telecommunications and Information Exchange between Systems-LAN/MAN Specific Requirements-Part 11: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications: High Speed Physical Layer in the 5 GHz Band. 30 December 1999; pp. 1–102. [CrossRef]
27. Souri, A.; Zarei, M.; Hemmati, A.; Gao, M. A Systematic Literature Review of Vehicular Connectivity and V2X Communications: Technical Aspects and New Challenges. *Int. J. Commun. Syst.* **2024**, *37*, e5780. [CrossRef]
28. Batalla, J.M.; Sujecki, S.; Oko, J.; Kelner, J.M. Cost-Effective Measurements of 5G Radio Resources Allocation for Telecom Market Regulator's Monitoring. In Proceedings of the PE-WASUN 2022—Proceedings of the 19th ACM International Symposium on Performance Evaluation of Wireless Ad Hoc, Sensor, and Ubiquitous Networks, Montreal, QC, Canada, 24–28 October 2022; pp. 83–90. [CrossRef]
29. Garrott, W.R. The anticipated effect of improvements to automotive electronics on safety during the years 2009 through 2015. In Proceedings of the 2008 IEEE International Conference on Vehicular Electronics and Safety, Columbus, OH, USA, 22–24 September 2008; IEEE Xplore: Piscataway, NJ, USA, 2008. ISBN 9781424423606.
30. Zhang, Y.; Zhang, J. Design and Optimization of Cluster-Based DSRC and C-V2X Hybrid Routing. *Appl. Sci.* **2022**, *12*, 6782. [CrossRef]
31. Petrov, T.; Sevcik, L.; Pocta, P.; Dado, M. A Performance Benchmark for Dedicated Short-Range Communications and Lte-Based Cellular-V2x in the Context of Vehicle-to-Infrastructure Communication and Urban Scenarios. *Sensors* **2021**, *21*, 5095. [CrossRef] [PubMed]
32. *IEEE Std 802.11bd-2022 (Amendment to IEEE Std 802.11-2020 as amended by IEEE Std 802.11ax-2021, IEEE Std 802.11ay-2021, IEEE Std 802.11ba-2021, IEEE Std 802.11-2020/Cor 1-2022, and IEEE Std 802.11az-2022)*; IEEE Standard for Information Technology--Telecommunications and Information Exchange between Systems Local and Metropolitan Area Networks--Specific Requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 5: Enhancements for Next Generation V2X. 10 March 2023; pp. 1–144. [CrossRef]
33. Nguyen, H.; Noor-A-Rahim, M.; Liu, Z.; Jamaludin, D.; Guan, Y.L. A Semi-Empirical Performance Study of Two-Hop DSRC Message Relaying at Road Intersections. *Information* **2018**, *9*, 147. [CrossRef]
34. Albattah, W.; Habib, S.; Alsharekh, M.F.; Islam, M.; Albahli, S.; Dewi, D.A. An Overview of the Current Challenges, Trends, and Protocols in the Field of Vehicular Communication. *Electronics* **2022**, *11*, 3581. [CrossRef]
35. Zhao, P.; Wang, X.; Zhang, K.; Jin, Y.; Zheng, G. Doppler Modeling and Simulation of Train-to-Train Communication in Metro Tunnel Environment. *Sensors* **2022**, *22*, 4289. [CrossRef] [PubMed]
36. Choi, J.; Ahn, S. A 5.9 GHz Channel Characterization at Railroad Crossings for Train-to-Infrastructure (T2I) Communications. *Electronics* **2023**, *12*, 2400. [CrossRef]
37. Wang, T.; Hussain, A.; Cao, Y.; Gulomjon, S. An Improved Channel Estimation Technique for Ieee 802.11p Standard in Vehicular Communications. *Sensors* **2019**, *19*, 98. [CrossRef] [PubMed]
38. ICICI-BME 4. 2015 Bandung; Bandung Institute of Technology; IEEE International Conference on Service-Oriented Computing and Applications; International Conference on Instrumentation, C.; ICICI-BME 4 2015.11.02-03 Bandung. In Proceedings of the 2015 4th International Conference on Instrumentation, Communications, Information Technology, and Biomedical Engineering (ICICI-BME), Bandung, Indonesia, 2–3 November 2015; ISBN 9781467378000.
39. Khan, U.A.; Lee, S.S. Multi-Layer Problems and Solutions in VANETs: A Review. *Electronics* **2019**, *8*, 204. [CrossRef]
40. Hota, L.; Nayak, B.P.; Kumar, A.; Sahoo, B.; Ali, G.G.M.N. A Performance Analysis of VANETs Propagation Models and Routing Protocols. *Sustainability* **2022**, *14*, 1379. [CrossRef]
41. Augustin, A.; Yi, J.; Clausen, T.; Townsley, W.M. A Study of Lora: Long Range & Low Power Networks for the Internet of Things. *Sensors* **2016**, *16*, 1466. [CrossRef]
42. Lima, W.G.; Lopes, A.V.R.; Cardoso, C.M.M.; Araújo, J.P.L.; Neto, M.C.A.; Tostes, M.E.L.; Nascimento, A.A.; Rodriguez, M.; Barros, F.J.B. LoRa Technology Propagation Models for IoT Network Planning in the Amazon Regions. *Sensors* **2024**, *24*, 1621. [CrossRef]
43. Zhang, Z.; Cao, S.; Wang, Y. A Long-Range 2.4G Network System and Scheduling Scheme for Aquatic Environmental Monitoring. *Electronics* **2019**, *8*, 909. [CrossRef]
44. Zhang, Z.; Zhou, C.; Sheng, L.; Cao, S. Optimization Schemes for UAV Data Collection with LoRa 2.4 GHz Technology in Remote Areas without Infrastructure. *Drones* **2022**, *6*, 173. [CrossRef]
45. Kane, L.; Liu, V.; McKague, M.; Walker, G. An Experimental Field Comparison of Wi-Fi HaLow and LoRa for the Smart Grid. *Sensors* **2023**, *23*, 7409. [CrossRef]
46. Janssen, T.; Bnilam, N.; Aernouts, M.; Berkvens, R.; Weyn, M. Lora 2.4 Ghz Communication Link and Range. *Sensors* **2020**, *20*, 4366. [CrossRef] [PubMed]
47. Havriliuc, S.; Coca, E.; Lavric, A.; Popa, V.; Petrariu, A.I. Expanding IoT Connectivity by Using 2.4 GHz LoRa Technology. In Proceedings of the 2024 17th International Conference on Development and Application Systems, DAS, Suceava, Romania, 23–25 May 2024; Institute of Electrical and Electronics Engineers Inc.: New York City, NY, USA, 2024; pp. 87–90.
48. Ragnoli, M.; Colaiuda, D.; Leoni, A.; Ferri, G.; Barile, G.; Rotilio, M.; Laurini, E.; De Berardinis, P.; Stornelli, V. A LoRaWAN Multi-Technological Architecture for Construction Site Monitoring. *Sensors* **2022**, *22*, 8685. [CrossRef]
49. Guo, Q.; Yang, F.; Wei, J. Experimental Evaluation of the Packet Reception Performance of LoRa. *Sensors* **2021**, *21*, 1071. [CrossRef]

50. Zadobrischi, E.; Dimian, M.; Negru, M. The Utility of Dsrc and V2x in Road Safety Applications and Intelligent Parking: Similarities, Differences, and the Future of Vehicular Communication. *Sensors* **2021**, *21*, 7237. [CrossRef]
51. Miao, L.; Virtusio, J.J.; Hua, K.L. Pc5-Based Cellular-V2x Evolution and Deployment. *Sensors* **2021**, *21*, 843. [CrossRef]
52. Cui, G.; Zhang, W.; Xiao, Y.; Yao, L.; Fang, Z. Cooperative Perception Technology of Autonomous Driving in the Internet of Vehicles Environment: A Review. *Sensors* **2022**, *22*, 5535. [CrossRef] [PubMed]
53. Zadobrischi, E.; Dimian, M. Inter-Urban Analysis of Pedestrian and Drivers through a Vehicular Network Based on Hybrid Communications Embedded in a Portable Car System and Advanced Image Processing Technologies. *Remote Sens.* **2021**, *13*, 1234. [CrossRef]
54. Wiseman, Y. Autonomous Vehicles Will Spur Moving Budget from Railroads to Roads. *Int. J. Intell. Unmanned Syst.* **2024**, *12*, 19–31. [CrossRef]
55. Colombi, D.; Joshi, P.; Xu, B.; Ghasemifard, F.; Narasaraju, V.; Törnevik, C. Analysis of the Actual Power and EMF Exposure from Base Stations in a Commercial 5G Network. *Appl. Sci.* **2020**, *10*, 5280. [CrossRef]
56. Saari, M.; Muzaffar Bin Baharudin, A.; Sillberg, P.; Hyrynsalmi, S.; Yan, W. LoRa-A Survey of Recent Research Trends. In Proceedings of the 2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia, 21–25 May 2018.
57. Brahim, A.; Khedher, L.; Górriz, J.M.; Ramírez, J.; Toumi, H.; Lespessailles, E.; Jennane, R.; El Hassouni, M. A proposed computer-aided diagnosis system for Parkinson's disease classification using 123I-FP-CIT imaging. In Proceedings of the 2017 International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), Fez, Morocco, 22–24 May 2017; ISBN 9781538605516.
58. Yun, Z.; Iskander, M.F. Ray Tracing for Radio Propagation Modeling: Principles and Applications. *IEEE Access* **2015**, *3*, 1089–1100. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

electronics

MDPI

Article

An Architecture of Enhanced Profiling Assurance for IoT Networks

Nut Aroon *, Vicky Liu, Luke Kane, Yuefeng Li, Aklilu Daniel Tesfamicael and Matthew McKague

Faculty of Science, Queensland University of Technology, Brisbane, QLD 4000, Australia; v.liu@qut.edu.au (V.L.); le.kane@qut.edu.au (L.K.); y2.li@qut.edu.au (Y.L.); aklilu.tesfamicael@qut.edu.au (A.D.T.); matthew.mckague@qut.edu.au (M.M.)

* Correspondence: nut.aroon@hdr.qut.edu.au

Abstract: Attacks launched from IoT networks can cause significant damage to critical network systems and services. IoT networks may contain a large volume of devices. Protecting these devices from being abused to launch traffic amplification attacks is critical. The manufacturer usage description (MUD) architecture uses pre-defined stateless access control rules to allow or block specific network traffic without stateful communication inspection. This can lead to false negative filtering of malicious traffic, as the MUD architecture does not include the monitoring of communication states to determine which connections to allow through. This study presents a novel solution, the enhanced profiling assurance (EPA) architecture. It incorporates both stateless and stateful communication inspection, a unique approach that enhances the detection effectiveness of the MUD architecture. EPA contains layered intrusion detection and prevention systems to monitor stateful and stateless communication. It adopts three-way decision theory with three outcomes: allow, deny, and uncertain. Packets that are marked as uncertain must be continuously monitored to determine access permission. Our analysis, conducted with two network scenarios, demonstrates the superiority of the EPA over the MUD architecture in detecting malicious activities.

Keywords: manufacturer usage description; network behaviour analysis; three-way decision theory; stateful inspection; malicious behaviour detection

1. Introduction

Protecting IoT networks from traffic amplification attacks is challenging due to a large volume of devices as attacks cause disruption on the networks. A prominent example is the attack on the domain name service provider DYN, conducted by infecting IoT devices with Mirai malware and using them to attack the DYN central server [1,2]. This attack disrupted many services as they were unable to process requests from users [3]. Recently, Cloudflare published a report on traffic amplification attacks for the first quarter of 2024 [4]. The report showed a 32% increase in attacks compared with the previous year.

The Internet Engineering Task Force (IETF) approved the RFC8520 as a standard for protecting IoT networks [5]. It defines the manufacturer usage description (MUD) architecture for protecting IoT devices from being abused as part of traffic amplification attacks. The MUD architecture uses a device behaviour profile provided by manufacturers or related entities, translated into access control list (ACL) rules, to filter out unwanted or malicious traffic.

Despite the promise of the MUD architecture, it still has limitations that should be addressed to improve the effectiveness of malicious activity detection. The MUD architecture uses pre-defined stateless ACL rules to allow specific communications. This can result in a false negative filter, as MUD architecture does not monitor communication states during a session. The MUD architecture only allows or denies network packets based on pre-defined ACL rules. However, uncertain network behaviours could occur in the traffic, requiring continuous analysis before allowing or rejecting them.

1.1. Objective

The existing limitations of the MUD architecture can cause false negative filtering, decreasing the effectiveness of network communication enforcement. This study proposed the enhanced profiling assurance (EPA) architecture to improve the effectiveness of malicious activity detection. This is achieved by incorporating stateless and stateful inspections to monitor communication states. Three-way decision theory has been adopted with allow, deny, and uncommitted permission to further analyse the uncertain network behaviours.

1.2. Scope

The MUD architecture consists of three main aspects: MUD profile format, MUD profile generation and retrieval, and MUD deployment. This study focuses on the following:

- improvement to the effectiveness of malicious activity detection over the current MUD standard;
- IP-based IoT networks;
- MUD deployment in small-scale networks.

This study will not focus on the following aspects:

- mitigation of other than traffic amplification attacks;
- software-defined network (SDN)-based MUD deployments as these are not found in small-scale networks;
- authentication/authenticity of MUD files;
- MUD profile format/MUD profile generation and retrieval;
- non-IP-based networks;
- large-scale IoT networks.

1.3. Research Method

Our study begins by studying the MUD architecture and MUD file data model to understand the operation of network traffic enforcement. This study then continues investigating recent works on MUD deployment and MUD usage for network traffic analysis. The limitations of the MUD architecture and existing works are then identified to detect the research gap. Based on the identified limitations, EPA is proposed to address the limitations found in the MUD architecture and existing works. The EPA is then analysed in various network communication scenarios to determine the effectiveness of malicious activity detection. This study then concludes with a summary discussion of the proposed EPA and outlines future work to implement the EPA.

1.4. Contribution

The security of the IoT network would be at risk of traffic amplification attacks as the current MUD architecture enforcement does not monitor communication states, leading to false negative filtering. This study is significant to IoT network security as it enhances the MUD architecture by improving the effectiveness of malicious activity detection and reducing false negative filtering. Combining stateless and stateful inspection with three-way decision theory is a unique approach to the MUD architecture, which allows further analysis of the communication states to detect malicious activity.

The contributions of this paper can be expressed as follows:

- proposes a new gateway-based MUD architecture that enables stateless and stateful communication inspection to improve the effectiveness of malicious activity detection;
- employs layered intrusion detection and prevention systems (IDPSs) to improve the effectiveness of malicious activity detection;
- introduces the network behaviour analysis (NBA) system with network behaviour knowledge databases to conduct a stateful inspection of the network communication states;
- presents three-way decision theory with allow, deny, and uncommitted to allow continuous monitoring for uncertain network behaviours.

1.5. Paper Structure

The remainder of this paper is structured as follows: Section 2 provides insight into the background of the MUD architecture and data model. Section 3 reviews the existing works on gateway-based MUD architecture and the usage of MUD for traffic to identify the research gap. Section 4 proposes the architectural design of the EPA, workflow, and network behaviour analysis and decision-making. Section 5 evaluates the different scenarios to determine the effectiveness of malicious activity detection. Finally, Section 6 concludes this study and outlines future work.

2. Background

2.1. Manufacturer Usage Description (MUD)

MUD is a standard defined in RFC8520 [5] for protecting IoT devices from being abused by malicious activities. It enforces communication using pre-defined ACL rules translated from the device network behaviour profile to monitor and filter the traffic. Figure 1 shows an overview of the MUD communication enforcement process.

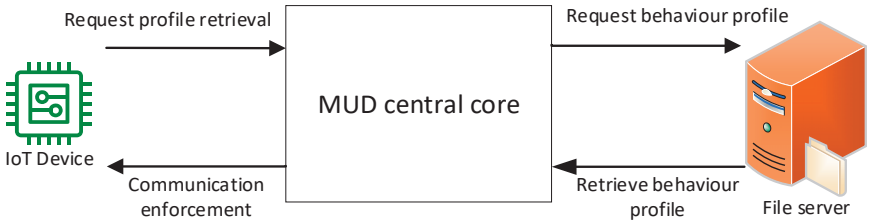


Figure 1. Overview of MUD communication enforcement process.

When the IoT device connects to the MUD-based network for the first time, it sends the request of its behaviour profile to the MUD central core to retrieve it from the hosted file server. The behaviour profile is then translated into ACL rules to allow or block the specific communication.

2.2. MUD Architecture

The MUD architecture describes the interaction between each component to receive the behaviour profile address (MUD URL) and retrieve the behaviour profile (MUD file) for traffic filtering. The diagram of the MUD architecture is shown in Figure 2.

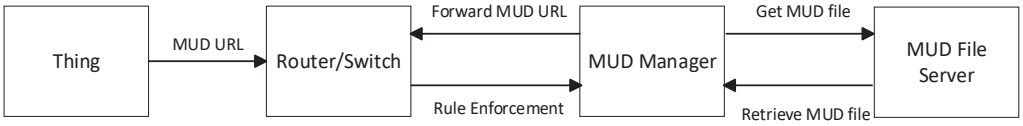


Figure 2. MUD architecture and interaction between each component.

Details of each component can be expressed as follows:

- Thing: the IoT device that transmits the MUD URL for MUD file retrieval.
- Router/Switch: provides the internet connection to the Thing.
- MUD Manager: the central core of MUD architecture. It retrieves and translates the MUD file into ACL rules for communication enforcement.
- MUD File Server: the file server that hosts the MUD file by the manufacturer or related entities.

MUD Manager functions for communication enforcement are as follows:

- extract the MUD URL from the requested device for MUD file retrieval;
- retrieve the MUD file from the hosted MUD file server;
- translate the MUD file into the ACL rules for communication enforcement at the IDPS;
- maintain and update the behaviour profile and network configuration.

Regarding the MUD file retrieval process, the IETF suggested using the HTTPS protocol to retrieve the MUD file securely. The IoT device must transmit the MUD URL to the MUD manager to extract the MUD URL and retrieve the MUD file.

This study describes the operation of the MUD architecture, as illustrated in Figure 3, to provide more details of the MUD enforcement process.

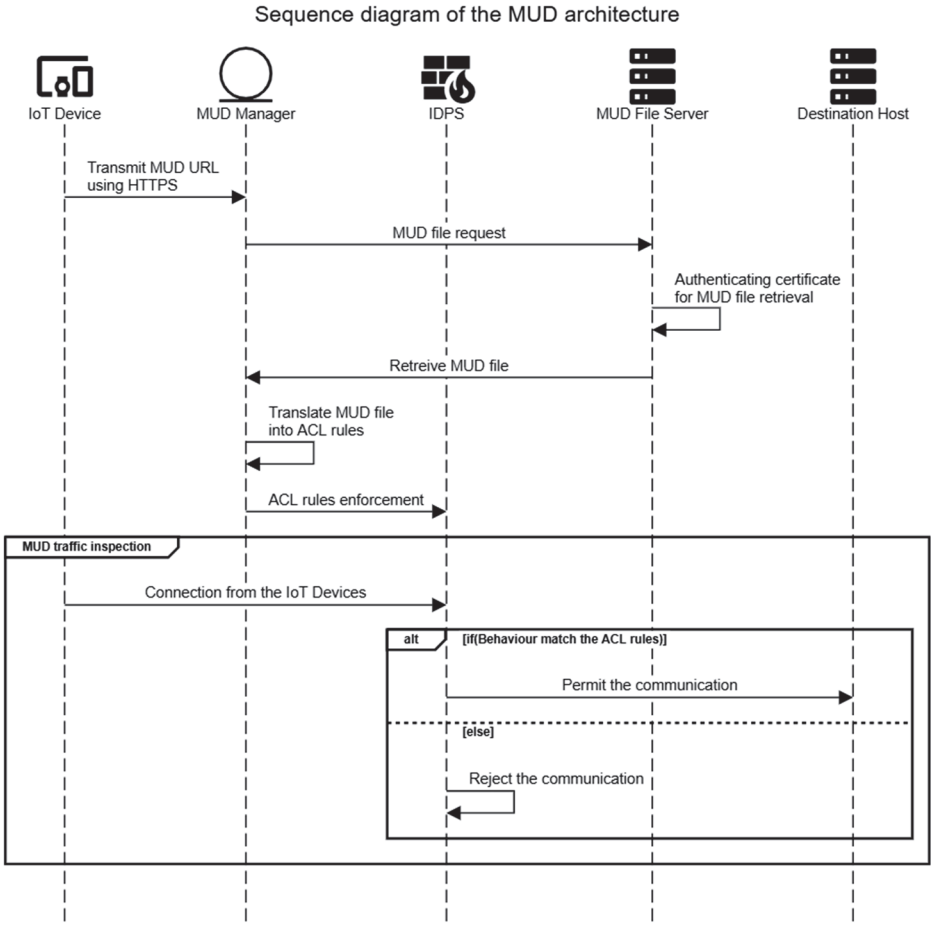


Figure 3. Sequence diagram of the MUD architecture.

The process of network enforcement in the MUD architecture can be expressed as follows:

1. In the first connection, the IoT device transmits the MUD URL to the MUD manager for the MUD profile request.
2. The MUD manager extracts the MUD URL for MUD file retrieval and then requests the MUD file from the hosted file server.
3. MUD file server authenticating certificate to ensure the authenticity of MUD file request.
4. The MUD manager retrieves the MUD file and translates it into the ACL rules to enforce at the IDPS.
5. Enforcing ACL rules at the IDPS for traffic monitoring.
6. The IDPS receives new connections from the IoT device and conducts stateless inspection based on the ACL rules. The IDPS will permit the communication if the behaviour matches the ACL rules and reject it if it does not.

Our study categorises the deployments of MUD managers as the following:

- software define network (SDN)-based MUD deployment;
- gateway-based MUD deployment.

Our study will only focus on gateway-based MUD deployment, as our study intends to propose an alternative approach to MUD deployment for securing small-scale networks.

2.3. MUD Behaviour Profile

RFC8520 proposed using the network management data modelling language Yet Another Next Generation (YANG) [6] for specifying the intended communication behaviour in the MUD file. The YANG module allows the behaviour profile to be translated into the configuration parameters to configure the IDPS for traffic monitoring.

The MUD file contains two main containers, the MUD and ACL containers. The MUD container is used for MUD file validation, which includes the information related to the MUD file and the access list names. The ACL container contains the intended network behaviours for communication enforcement at the IDPS. Details of the device behaviour information in the ACL containers are as follows:

- Name: Name of the ACL and access control entry (ACE);
- Type: Connection type, either IPv4 or IPv6;
- Protocol: Protocol number defined by the Internet Assigned Numbers Authority (IANA) [7];
- Destination (dst): The address of the destination host to which it is to be connected; Port: Port number usage of each connectivity. Depending on the manufacturer, this can be a well-known port assigned by IANA [8] or a custom port;
- Direction-initiated: Description of connection initiation direction for TCP-based communication. It can be "from-device" or "to-device";
- Forwarding: Denotes that it will either allow or deny this communication behaviour.

The example of behaviour information in the ACL container of the MUD file is shown in Figure 4.

```
}, {
  "name" : "from-ipv4-amazon_echo-2",
  "matches" : {
    "ipv4" : {
      "protocol" : 6,
      "ietf-acldns:dst-dnsname" : "softwareupdates.amazon.com"
    },
    "tcp" : {
      "destination-port" : {
        "operator" : "eq",
        "port" : 443
      },
      "ietf-mud:direction-initiated" : "from-device"
    }
  },
  "actions" : {
    "forwarding" : "accept"
  }
}, {
```

Figure 4. Code snippet of device communication behaviour.

This code snippet describes the behaviour information of the ACE "from-ipv4-amazon _echo-2". It specifies that it allows the outbound HTTPS (port 443) to a server (softwareupdates.amazon.com) from the device. If the IDPS detects a matched network packet, it will allow a packet to reach the destination.

3. Related Works

3.1. MUD Impact on IoT Security and Supporters

The introduction of MUD architecture provides a means for reducing risk from malicious activity. Heeb et al. [9] analysed the impact of MUD on IoT security in different attack scenarios. Their analysis showed that the MUD architecture can reduce attacks on the transport layer, especially traffic amplification-based attacks. However, their analysis has yet to point out the limitation of stateless inspection in MUD that can cause false negative filtering.

Due to being a standard for IoT security, MUD has gained support from organisations such as the National Institute of Standards and Technology (NIST). They demonstrated the

deployment of MUD-based networks and published guideline documents based on their demonstration for securing home and small business networks [10]. NIST also developed MUD-PD, a companion tool for MUD profile generation, which can be accessed at [11].

Cisco plays a vital role in supporting MUD usage [12]. Cisco has proposed its MUD architecture with the FreeRADIUS-based Identity Services Engine (ISE) for MUD URL authentication [13,14]. Cisco has also developed the Cisco MUD manager as open-source software for its proposed architecture, which can be accessed at [15].

The University of New South Wales (UNSW) IoT Traffic Analytics Research Group also contributed to MUD usage. The team has developed the MUD profile generation tool, MUDGEE [16], which can be accessed at [17]. The team also published the MUD profiles dataset generated from devices in the testbed network [18], which can be accessed at [19].

3.2. Gateway-Based MUD Architecture

The gateway-based MUD network architecture provides an alternative approach for deploying MUD to secure IoT networks. It has the benefit of lower deployment costs when compared to the SDN-based network deployment cost, which makes it suitable for small-scale networks such as smart homes.

To achieve this, the osMUD [20] has been developed for the OpenWRT-based router [21] with the customised version of Dnsmasq [22] for DNS and DHCP services. It was developed for deployment on a network gateway. Currently, the osMUD only supports stateless communication inspection without considering communication states.

Andalibi et al. [23] proposed deploying the MUD manager in fog computing. Two components, the Fog MUD-Manager and peak request rate, have been introduced. The Fog MUD-Manager retrieves the MUD file and enforces at the IDPS to filter communication from local networks. The peak request rate is the parameter used to limit the maximum communication rate with the Fog MUD-Manager. This research only focuses on stateless inspection without considering communication state inspection.

Corno and Manella [24] extended the gateway-based MUD architecture to enhance smart home security by proposing a plug-in for open-source home automation, Home Assistant [25]. The plug-in allows developers to create network behaviour policies for automatic MUD profile generation at local networks. This could benefit legacy devices that cannot transmit the MUD URL or end-of-life devices. However, their research only concerns MUD profile generation and relies on stateless inspection without considering communication state monitoring.

Feraudo et al. [26] introduced the traffic amplification attack mitigation system using MUD and extended the Berkeley packet filter (eBPF)-based traffic filtering. Their research proposed rate limit extension, packet rate, and byte rate to define the communication rate threshold in the MUD profile. Additionally, they implemented the eBPF-based packet filtering as a backend service for traffic filtering. This research only uses a stateless communication inspection without monitoring communication states.

Sajjad et al. [27] proposed enhancing the MUD architecture for vulnerability detection in home routers and connected devices by adopting the OWASP firmware security testing methodology [28]. The vulnerability information is shared with stakeholders via blockchain smart contracts to notify security patching if detected. The vulnerability information also calculates the overall risk for MUD profile retrieval authorisation. Their proposal only focuses on vulnerability detection rather than stateful communication inspection for monitoring communication states.

Feraudo et al. [29] proposed Colearn, the architecture that implemented the osMUD and Federated Learning (FL) framework, PySyft [30], to protect MQTT-based IoT networks. The MUD manager fetches the MUD file from the local User Policies Server (UPS), which is included by network administrators, to assist the FL coordinator in excluding unauthorised devices. Their research focuses on the device connection authorisation and does not concern the stateful communication inspection to monitor communication states.

Datta et al. [31] proposed iDAM, the two-level defence MUD architecture for mitigating volumetric-based attacks. The first level is at the iDAM gateway, which permits MUD-compliant traffic and detects attacks using federated learning-based machine learning. The iDAM manager acts as a second level of protection to enforce and detect anomalies in gateways while maintaining and aggregating the anomaly detection model. This research has not considered stateful traffic inspection for continuously monitoring communication states, which could lead to false negative filtering.

Nisa et al. [32] integrated the blockchain into a gateway-based MUD network for intelligence sensor protection in smart city networks. Their architecture conducts penetration testing to detect vulnerabilities and share information with other gateways via blockchain. Their research concerns vulnerability detection without considering network behaviour analysis.

Afek et al. [33] introduced MUDirect, the MUD architecture for peer-to-peer (P2P) IoT networks. MUDirect contains three main components: a tracking application for tracking IP address changes on mobile devices, a MUDirect service for mapping the P2P connection, and a MUD manager for extracting P2P communication information. The MUD profile has been extended to support the definition of a secondary virtual manufacturer (SVM) for managing P2P communication. This research only focuses on MUD deployment for P2P IoT networks without considering stateful inspection for malicious activity detection in the network packets.

A summary of the literature review on gateway-based MUD deployments is shown in Table 1.

Table 1. Summary of the literature review on gateway-based MUD deployments.

Reference	Application	Inspection Techniques	Description
[23]	Unspecified	Stateless	MUD manager deployment on fog computing.
[24]	Smart home	Stateless	Local behaviour policies and MUD profile generation.
[26]	Unspecified	Stateless	Rate limit extension and backend deployment for packet filtering.
[27]	Smart home	-	Firmware testing and vulnerability detection on the gateway.
[29]	-	Stateless	Focus on employing federated learning for authorising device connection.
[31]	Smart home	Stateless	Employing federated learning with two-level defence.
[32]	Smart cities	-	Firmware testing and vulnerability detection on the gateway.
[33]	Smart home	Stateless	Extending MUD profile to support P2P networks.

3.3. Network Traffic Analysis in MUD-Based Networks

Aside from deploying the MUD to protect the IoT networks from malicious activities, it has also been applied for traffic analysis applications. Hadi et al. [34] presented the MUD-based botnet detection and mitigation system for home routers. They have implemented Snort [35] for analysing MUD-rejected traffic to identify the attack source and notify the user via email. This research only focuses on the analysis of MUD-rejected traffic.

Zangrandi et al. [36] developed MUDScope [37] to analyse and detect anomalies in the rejected traffic. The system can identify four types of attacks: Telnet/SSH port scan, OS scan, vulnerability scan, and TCP SYN flood DDoS. The authors also published the produced dataset, which can be accessed at [38]. Their research only focuses on analysing MUD-rejected traffic.

Andalibi et al. [39] presented MUD-Visualizer, the MUD traffic visualisation tool for displaying MUD communication behaviours. The tool receives sets of MUD files, merges sets of ACEs, and removes redundancy information to visualise MUD traffic information. MUD-Visualizer could assist in MUD deployment by graphically displaying the MUD traffic information, but it is not concerned with traffic inspection for malicious activity detection. The tool can be accessed at [40].

Bremler-Barr et al. [41] proposed MUDIS, a tool for inspecting the similarity of two MUD files and merging them into a single MUD file. The tool will detect the similarity of network behaviours in two MUD files. Then, the remaining unmatched behaviour profiles are clustered into two groups before being merged into a new MUD file. The authors

have implemented MUDIS in [42] to demonstrate the impact of location on MUD profile generation. Their research does not focus on malicious activity detection.

A summary of the literature review on traffic analysis in MUD-based networks is shown in Table 2.

Table 2. Summary of the literature review on traffic analysis in MUD-based networks.

Reference	Application Usage	Description
[34]	Smart home protection	MUD-rejected traffic analysis for attack source detection and notification.
[36]	Attack type identification	MUD-rejected traffic analysis to identify attack type.
[39]	Visualising MUD traffic	Development of MUD traffic visualiser tool.
[41]	Behaviour profile inspection and merging	Development of MUD behaviour profile analysis and profile merging tool, MUDIS.
[42]	Impact of location on MUD profile generation demonstration	Implemented MUDIS from [41] to demonstrate the impact of location on MUD profile generation.

3.4. Three-Way Decision Theory for Intrusion Detection

As the attack patterns have evolved, they have become more sophisticated and challenging to distinguish. Some network behaviours might require further investigation to determine the actual type of behaviour. The implementation of three-way decision theory [43–46] has gained interest as it can support allowing, denying, and uncommitted decisions to further analyse uncertain network behaviours before making a decision.

Shen [47] and Du et al. [48] introduced the usage of a deep belief network (DBN) with three-way decision theory for intrusion detection. DBN has been adopted to extract features from pre-processed data for data classification using three-way decision theory to detect the attack. In case of an uncommitted decision, the authors proposed using a K-nearest neighbour (KNN) to reclassify until no data remains. Shen [47] experimented with the switch flow tables and control domain bandwidth attack detection on SDN. Du et al. [48], on the other hand, evaluated with KDD Cup 99 and NSL-KDD datasets. Their research did not focus on the real-time monitoring of IoT networks.

Zhang et al. [49] and Geng et al. [50] proposed the autoencoder and three-way decision approach for intrusion detection. The character data were pre-processed before being denoised by an autoencoder to construct the multi-granular feature structure. The difference is that Zhang et al. [49] applied using three-way decision while Geng et al. [50] implemented with sequential three-way decision. Both research studies conducted an evaluation with the NSL-KDD dataset but it is yet to be implemented with real-time IoT network monitoring.

Zhang et al. [51] presented the three-branch random forest model for intrusion detection. Each attribute from the training dataset was extracted to calculate the attribute importance before being divided into three domains by three-way decision theory. Then, the model will randomly select the attributes from each domain to train the decision tree to form the random forest for intrusion detection. The proposed model was evaluated with the NSL-KDD dataset. This research is not concerned with real-time IoT communication monitoring.

A summary of the literature review on three-way decision theory for intrusion detection is shown in Table 3.

To identify the research gaps in prior studies of MUD architecture and MUD-based traffic analysis, a summary of MUD architecture and MUD-based traffic analysis is listed as follows:

Heeb et al. [9] discussed the impact of MUD on IoT security by describing its benefits in various threat scenarios. However, they have yet to discuss the limitation of stateless communication in the MUD architecture.

Multiple organisations [10–13,15,16,18] have shown support for MUD usage. NIST [10,11] only focuses on the demonstration of MUD deployment. Cisco [12,13,15] proposed its

MUD architecture but still relies on stateless communication inspection. UNSW [16,18] only focuses on the MUD profile generation.

Table 3. Summary of the literature review on three-way decision theory for intrusion detection.

Reference	Techniques	Evaluation	Description
[47]	DBN, Three-way decision, and KNN	Real SDN traffic	Integrated three-way decision with DBN for intrusion detection in SDN.
[48]	DBN, Three-way decision, and KNN	KDD Cup 99 and NSL-KDD	Integrated three-way decision with DBN for intrusion detection.
[49]	Autoencoder and Three-way decision	NSL-KDD	Using autoencoder for denoising before classification for intrusion detection.
[50]	Autoencoder and Sequential three-way decision	NSL-KDD	Using autoencoder to extract multi-granularity with sequential three-way decision.
[51]	Three-way decision and random forest	NSL-KDD	Using three-way decision to divide attributes for random forest forming.

osMUD [20] provides an alternative approach to deploying MUD in a network gateway for small-scale IoT networks. However, the operation is based on standard MUD architecture, which only conducts stateless communication inspection for traffic filtering.

Numerous studies [23,26] have extended the MUD profile to support the communication rate limit but do not consider the network behaviour analysis for malicious activity detection.

Corno and Manella [24] proposed a plug-in to create intended behaviour profiles for local MUD file generation. Therefore, stateful inspection and malicious activity detection are not the focus of their study.

Various studies [27,32] have only emphasised vulnerability detection in network gateways without considering malicious activity detection in network traffic.

Two studies [29,31] have employed federated learning for malicious activity detection in MUD-based networks. Feraudo et al. [29] only focused on device connection authorisation. Datta et al. [31] employed two-level defences in their architecture but did not support uncommitted decisions to allow ongoing monitoring of uncertain network behaviours.

Afek et al. [33] focused on extending the MUD capability to support peer-to-peer communication in IoT networks. Therefore, their study is specific to peer-to-peer networks.

The research works in [34,36] only focused on analysing MUD-rejected traffic. Hadi et al. [34] analysed MUD-rejected traffic for attack source detection. Zangrandi et al. [36] inspected rejected traffic for attack-type identification.

Andalibi et al. [39] focused on proposing a tool for visualising MUD traffic without considering malicious activity detection.

Bremler-Barr et al. [41,42] emphasised merging two MUD files into one generalised MUD file and discussed the impact of location on MUD file generation.

Different studies [47,48] have implemented DBN for feature extraction to be used in three-way decision theory with different applications. Shen [47] focused on two DDoS attack prevention on SDN and Du et al. [48] evaluated using two datasets. They did not focus on monitoring the IoT network traffic in real-time.

Several studies [49,50] have employed an autoencoder for feature extraction to form the multi-granularity for three-way decisions. They were not concerned with real-time IoT network monitoring.

Zhang et al. [51] employed three-way decision theory for grouping attributes based on the attribute importance calculation for generating a decision tree to form a random forest. This research did not focus on the implementation of real-time IoT network monitoring.

3.5. Limitations of the MUD Architecture

Our study identified the limitations of the MUD architecture and existing works on MUD architecture, which can be expressed as follows:

- The MUD architecture only conducts stateless traffic inspections without monitoring communication states. This could cause false negative filtering, as malicious activities might pass through the MUD enforcement.
- The MUD architecture does not support uncommitted decisions for uncertain network behaviours, which requires continuous analysis.

The MUD architecture only supports stateless inspection to allow or deny network traffic based on pre-defined ACL rules from the MUD file. This can cause false negative filtering as MUD does not monitor communication states to identify malicious activity. Moreover, uncertain network behaviours could occur in the traffic, which requires continuous monitoring to determine access permission. The MUD architecture only supports binary decisions with allow or deny, which is inadequate for supporting the uncommitted decision for ongoing monitoring of uncertain behaviours.

The identified limitations of the MUD architecture and existing works could jeopardise IoT networks as they have yet to consider the stateful inspection for detecting malicious activity. This leads to the intention of the proposal of the new gateway-based MUD architecture, EPA, to improve the effectiveness of malicious activity detection.

4. Proposal of the EPA

This study proposes a new architecture design, EPA, to improve the effectiveness of malicious activity detection. Our EPA considers the following two main aspects to improve the effectiveness of malicious activity detection:

- stateless and stateful communication inspection to detect malicious activity effectively;
- three-way decision-making with allow, deny, and uncommitted to continuously analyse uncertain behaviours.

The overview of the EPA is displayed in Figure 5.

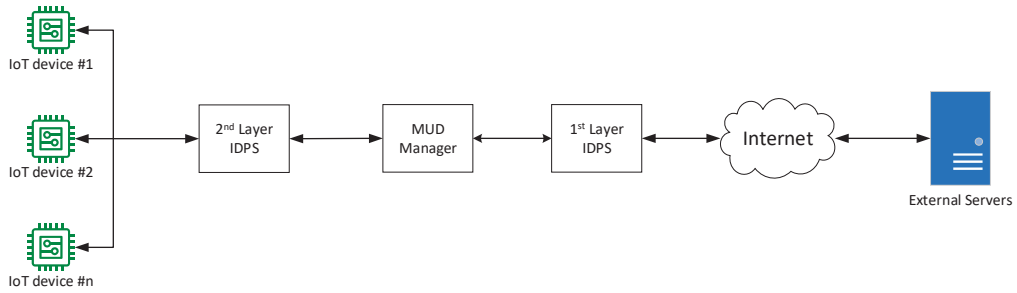


Figure 5. Overview of the EPA.

The EPA employs two layers of an IDPS for stateless and stateful communication monitoring on inbound and outbound traffic to improve the effectiveness of malicious activity detection. The first layer of the IDPS conducts a stateless inspection based on the pre-defined ACL rules from the MUD file. It is placed between the MUD manager and the external connection, focusing on monitoring inbound traffic. The second layer of the IDPS is placed between the local IoT network and the MUD manager to conduct stateful communication inspection on outbound traffic. It receives the communication packets and forwards them to the NBA system for further inspection to detect malicious activity. The behaviour analysis result then updates the ACL rules of both IDPS layers to allow or deny the analysed packet and similar packets.

The reasons for employing two-layer IDPSs are as follows:

- The MUD architecture relies on stateless inspections for packet filtering, which may be insufficient to detect and prevent malicious activities as the attack patterns evolve.
- A two-layer IDPS can separate simple and complex traffic monitoring tasks to reduce the traffic congestion that occurs in the first layer, thus reducing the latency in network inspection.

- A two-layer IDPS design can avoid a single point of failure. If one IDPS is down, another active IDPS can still operate to prevent malicious activity from reaching the destination host.
- The second layer of the IDPS can protect the MUD manager and NBA system by segregating the IoT networks from being directly connected to the MUD manager and NBA system.
- A two-layer IDPS in the EPA is the conceptual design that displays how it can protect against malicious activities. The IDPS can be deployed on a single or separate entity as EPA does not specify the physical setup.

4.1. Components

The EPA incorporates different components to improve the effectiveness of detecting malicious activity in the MUD-based network. The components comprise the MUD manager, NBA system, network behaviour knowledge databases, and two-layer IDPSs. Figure 6 displays the data flow of each component of the EPA.

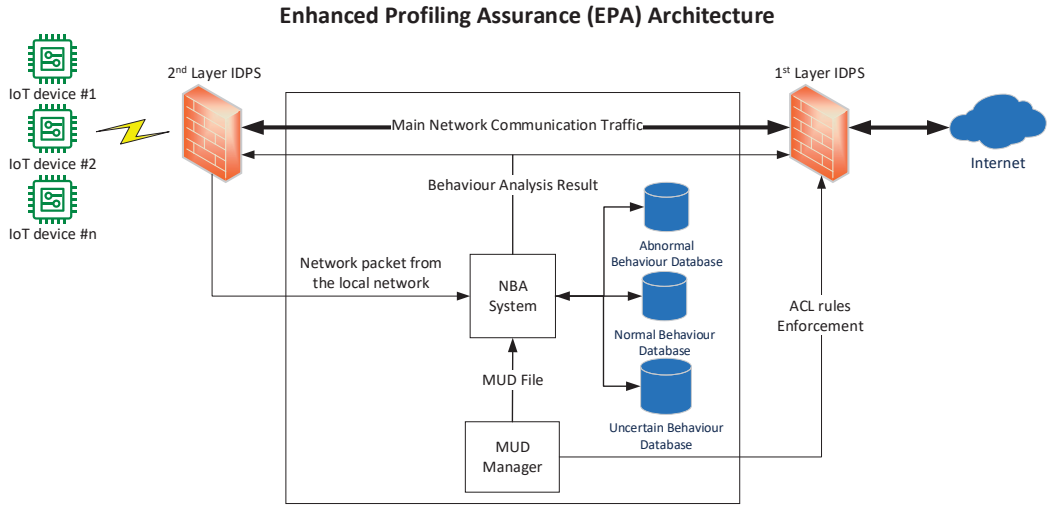


Figure 6. Internal components of the EPA.

The functionality of each component can be described as follows:

1. MUD manager

The MUD Manager retrieves the MUD file and translates it into ACL rules to enforce at the first layer of the IDPS. The MUD file is also forwarded to the NBA system to analyse incoming traffic before conducting further analysis.

2. NBA system

The NBA system conducts stateless and stateful inspections to monitor the communication from local IoT networks. The NBA system should be deployed on edge computing devices.

To analyse the network behaviour, the NBA system includes a packet sniffer to capture the network packet. The network behaviour analysis process uses the behaviour information from the MUD file and network behaviour knowledge database to identify malicious activity in the traffic. The examination process can be expressed as follows:

- the NBA system first conducts stateless inspection using ACL rules from the MUD file for packet filtering;
- the stateful communication inspection is conducted after stateless inspection to monitor communication states for malicious activity detection.

To achieve the objective of stateful communication inspection, the NBA system includes a communication tracking system that tracks and records each communication status.

3. Network behaviour knowledge databases

Knowledge-based databases are part of the NBA system. They store the network behaviour signatures to conduct a stateful inspection. The network behaviour signatures are stored in three separate databases as follows:

- Abnormal behaviour database

This database stores known abnormal network behaviour signatures that must be rejected, such as ICMP flooding and HTTP flooding patterns.

- Normal behaviour database

This database stores known normal network behaviour signatures to allow communication between the client and server.

- Uncertain behaviour database

This database stores known uncertain behaviour signatures stores for continuous observation.

Our knowledge-based database design supports constantly updating new behaviour signatures, which can be added to existing databases without redesigning the architecture.

4. First layer of the IDPS

The first layer of the IDPS conducts stateless communication inspection to filter packets. Packet behaviour inspection is based on the ACL rules from the MUD file.

5. Second layer of the IDPS

The second layer of the IDPS emphasises preventing malicious behaviour from infected devices. It forwards incoming packets to the NBA system to conduct stateless and stateful communication inspections to detect malicious activities and deny them.

4.2. Implementation of the EPA

This section provides an example of a possible implementation of the EPA. In this example, the EPA is constructed from numerous virtual machines (VMs) located in one physical server. This implementation can be set up as per the network topology in Figure 7.

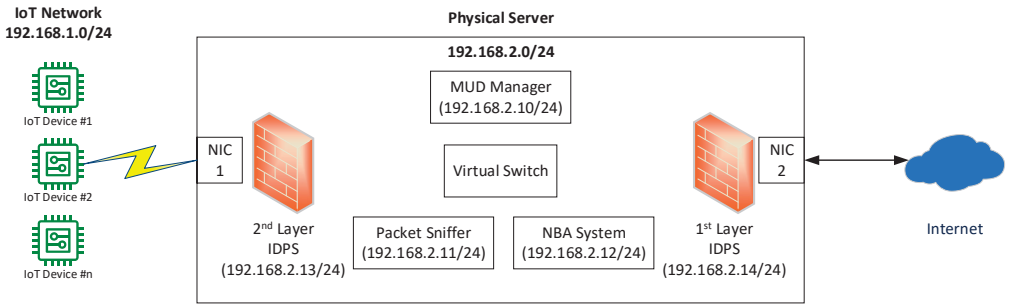


Figure 7. Example of EPA implementation.

The components required for this implementation are as follows:

1. one physical server with two network interface cards (NICs) and a Linux-based OS/Hypervisor;
2. five Linux-based VMs for the following:
   - MUD manager (OSMUD);
   - packet sniffer (tcpdump)
   - NBA system;
   - first layer of the IDPS (iptables);
   - second layer of the IDPS (iptables).

The EPA does not specify the physical setup. The EPA is flexible and can be integrated into the existing network. For example, it can be deployed into an existing network edge device or appliance.

4.3. EPA Workflow and Decision-Making

4.3.1. Workflow of the EPA

The sequence diagram of the workflow of the EPA is shown in Figure 8.

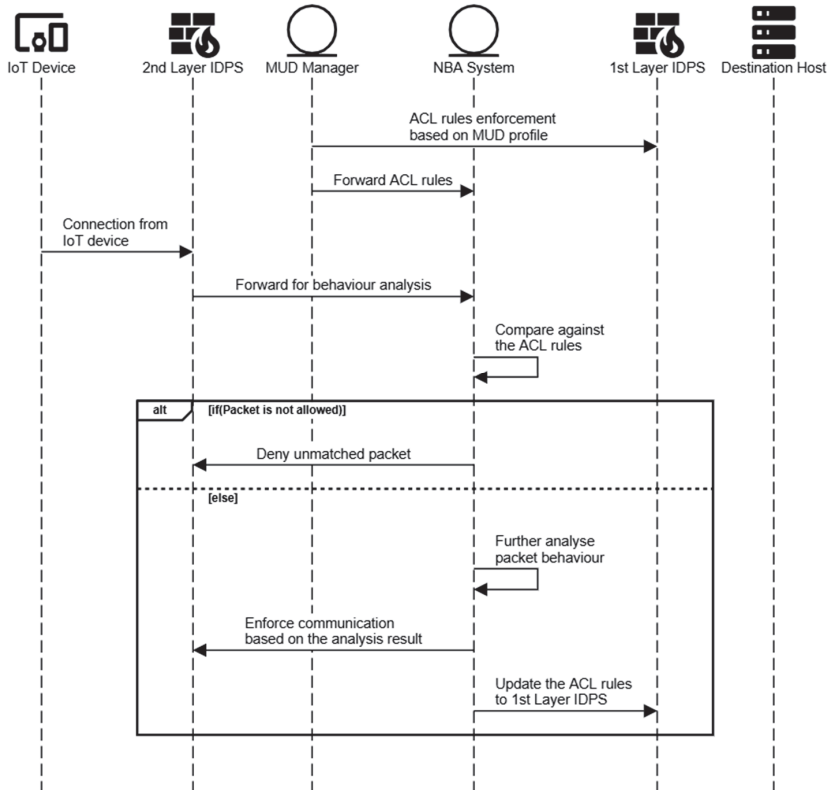


Figure 8. Workflow of the EPA.

The workflow of the EPA for network behaviour analysis can be expressed as follows:

1. the MUD manager retrieves the MUD file, translates it into ACL rules, and then enforces it at the first layer of the IDPS;
2. the MUD Manager forwards the MUD file to the NBA system;
3. the IoT device sends communication packets through the second layer of the IDPS;
4. the second layer of the IDPS forwarded the received packet to be analysed by the NBA system;
5. the NBA system compares the received packet against the ACL rules specified in the MUD file;
6. if the packet behaviour does not match the ACL rules, it will be rejected;
7. if the packet behaviour matches the ACL rules, further analysis will be conducted to detect malicious activity;
8. based on the result of the analysis, the NBA system informs the second layer of the IDPS to allow or deny the connection;
9. the NBA system informs the first layer of the IDPS to reject similar connection patterns.

4.3.2. Network Behaviour Analysis and Decision-Making Process

The introduction of the network behaviour analysis and decision-making process in EPA is intended to improve the effectiveness of malicious behaviour detection. This extends the decision-making capability of the MUD-based network by allowing an ongoing

monitoring process of uncertain network behaviours. The three-way decision-making process of the EPA is shown in Figure 9.

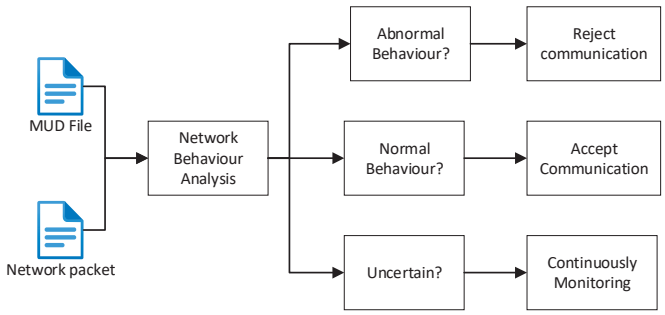


Figure 9. Three-way decision-making process of the EPA.

The flowchart of network behaviour analysis and decision-making is shown in Figure 10.

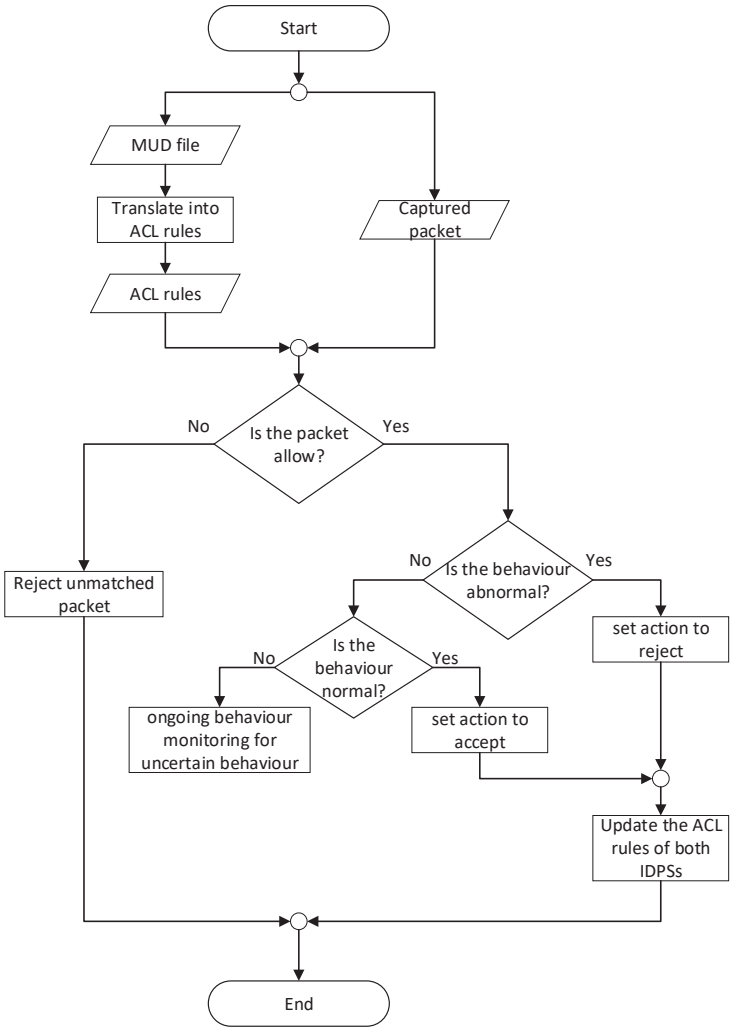


Figure 10. Flowchart of network behaviour analysis and decision-making process.

The sequence diagram of network behaviour analysis and decision-making process is shown in Figure 11.

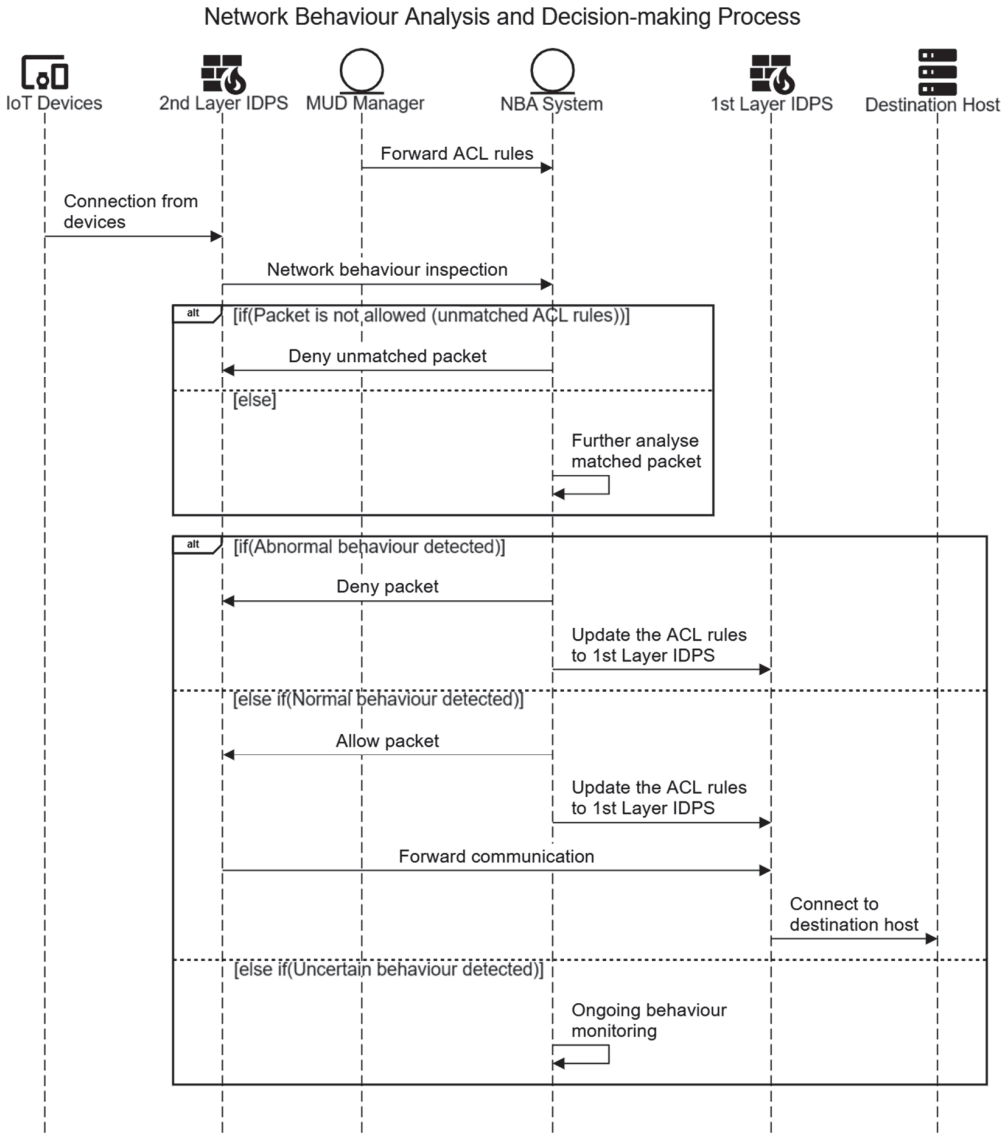


Figure 11. Sequence diagram of network behaviour analysis and decision-making process.

The detailed algorithm of the network behaviour analysis and decision-making process can be expressed in Algorithm 1.

The analysis begins with capturing the network packet and conducting a stateless packet inspection. This process filters the incoming packet by using ACL rules translated from the MUD file. Two possible outcomes of this process are as follows:

- rejection of the unmatched packet;
- further analysis of the packet behaviour to detect malicious activity.

To further analyse the packet behaviour, the NBA system conducts the stateful inspections of packets by investigating and monitoring the communication state to detect malicious activities that the MUD cannot detect. Examples include, but are not limited to, the following:

- the transmission rate is higher than the usual rate;
- the transmission rate does not match the previous rate;
- there is a mismatched protocol and port number usage (for example, traffic arrived at port 80, but it is not the HTTP communication).

Algorithm 1 Network behaviour analysis and decision-making

```
Require:
        MUDretrieval {MUD file retrieval and ACL translation function of MUD manager}
        ACLrules {translated ACL rules from MUD file}
        packet {captured packet for analysis}
        IDPSupdate {function for updating the ACL rules of both IDPSs}
        action {action for allowing or denying packet for updating IDPSs}
        ongoingmonitor {ongoing behaviour monitoring function}
ACLrules ← MUDretrieval()
if packet = ACLrules then
    if packet = abnormal then
        action ← deny
        IDPSupdate()
        break
    else if packet = normal then
        action ← allow
        IDPSupdate()
        break
    else if packet = uncertain then
        ongoingmonitor()
    end if
else
    return deny
    end if
```

The MUD-permitted packet is compared against the abnormal behaviour database to detect malicious activity. If abnormal behaviour is detected, the behaviour analysis system informs the second layer of the IDPS to reject the packet to prevent it from reaching its destination. The analysis result is also updated at the first layer of the IDPS to block similar communication patterns.

Suppose the analysis cannot detect malicious behaviour. The analysis then compares with the normal behaviour database, which will be permitted if it is normal. The first layer of the IDPS has also been updated to allow similar communication patterns to reach the destination.

If the uncertain behaviour is detected, the analysis will call the function for ongoing monitoring for uncertain behaviours to further analyse before allowing or rejecting the packet.

5. Analysis

This section uses two scenarios to demonstrate how the EPA can effectively detect malicious ICMP and HTTP traffic.

5.1. Malicious ICMP Traffic Detection

Suppose an IoT device needs to track the connectivity status of a host vpn.netatmo.net. The MUD file specifies that it allows the IoT device to send ICMP request messages to this host and receive ICMP reply packets from the host, as shown in Figure 12. If the IoT device were compromised and made part of botnets to launch an ICMP flood attack against the host, then the MUD file would fail to block the attack. As the MUD file grants access based on pre-defined access rules, it does not analyse the communication patterns to block this malicious traffic.

The EPA considers the permitted access as an unknown behaviour and conducts behaviour analysis. In the case of an ICMP flood attack, the message volume is usually larger than usual in a short period. The EPA can detect this attack using the knowledge of time difference and communication rate from the abnormal behaviour database. It informs the first layer of the IDPS of the analysis result. Figure 13 shows the process of detecting

malicious ICMP behaviour. The ICMP communication behaviour analysis and malicious activity detection process can be described as follows:

1. the IoT device sends ICMP echo-request packets to the destination host through the second layer of the IDPS;
2. the second layer of the IDPS forwards the arrived ICMP packets to be analysed by the NBA system;
3. the NBA system compares the received ICMP packets against the access control list (ACL) rules specified in the MUD file;
4. the ICMP packets match the ACL rule in the MUD file, and then the NBA conducts malicious detection against the abnormal behaviour database, normal database, and uncertain behaviour databases;
5. it identifies that this is a malicious ICMP flood attack;
6. the NBA system informs the first and second layers of the IDPS to block this malicious ICMP flood traffic.


```
}, {
  "name" : "from-ipv4-netatmo_welcome-3",
  "matches" : {
    "ipv4" : {
      "protocol" : 1,
      "ietf-acldns:dst-dnsname" : "vpn.netatmo.net"
    },
    "icmp" : {
      "type" : 8,
      "code" : 0
    }
  },
  "actions" : {
    "forwarding" : "accept"
  }
}, {
```

(**a**)

```
}, {
  "name" : "to-ipv4-netatmo_welcome-12",
  "matches" : {
    "ipv4" : {
      "protocol" : 1,
      "ietf-acldns:src-dnsname" : "vpn.netatmo.net"
    },
    "icmp" : {
      "type" : 0,
      "code" : 0
    }
  },
  "actions" : {
    "forwarding" : "accept"
  }
}, {
```

(**b**)

Figure 12. Code snippets of a MUD file—allowing ICMP echo-request (**a**) and echo-reply (**b**).

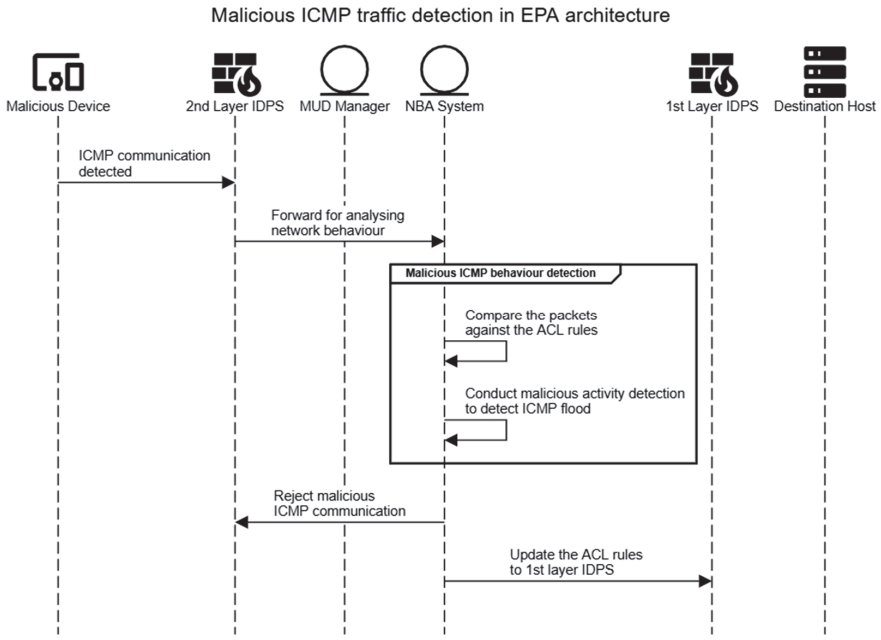


Figure 13. Malicious ICMP behaviour detection process in the EPA.

5.2. Malicious HTTP Traffic Detection

Assume that an IoT device needs to access the data on the web server clients3.google.com via HTTP. The MUD file specifies that it allows the HTTP connection to port 80 of the web server over TCP (identified as protocol 6), as shown in Figure 14. If the IoT device was manipulated as part of botnets to launch an HTTP flood attack against the web server, then

the MUD file would not be effective in blocking this HTTP flood attack. As the MUD file grants access based on pre-defined access rules, it does not examine the network traffic for unusual communication patterns to block this malicious traffic.

```
}, {
  "name" : "from-ipv4-netatmo_welcome-6",
  "matches" : {
    "ipv4" : {
      "protocol" : 6,
      "ietf-acldns:dst-dnsname" : "clients3.google.com"
    },
    "tcp" : {
      "destination-port" : {
        "operator" : "eq",
        "port" : 80
      },
      "ietf-mud:direction-initiated" : "from-device"
    }
  },
  "actions" : {
    "forwarding" : "accept"
  }
}, {
```

Figure 14. Code snippet of a MUD file—allowing the HTTP connections to port 80 of the web server.

The EPA considers the permitted access as an unknown behaviour. It conducts network behaviour analysis by monitoring traffic and observing unusual activity and departures from normal network operation patterns. As the HTTP protocol invokes the underlying transport protocol TCP, the communication inspection monitors the three-way handshake process in TCP communication before inspecting the HTTP protocol. The NBA system monitors the traffic and compares it against the abnormal behaviour, followed by the normal behaviour database. It blocks this traffic, as it examines the pattern in the packets that match one of the abnormal behaviour database signatures. The process of malicious HTTP behaviour detection is shown in Figure 15.

The HTTP communication behaviour analysis and malicious activity detection process can be described as follows:

1. The IoT device sends HTTP packets through the second layer of the IDPS;
2. The second layer of the IDPS forwards the arrived packets to be analysed by the NBA system;
3. The NBA system compares the received HTTP packets against the ACL rules specified in the MUD profile;
4. The HTTP packets match the ACL rule in the MUD profile. Then, the NBA conducts malicious detection against the abnormal behaviour database, normal database, and uncertain behaviour databases;
5. The NBA system inspects the three-way handshake process in TCP communication. Then, it inspects the HTTP request packets to identify that it is a malicious HTTP flood packet;
6. The NBA system informs the second layer of the IDPS to reject this HTTP flood communication;
7. The second layer of the IDPS forwards the analysis result to the first layer of the IDPS to prevent this malicious communication.

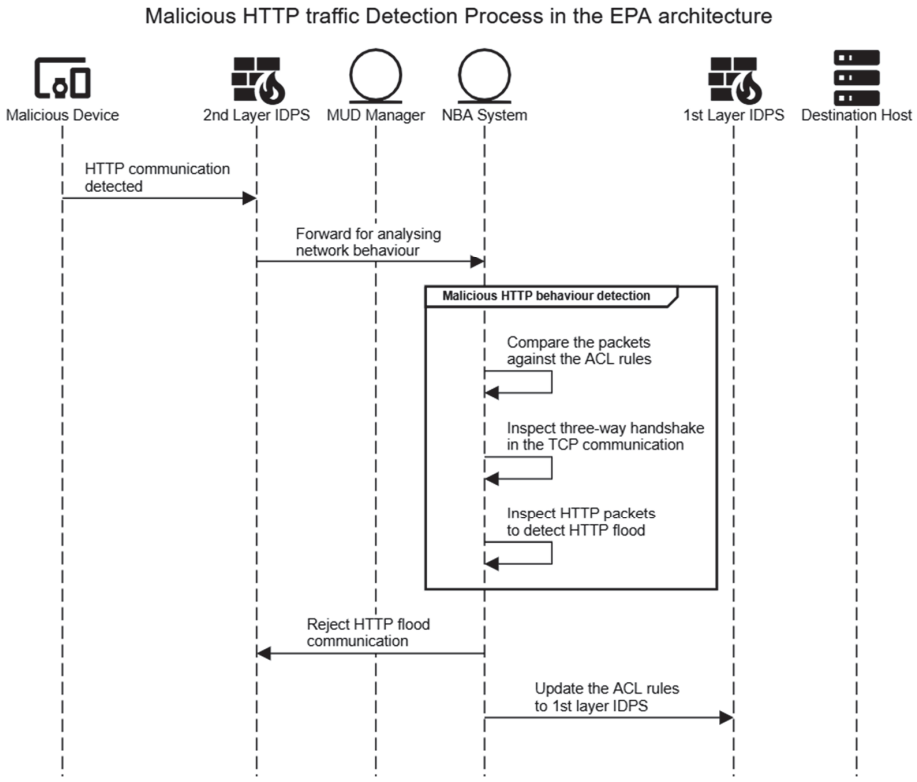


Figure 15. Malicious HTTP traffic detection process in the EPA.

6. Conclusions and Future Work

This work has proposed the EPA to improve the effectiveness of malicious activity detection, focusing on small-scale IP-based IoT network deployment. EPA addresses the limitations of the MUD architecture by incorporating both stateless and stateful inspection to effectively detect malicious activities. EPA adopts a three-way decision theory for continuously analysing uncertain network behaviour by calling the ongoing behaviour monitoring function. We demonstrate the effectiveness of the EPA through the analysis of two communication scenarios, ICMP and HTTP. This shows that our EPA achieves superiority over the MUD architecture in detecting malicious activities in the network traffic. This is because the EPA monitors each communication state to detect and prevent malicious traffic, which differs from the MUD architecture as it relies only on stateless inspection.

The EPA is designed for small-scale networks. The EPA's ability to scale to larger networks is currently unknown. Incorporating stateless and stateful inspection could increase the computational complexity. We have yet to conduct the complexity analysis as this study primarily focused on the architectural design proposal.

In future work, we will develop a prototype proof of concept and implement three-way decision theory to demonstrate the effectiveness of malicious activity detection and prevention. We will also conduct a complexity analysis to determine the efficiency of the EPA.

Author Contributions: Conceptualisation and design, N.A., V.L., L.K. and Y.L.; investigation, N.A. and V.L.; analysis, N.A., V.L. and L.K.; methodology, N.A., V.L. and L.K.; writing, N.A. and V.L.; review and editing, V.L., L.K. and A.D.T.; supervision, V.L., L.K., Y.L. and M.M.; original draft preparation, N.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data supporting this study is publicly available in the UNSW IoT Analytics at https://iotanalytics.unsw.edu.au/mudprofiles.html (accessed on 30 May 2024).

Acknowledgments: The Royal Thai Government has sponsored Nut Aroon for his Ph.D. study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Nayak, G.; Mishra, A.; Samal, U.; Mishra, B.K. Depth Analysis on DoS & DDoS Attacks. In *Wireless Communication Security*; Scrivener Publishing: Beverly, MA, USA, 2022; pp. 159–182.
2. Gamblin, J. Mirai BotNet. Available online: https://github.com/jgamblin/Mirai-Source-Code (accessed on 15 March 2024).
3. Greenstein, S. The Aftermath of the Dyn DDOS Attack. *IEEE Micro* **2019**, *39*, 66–68. [CrossRef]
4. Yoachimik, O.; Pacheco, J. DDoS threat report for 2024 Q1. Available online: https://blog.cloudflare.com/ddos-threat-report-for-2024-q1 (accessed on 17 April 2024).
5. Lear, E.; Droms, R.; Romascanu, D. RFC 8520: Manufacturer Usage Description Specification. Available online: https://datatracker.ietf.org/doc/html/rfc8520 (accessed on 15 March 2024).
6. Jethanandani, M.; Agarwal, S.; Huang, L.; Blair, D. YANG Data Model for Network Access Control Lists (ACLs). Available online: https://datatracker.ietf.org/doc/html/rfc8519 (accessed on 9 May 2024).
7. Boehm, B.; Howard, B.; Aboba, B.; Petri, B.; Nguyen, B.; McIntosh, B.; Braden, B.; Hinden, B.; Kantor, B.; Lee, C.; et al. Protocol Numbers. Available online: https://www.iana.org/assignments/protocol-numbers/protocol-numbers.xhtml (accessed on 17 May 2024).
8. Touch, J.; Lear, E.; Ono, K.; Eddy, W.; Trammell, B.; Iyengar, J.; Scharf, M.; Tuexen, M.; Kohler, E.; Nishida, Y. Service Name and Transport Protocol Port Number Registry. Available online: https://www.iana.org/assignments/service-names-port-numbers/service-names-port-numbers.xhtml (accessed on 17 May 2024).
9. Heeb, Z.; Kalinagac, O.; Soussi, W.; Gur, G. The Impact of Manufacturer Usage Description (MUD) on IoT Security. In Proceedings of the 2022 1st International Conference on 6G Networking (6GNet), Paris, France, 6–8 July 2022.
10. Souppaya, M.; Montgomery, D.; Polk, T.; Ranganathan, M.; Dodson, D.; Barker, W.; Johnson, S.; Kadam, A.; Pratt, C.; Thakore, D.; et al. *Securing Small-Business and Home Internet of Things (IoT) Devices: Mitigating Network-Based Attacks Using Manufacturer Usage Description (MUD)*; National Institute of Standards and Technology: Gaithersburg, MD, USA, 2021.
11. Watrobski, P.; Klosterman, J. MUD-PD. Available online: https://github.com/usnistgov/MUD-PD (accessed on 24 March 2024).
12. Lear, E.; Weis, B. Slinging MUD: Manufacturer usage descriptions: How the network can protect things. In Proceedings of the 2016 International Conference on Selected Topics in Mobile & Wireless Networking (MoWNeT), Cairo, Egypt, 11–13 April 2016.
13. What is MUD? Available online: https://developer.cisco.com/docs/mud/what-is-mud/#what-is-mud (accessed on 16 May 2024).
14. DeKok, A.; Cudbard-Bell, A.; Newton, M.; Clouter, A. FreeRADIUS. Available online: https://freeradius.org/ (accessed on 16 May 2024).
15. Shah, R.; Madson, C.; Lear, E. CiscoDevNet MUD-Manager. Available online: https://github.com/CiscoDevNet/MUD-Manager (accessed on 16 May 2024).
16. Hamza, A.; Ranathunga, D.; Gharakheili, H.H.; Roughan, M.; Sivaraman, V. Clear as MUD: Generating, Validating and Applying IoT Behavioral Profiles. In Proceedings of the 2018 Workshop on IoT Security and Privacy, Budapest, Hungary, 20 August 2018.
17. Hamza, A. MUDGEE. Available online: https://github.com/ayyoob/mudgee (accessed on 4 March 2024).
18. Hamza, A.; Ranathunga, D.; Gharakheili, H.H.; Benson, T.A.; Roughan, M.; Sivaraman, V. Verifying and Monitoring IoTs Network Behavior Using MUD Profiles. *IEEE Trans. Dependable Secur. Comput.* **2022**, *19*, 1–18. [CrossRef]
19. Hamza, A.; Ranathunga, D.; Habibi Gharakheili, H.; Benson, T.A.; Roughan, M.; Sivanathan, A. MUD Profiles. Available online: https://iotanalytics.unsw.edu.au/mudprofiles.html (accessed on 9 May 2024).
20. osMUD—The Open Source MUD Manager. Available online: https://osmud.org/ (accessed on 20 March 2024).
21. OpenWRT. Available online: https://openwrt.org/ (accessed on 20 March 2024).
22. Kelly, S. Dnsmasq. Available online: https://thekelleys.org.uk/dnsmasq/doc.html (accessed on 20 March 2024).
23. Andalibi, V.; Kim, D.; Camp, J. Throwing MUD into the FOG: Defending IoT and Fog by expanding MUD to Fog network. In Proceedings of the 2nd USENIX Workshop on Hot Topics in Edge Computing, HotEdge 2019, Co-Located with USENIX ATC 2019, Renton, WA, USA, 9 July 2019.
24. Corno, F.; Mannella, L. A Gateway-based MUD Architecture to Enhance Smart Home Security. In Proceedings of the 2023 8th International Conference on Smart and Sustainable Technologies (SpliTech), Split/Bol, Croatia, 20–23 June 2023; pp. 1–6.
25. Home Assistant. Available online: https://www.home-assistant.io/ (accessed on 26 March 2024).
26. Feraudo, A.; Popescu, D.A.; Yadav, P.; Mortier, R.; Bellavista, P. Mitigating IoT Botnet DDoS Attacks through MUD and eBPF based Traffic Filtering. In Proceedings of the 25th International Conference on Distributed Computing and Networking, Chennai, India, 4–7 January 2024; pp. 164–173.
27. Sajjad, S.M.; Yousaf, M.; Afzal, H.; Mufti, M.R. eMUD: Enhanced Manufacturer Usage Description for IoT Botnets Prevention on Home WiFi Routers. *IEEE Access* **2020**, *8*, 164200–164213. [CrossRef]
28. OWASP Firmware Security Testing Methodology. Available online: https://github.com/scriptingxss/owasp-fstm (accessed on 25 May 2024).
29. Feraudo, A.; Yadav, P.; Safronov, V.; Popescu, D.A.; Mortier, R.; Wang, S.; Bellavista, P.; Crowcroft, J. CoLearn: Enabling federated learning in MUD-compliant IoT edge networks. In Proceedings of the Third ACM International Workshop on Edge Systems, Analytics and Networking, Heraklion, Greece, 7 April 2020.

30. Ziller, A.; Trask, A.; Lopardo, A.; Szymkow, B.; Wagner, B.; Bluemke, E.; Nounahon, J.-M.; Passerat-Palmbach, J.; Prakash, K.; Rose, N.; et al. PySyft: A Library for Easy Federated Learning. In *Federated Learning Systems: Towards Next-Generation AI*.; Rehman, M.H.u., Gaber, M.M., Eds.; Springer International Publishing: Cham, Switzerland, 2021; pp. 111–139.
31. Datta, S.; Bhattacharya, A.; Rana, R.; Venkanna, U. iDAM: A Distributed MUD Framework for Mitigation of Volumetric Attacks in IoT Networks. In Proceedings of the 2022 13th International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP), Porto, Portugal, 20–22 July 2022.
32. Un Nisa, K.; Alhudhaif, A.; Qureshi, K.N.; Hadi, H.J.; Jeon, G. Security provision for protecting intelligent sensors and zero touch devices by using blockchain method for the smart cities. *Microprocess. Microsyst.* **2022**, *90*, 104503. [CrossRef]
33. Afek, Y.; Bremler-Barr, A.; Hay, D.; Shalev, A. MUDirect: Protecting P2P IoT Devices with MUD. In Proceedings of the 2021 IEEE International Conferences on Internet of Things (iThings) and IEEE Green Computing & Communications (GreenCom) and IEEE Cyber, Physical & Social Computing (CPSCom) and IEEE Smart Data (SmartData) and IEEE Congress on Cybermatics (Cybermatics), Melbourne, Australia, 6–8 December 2021.
34. Hadi, H.J.; Sajjad, S.M.; Nisa, K. BoDMitM: Botnet Detection and Mitigation System for Home Router Base on MUD. In Proceedings of the 2019 International Conference on Frontiers of Information Technology (FIT), Islamabad, Pakistan, 16–18 December 2019; pp. 139–1394.
35. Cisco. Snort—Network Intrusion Detection & Prevention System. Available online: https://www.snort.org/ (accessed on 18 March 2024).
36. Zangrandi, L.M.; Ede, T.V.; Booij, T.; Sciancalepore, S.; Allodi, L.; Continella, A. Stepping out of the MUD: Contextual threat information for IoT devices with manufacturer-provided behavior profiles. In Proceedings of the 38th Annual Computer Security Applications Conference, Austin, TX, USA, 5–9 December 2022.
37. Zangrandi, L.M.; Ede, T.V. MUDscope. Available online: https://github.com/lucamrgs/MUDscope (accessed on 3 April 2024).
38. Morgese Zangrandi, L.; van Ede, T.; Booij, T.; Sciancalepore, S.; Allodi, L.; Continella, A. MUDscope dataset. Available online: https://zenodo.org/records/7182597 (accessed on 30 May 2024).
39. Andalibi, V.; Dev, J.; Kim, D.; Lear, E.; Camp, L.J. Is Visualization Enough? Evaluating the Efficacy of MUD-Visualizer in Enabling Ease of Deployment for Manufacturer Usage Description (MUD). In Proceedings of the Annual Computer Security Applications Conference, Virtual Event, 6–10 December 2021; pp. 337–348.
40. Lear, E.; Andalibi, V. MUD Visualizer. Available online: https://github.com/iot-onboarding/mud-visualizer (accessed on 30 March 2024).
41. Bremler-Barr, A.; Meyuhas, B.; Shister, R. MUDIS: MUD Inspection System. In Proceedings of the NOMS 2022–2022 IEEE/IFIP Network Operations and Management Symposium, Budapest, Hungary, 25–29 April 2022.
42. Bremler-Barr, A.; Meyuhas, B.; Shister, R. One MUD to Rule Them All: IoT Location Impact. In Proceedings of the NOMS 2022–2022 IEEE/IFIP Network Operations and Management Symposium, Budapest, Hungary, 25–29 April 2022.
43. Li, Y.; Zhang, L.; Xu, Y.; Yao, Y.; Lau, R.Y.K.; Wu, Y. Enhancing Binary Classification by Modeling Uncertain Boundary in Three-Way Decisions. *IEEE Trans. Knowl. Data Eng.* **2017**, *29*, 1438–1451. [CrossRef]
44. Subhashini, L.D.C.S.; Li, Y.; Zhang, J.; Atukorale, A.S. Assessing the effectiveness of a three-way decision-making framework with multiple features in simulating human judgement of opinion classification. *Inf. Process. Manag.* **2022**, *59*, 102823. [CrossRef]
45. Subhashini, L.D.C.S.; Li, Y.; Zhang, J.; Atukorale, A.S. Integration of Fuzzy and Deep Learning in Three-Way Decisions. In Proceedings of the 2020 International Conference on Data Mining Workshops (ICDMW), Sorrento, Italy, 17–20 November 2020; pp. 71–78.
46. Subhashini, L.D.C.S.; Li, Y.; Zhang, J.; Atukorale, A.S. Integration of Fuzzy and LSTM in Three-Way Decisions. In Proceedings of the 2020 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT), Melbourne, Australia, 14–17 December 2020; pp. 975–980.
47. Shen, Y. An Intrusion Detection Algorithm for DDoS Attacks Based on DBN and Three-way Decisions. *J. Phys. Conf. Ser.* **2022**, *2356*, 012044. [CrossRef]
48. Du, X.; Li, Y.; Zhang, S. Research on Intrusion Detection Algorithm Based on Deep Belief Networks and Three-way Decisions. In Proceedings of the 2020 4th International Conference on Electronic Information Technology and Computer Engineering, Xiamen, China, 6–8 November 2020; pp. 50–56.
49. Zhang, S.; Li, Y.; Du, X. An Intrusion Detection Approach Based on Autoencoder and Three-way Decisions. In Proceedings of the 2020 4th International Conference on Electronic Information Technology and Computer Engineering, Xiamen, China, 6–8 November 2020; pp. 495–500.
50. Geng, Y.; Li, Y.; Zhang, S. Research on Multi-granularity Intrusion Detection Algorithm Based onSequential Three-Way Decision. In Proceedings of the 2021 5th International Conference on Electronic Information Technology and Computer Engineering, Xiamen, China, 22–24 October 2021; pp. 1154–1160.
51. Zhang, C.; Wang, W.; Liu, L.; Ren, J.; Wang, L. Three-Branch Random Forest Intrusion Detection Model. *Mathematics* **2022**, *10*, 4460. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

electronics

MDPI

Article

Resilience against Catastrophic Cyber Incidents: A Multistakeholder Analysis of Cyber Insurance

Brianna Bace [1], Elisabeth Dubois [2] and Unal Tatar [1,*]

[1] College of Emergency Preparedness, Homeland Security and Cybersecurity, University at Albany, State University of New York, Albany, NY 12203, USA; bbace@albany.edu
[2] NYMIR Division, Wright Public Entity, Garden City, NY 11530, USA; evdubois@albany.edu
* Correspondence: utatar@albany.edu

Abstract: Catastrophic cyber incidents—events of low probability but high impact, with the potential to incur billions of dollars in damages—are prompting insurers to elevate premiums, create higher barriers for potential buyers, and tighten policies with exclusions. While these responses of the insurance industry are important to prevent its insolvency during catastrophic incidents due to excessive claims, they lead to a notable gap in market protection. Using a content analysis of multistakeholder comments submitted in response to a Treasury Department Request for Information (RFI), this study seeks to define what constitutes a catastrophic cyber event, identify mitigation strategies, evaluate the current capacity of the cyber insurance sector to handle such incidents, and investigate the potential roles and support mechanisms that the government can provide to enhance the insurance sector's capacity to manage these extreme risks. This paper is one of the pioneering studies using data and a multistakeholder perspective to provide essential guidance for policymakers, regulators, the insurance industry, and the cybersecurity sector in formulating robust policies and strategies to address catastrophic cyber risks, ultimately enhancing national economic and technological resilience.

Keywords: catastrophic cyber risk; cyber insurance; cyber resilience; economic resilience; regulation; content analysis; federal backstop

1. Introduction

In the realm of cybersecurity, the evolving nature of threats and the increasing sophistication of cyberattacks have necessitated a reevaluation of traditional risk management strategies. The role of cyber insurance, in this regard, has become a focal point for both the private sector and policymakers. As cyber threats transcend the digital domain, impacting the physical, financial, and reputational realms, businesses are increasingly turning to cyber insurance as a financial safeguard. However, the rise in catastrophic cyber incidents threatens the economic stability of the insurance sector and the feasibility and effectiveness of cyber insurance as a risk management tool.

The 2017 WannaCry and NotPetya attacks are two of the closest examples of catastrophic cyber incidents to date. The WannaCry ransomware attack, widely attributed to the North Korean hacking group Lazarus Group, targeted computers running the Microsoft Windows operating system in over 150 countries. Government agencies, healthcare institutions, universities, and businesses were impacted, and the costs of the hack are estimated to be USD 4 billion [1]. The NotPetya malware attack was a systemic cyber incident, as the data-destroying malware deployed by Russia against Ukraine spread outside of Ukraine to other companies and systems across the globe. It is estimated that this attack led to USD 10 billion in damages [2].

Cyberattacks like WannaCry and NotPetya have contributed to the significant hardening of the cyber insurance market. In the insurance business, it is better to have a soft

market than a hard one, as hardening can cause changes to prices, terms, and conditions, as well as the availability of coverage and capacity [3]. Specifically for the cyber insurance market, the price of premiums, deductibles, and retentions has increased, while policy coverage in areas of extortion and exclusions for nation-state attacks has tightened.

In response to the hardening of the cyber insurance market, the Government Accountability Office (GAO) published a report in June 2022 recommending the Cybersecurity and Infrastructure Security Agency (CISA) and the Federal Insurance Office (FIO) "jointly assess the extent to which risks to critical infrastructure from catastrophic cyber incidents and potential financial exposures warrant a federal insurance response" [4] (para. 12). In response to the GAO's recommendation, the FIO in the Treasury Department issued a Request for Information (RFI) on 28 September 2022, with submissions open until 14 December 2022 [5]. The public had just under three months to submit their comments to answer the Treasury Department's questions on this issue and provide additional information relating to cyber insurance and catastrophic cyber incidents. A total of 60 comments were submitted from private citizens and firms from different industries, including but not limited to cybersecurity and insurance.

This research aims to explore resilience mechanisms against catastrophic cyber incidents through a comprehensive multistakeholder analysis, with a focus on the role of cyber insurance. This study seeks to define what constitutes a catastrophic cyber event, identify mitigation strategies, evaluate the current capacity of the cyber insurance sector to handle such incidents, and investigate the potential roles and support mechanisms that the government can provide to enhance the insurance sector's capacity to manage these extreme risks. In this paper, we perform a content analysis of the comments submitted in response to the Treasury Department's RFI to identify concepts and themes present in the comments that answer our four research questions. This analysis reveals how government intervention, particularly through a federal insurance backstop, can mitigate the economic repercussions of catastrophic cyber risks. It also brings to the forefront the critical legal and regulatory aspects that need to be considered. These findings offer essential guidance for policymakers and legal experts working on formulating robust laws and regulations in the dynamically changing domain of cyber risks and insurance solutions.

This paper is structured as follows: Section 2 provides a comprehensive overview of the cyber insurance market, encompassing its historical evolution, current landscape, notable past cyberattacks, and prior federal insurance programs used to cover catastrophic risk. Section 2 also summarizes previous research efforts in this space and lays out the four research questions that guided the analysis performed in this paper. Section 3 outlines the methodology adopted in this study. In Section 4, we answer the research questions using the collected data. Section 5 includes a discussion of the concerns and benefits of creating a federal backstop for catastrophic cyber incidents as a government support mechanism for improving the capacity of the cyber insurance market. Section 6 offers suggestions as to the design of a federal backstop, considering the suggestions and concerns expressed by the commenters. Finally, in Section 7, we conclude.

2. Background

This section provides background on the cyber insurance market and catastrophic cyber incidents (CCIs), including the history of the market and past cyberattacks that are considered to be the closest examples of CCIs to date. We also discuss previous federal insurance programs and whether, in its present form, the Terrorism Risk Insurance Program (TRIP) can cover losses from a CCI. This section ends with an overview of previous research efforts in this space and lays out the four research questions that guide our investigation.

2.1. Cyber Insurance History and Current Market

The history of cyber insurance spans over two decades and reflects the ever-evolving landscape of digital threats and vulnerabilities [6]. As of 2020, it had grown to become a substantial industry, with direct premiums written amounting to around USD 4 billion [7].

Today, the United States stands as the global leader in the cyber insurance market [8,9]. This market segment is part of the broader property and casualty (P&C) insurance industry, and it offers various types of coverage to mitigate the financial impacts of cyber-related incidents, including first-party, third-party, and implicit silent cyber coverage. According to Baker and Shortland, there appears to be a trend of insurers shifting coverage for cyber risks out of general-purpose liability and property insurance policies and into specialized cyber insurance policies. These policies, which narrowly define coverage, help insurers reduce uncertainty by insuring losses of only a certain kind and for a certain amount, and not for everything cyber [10].

In its early years, cyber insurance was typically included as an addendum to commercial P&C insurance policies; however, as cyber threats became more sophisticated and the need for specialized coverage grew, standalone cyber insurance policies emerged as a significant trend. These policies are designed specifically to address the unique risks associated with cyberattacks and data breaches. The adoption of standalone cyber insurance policies varies by industry. Sectors such as financial services (59%), retail (56%), and media and entertainment (54%) are more likely to opt for standalone coverage [11]. In contrast, information technology and telecoms (35%) and energy, oil/gas, and utilities (39%) are sectors less likely to have standalone cyber insurance policies [11].

Factors such as regulatory changes, diverse cyber risks, digitalization, and supply chain vulnerabilities have propelled demand [12,13]. Threats such as ransomware, business email compromise, and fund transfer fraud are also moving organizations to seek coverage [14]. The industry grapples with challenges like risk assessment and policy standardization, which make the underwriting process for cyber insurance more complex and longer compared to traditional insurance [14]. The size of a business also plays a role in the adoption of cyber insurance [9]. According to a report by Cowbell Cyber, a cyber insurance company, 65% of small- and mid-size enterprises plan to increase spending on cyber insurance over the next two years compared to 58% of large U.S.-based enterprises [15].

By 2025, the global cost of cybercrime is estimated to be USD 10.5 trillion annually [16], underlining the significance of the cyber insurance market. As cyber threats continue to evolve and adapt to technological advancements, the demand for cyber insurance is expected to grow further [17]. Insurers will need to continually refine their policies and risk assessments to keep pace with the ever-changing landscape of cybersecurity threats and vulnerabilities. Additionally, regulatory developments and industry standards are likely to shape the future of the cyber insurance market as it continues to mature [6,18].

Unfortunately, there have been instances where insurance companies, despite entities having policies, attempt to avoid payouts due to stringent language and exclusions within the policy. We saw this in the months following NotPetya. Mondelez International, a multinational food company based in Chicago, was a victim of the NotPetya attack, with losses totaling more than USD 180 million for the company [19]. After the attack, Mondelez, who has an all-risk property insurance policy with Zurich American Insurance, filed an insurance claim seeking compensation for their NotPetya losses. Zurich denied the claim and refused to provide coverage for damages by invoking the "hostile or war-like action" exclusion [19]. The Zurich case marked the inaugural major legal conflict in the insurance sector over the recovery costs stemming from a cyber breach [20]. After a four-year legal battle, both parties came to a settlement at the end of 2022 before a judge could make a definitive ruling on the issue of war exclusions.

The Zurich case also provided important lessons for the cyber insurance market. For one, traditional property insurance or silent coverage should not be relied upon to cover cyber risks as this type is likely to fall short of needs and leave important risks outside coverage since it is not sufficiently tailored [21]. Of course, insurers also need to make sure that cyber insurance policies are crafted with clear wording that avoids ambiguity [21]. Given the likelihood of future state-sponsored or widespread attacks like WannaCry and NotPetya, several improvements to the cyber insurance market need to take place to prevent companies from invoking arguments like those used by Zurich. The creation of

laws and regulations can also play a crucial role in ensuring cyber resilience for both the insurance industry and policyholders. These measures can safeguard against cyber threats at both sectoral and national levels, thereby fostering a secure environment for insurers and protecting the interests of those they insure.

It is crucial to highlight that cyber incidents are not just an issue for underwriters of cyber insurance policies; they are a significant problem for underwriters of non-cyber policies as well due to "non-affirmative" or "silent" risk. Silent risk can reside in many different types of policies, even policies written by insurers that write little or no cyber insurance policies. Many traditional liability insurance policies neither explicitly include nor exclude cyber coverage or coverage of cybersecurity incidents. This vagueness results in silent cyber coverage [21]. According to a 2020 global survey, 65% of underwriters were concerned about silent cyber coverage exposure in P&C policies [22]. This is a valid concern, as USD 2.7 billion of the USD 3 billion insurance claims that resulted from the NotPetya attack were made under P&C policies that did not explicitly cover cyber risks [22].

2.2. Catastrophic Cyber Incidents

There is not yet a widely accepted definition of what constitutes a "catastrophic" cyber incident, and the ambiguity that exists around the term is a major problem for the insurance industry. The Treasury Department states the term "catastrophic" is generally related to the magnitude of the loss, its dispersion among multiple entities, and the degree of critical services affected [5]. In many cases, the way we define catastrophic cyber incidents/risk is "in the eye of the beholder" [23]. Depending on the industry or sector, the definition of catastrophe varies widely based on the scale and scope of the impacts considered. Yet, there is broad agreement that a widespread cyberattack on critical infrastructure would be catastrophic [23].

From an insurance perspective, catastrophic cyber risk is typically defined as a single event exceeding a certain amount in insurable losses or a certain number of affected insurers [23]. They are often estimated to be 1-in-100 low-probability but high-impact events, while some vendors believe an industry-wide cyber event could result in 1-in-200 year potential losses [24]. Without a clear definition, measuring and modeling catastrophic cyber risk is a challenge. In risk modeling, access to accurate and relevant data is important to develop disaster scenarios [25]. The more data available, the more credibility the modeled outcomes have [25]. Cyber risk modeling has always been a challenge due to the lack of historical data for cyberattacks. There is even less data available for those that may be considered "catastrophic", if there is any.

Fortunately, efforts have been made by researchers to gather and analyze datasets for cyber risk quantification and modeling. Dubois et al. (2022) conducted a systematic interdisciplinary literature review for actuaries that consisted of collecting and classifying twenty different datasets including the academic and grey literature across multiple disciplines [26]. Cremer et al. (2022) performed a similar systematic review of data availability of the academic and industry literature on cybersecurity and cyber risk management that resulted in seventy-nine unique datasets [27]. Despite these efforts, several issues and challenges were identified in both projects that severely limit actuarial work in cyber insurance. Challenges identified include a lack of cyber data, information asymmetry, and correlated and interdependent risks that hinder efficient pooling [26]. Another issue identified by Cremer et al. (2022) was that the majority of datasets identified were not on cyber risks but related to the field of intrusion detection and machine learning, making them useful for technical cybersecurity aspects but less useful for actuarial purposes [27].

In terms of measuring catastrophic risk, it was suggested in an expert panel discussion that the economic impact, network effects, and severity are three aspects that make a cyber incident "catastrophic" [23]. Other dimensions include physical manifestations such as property damage and injury, irreversibility, and systemic nature [23]. Both the incident itself and the aftermath of the event can make a cyberattack catastrophic.

The SolarWinds, Colonial Pipeline, and Maersk attacks are a few examples of cyber incidents that caused significant economic and network impacts to critical infrastructure inside the United States. The SolarWinds supply chain attack spread to thousands of computers in both private corporations and U.S. government agencies and departments [28]. The Colonial Pipeline had knock-off effects that impacted multiple sectors, including gas stations across the East Coast [29]. Finally, the world's biggest shipping company, A.P. Moller-Maersk MAERSKb. C.O., experienced a global I.T. outage as a result of the NotPetya cyberattack that disrupted the global supply chain [2]. Seventy-six of the company's ports around the world, including those in the U.S., were affected, and it took several weeks for Maersk to restore its systems [2].

The NotPetya attack demonstrated that cyber catastrophes are not just a problem for the U.S.; they are a looming threat to nations around the world. The 2021 Conti ransomware attack on Ireland's health service was described as "catastrophic" by the head of the Republic of Ireland's health service [30]. The attack was said to have impacted every aspect of patient care, including causing substantial cancellations to outpatient services and encrypting 80% of electronic health records [31]. Recovery from this incident took over four months [31].

Governments abroad recognize the threat of catastrophic cyber incidents and are taking steps to improve their cyber resiliency. For example, a cybersecurity strategy was developed for the energy sector at the request of the European Parliament's Committee on Industry, Research, and Energy as it has become evident that "[t]he immediate and potentially catastrophic nature of the cyber threat across the Energy sector. . . demands an urgent and focused policy response" [32] (p. 8). Furthermore, the Global Cybersecurity Outlook 2023 report published by the World Economic Forum found that 91% of all respondents—global organizational leaders—believe that due to global geopolitical instability "a far-reaching, catastrophic cyber event is at least somewhat likely in the next two years" [33] (p. 8). This includes 86% of business leaders and 93% of cyber leaders [33].

The European Union Agency for Cybersecurity (ENISA) has also issued several publications related to cyber insurance and cyber risk. Though ENISA does not have any publications related to catastrophic cyber risk, the agency has been studying the cyber insurance market—including the barriers to a strong and mature market and ways to incentivize its development—since at least 2012 [34]. ENISA has also studied the lack of commonality in risk assessment language that creates a major obstacle to market growth [35]. They have also surveyed over 260 operators of essential services across twenty-five EU member states to collect their opinions and recommended improvements regarding the contracting of cyber insurance [36]. Most recently, the agency published a study that looks at how machine learning models and artificial intelligence can be used in cyber risk modeling and insurance [37].

Insurance providers have tried to assess catastrophic cyber risk by creating disaster scenarios that could conceivably be considered "catastrophic". For instance, Lloyd's of London developed a realistic attack scenario in which the U.S. power grid was targeted, and a blackout in 15 states occurred for approximately 24 h [38]. In this scenario, multiple companies and insurers are impacted at the same time as effects are spread across several critical infrastructure sectors. In addition to the loss of life, as health and safety systems become inaccessible and fail, property damage of this size is estimated to cause USD 243 billion in losses. In the most extreme version of the scenario, losses could rise to more than USD 1 trillion [38].

The Geneva Association, in their study, explored the insurability of catastrophic cyber risks, utilizing scenario analysis by having its member cyber re/insurers rank potential cyber incidents by the level of concern based on the size of possible extreme cyber losses. A denial-of-service attack, an interruption to operations caused by a worm-like malware epidemic, and a widespread ransomware attack were among the highest-ranked scenarios [24]. Disruptions to critical infrastructure, such as a cyberattack on a key utility provider or a cross-sector IT failure, were also among the scenarios of most concern to cyber re/insurers.

Drawing on insights from the study, it is evident that while the insurers and reinsurers play a pivotal role in mitigating financial losses, their capacity is not unlimited. This underscores the potential benefits of creating a government support mechanism to handle major cyber incidents, bolstering insurers' ability to extend coverage and enhance their risk absorption capabilities [24].

2.3. Past Federal Insurance Programs

The U.S. government has a long history of creating federal insurance programs, such as the National Flood Insurance Program (NFIP), the Federal Crop Insurance Program (FCIP), and the Terrorism Risk Insurance Program (TRIP). These programs are mechanisms for the FIO to provide federal assistance for coverage of catastrophic incidents caused by natural disasters or acts of terrorism. Essentially, the U.S. government becomes an insurer of last resort, taking on the risk that the private insurance market cannot handle through the creation of a backstop. Each program was created for a specific purpose with a unique structure, funding mechanism, and requirements. Reviewing these programs is useful while investigating potential governmental support mechanisms to enhance the insurance sector's capacity to manage extreme cyber risks.

The NFIP began in 1968 to reduce future flood damage and protect property owners in the event of flood disasters. The NFIP is managed by FEMA and covers more than 22,000 communities with nearly 5 million policies and almost USD 1.3 trillion in coverage [39]. Communities are not required to participate in the program, though some homeowners in high-risk flood zones are required to purchase flood insurance. If communities do choose to participate, they are required to adopt land use and control measures with effective enforcement provisions and to regulate the floodplain [39]. The NFIP is funded from premiums, fees, and surcharges paid by NFIP policyholders and from direct annual appropriations. When the balance of the NFIP is insufficient to pay insurance claims, the program can also borrow from the Treasury.

The FCIP began in 1938 and is a central component of the federal farm safety net as it offers farmers coverage against financial losses caused by adverse growing and market conditions. The FCIP is not mandatory but is heavily utilized, with more than 2 million policies sold and USD 116 billion in coverage provided as of 2019 [40]. The FCIP is funded with mandatory appropriations of "such sums as necessary" [41] (p. 1).

The TRIP was established by the Terrorism Risk Insurance Act (TRIA) of 2002 following the 9/11 terrorist attacks when coverage for acts of terrorism became impossible to obtain. TRIP requires insurers to make terrorism risk coverage available within certain lines of commercial P&C insurance. Through the program, certain insurance losses, like those resulting from an "act of terrorism", are eligible for reimbursement [7]. TRIP uses a recoupment mechanism, meaning that if federal payments are made to insurers, the Secretary of the Treasury will collect "terrorism loss risk-spreading premiums" from insurers [7].

When it comes to cyber risk management, the creation of a federal insurance program or backstop is an attractive method for addressing risks that the private market cannot be expected to handle [42]. Vicevich (2018) outlines several potential benefits of creating such a backstop, including improving information-sharing between government and businesses, promoting education by requiring the disclosure of risk to those purchasing insurance, and creating a mechanism to cap private market losses and recoup insurance costs [42].

2.4. Using TRIP for Cyber Catastrophes

The TRIP/TRIA is especially important when discussing catastrophic cyber incident coverage as the Treasury Department has a "longstanding interest in terrorism risk insurance for cyber losses" [7] (p. 58). The Treasury has confirmed that requirements found in the TRIA apply to any policy covering cyber risk written in a line of insurance. This means that the TRIP can be triggered if a policy does not specifically exclude losses arising from a cyber incident. To maintain this possibility, some insurers have worked to create carveouts in their policies for cyber incidents certified by the Treasury as acts of terrorism, as many

war exclusions already extend beyond war to "encompass a range of hostile activity that could include crime, political unrest, and terrorism" [7] (p. 61). Some proponents view the TRIP/TRIA as the remedy for the catastrophic cyber risk issue, advocating for minor legislative tweaks and the inclusion of cyber definitions within the TRIA to eliminate the need for a separate federal cyber risk backstop [43]; however, this approach faces several shortcomings.

For the TRIP to be triggered, the incident must be a certified act of terrorism. Achieving this certification would be extremely difficult for a cyber incident as the TRIA requires that the incident be violent or dangerous to human life, property, or infrastructure. While it is certainly true that cyberattacks can meet these requirements, especially if the attack takes place against critical infrastructure, not all cyberattacks have these types of impacts. Instead, cyberattacks can cause significant business impacts and monetary losses due to exfiltrated data or disruptions to command-and-control systems. These types of behaviors are serious but not necessarily violent or dangerous to human life. Another impediment to triggering the TRIP is that a significant amount of cyber coverage is included in non-standalone insurance policies, which are specifically excluded from the TRIA. Further, experts believe that it would be "legislatively awkward" [44] (p. 50) to add cyber provisions into the TRIA that only apply to catastrophic cyber incidents.

2.5. Prior Research on Catastrophic Risk and Insurance

A review of past research reveals that there is a modest amount of literature in this area, with researchers having previously examined the potential effect of catastrophic cyber incidents and government mechanisms to improve the resiliency of the private insurance market. Researchers have investigated the severity of catastrophic cyber risk in the digital domain, in particular, the potential effects of cyber catastrophes in Internet of Things (IoT) societies or smart cities in which IoT technologies connect industries and sectors, increasing the chance for widespread cyberattacks [45,46]. Other research has focused on identifying the problems in the private insurance market and discussing potential government interventions to combat these problems, including creating new models for cyber war exclusions that no longer result in coverage gaps [47,48], developing safety standards to reduce risky activity through regulation [49], and increasing investment in public funds that directly aim to reduce risks [49]. In an effort to create a consistent economic-impact analysis for cyber-risk scenarios, Eling et al. developed a dynamic inoperability input–output model that allows users to compare the economic impacts of certain scenarios while also accounting for the frequently omitted qualitative context of the hypothetical cyber events [50].

As we discuss in this paper, a commonly proposed government mechanism to combat the escalating threat of cyber catastrophes is the implementation of a federal backstop specific to cyber risk. Previous research has examined the potential benefits and drawbacks of this approach [42] and considered if existing insurance programs and legislation, like the TRIP and the SAFETY Act, could be used as a starting point in developing the backstop [51]. In their paper, Cunningham and Talesh (2021) went as far as to draft the "Catastrophic Cyberattack Resilience Act", a bill that would create a federal backstop for catastrophic cyber events [44]. Since the creation of a federal backstop for cyber risk is not expected soon, short-term solutions have also been discussed, including a proposal for a state solution that includes minimum security requirements for policyholders [52].

2.6. Research Questions

Despite existing research in this area, there are still significant knowledge gaps in the field that this paper seeks to fill. While catastrophic cyber risk has been discussed broadly by researchers, there has been little investigation into the specifics of the issue, such as how to define or measure catastrophic cyber risk or how to mitigate against catastrophic cyber incidents. Furthermore, this paper investigates the capacity of the cyber insurance market and the roles the government can employ to enhance the market's capacity. Our approach

is distinctive in that we answer our research questions utilizing public comments shared by stakeholders that are directly related to the issue. This paper's analysis was guided by the four following research questions:

- What specific factors and conditions elevate a cyber incident to the level of a catastrophe, impacting businesses and society at large?;
- How can the impact of catastrophic cyber incidents be effectively mitigated, and what role does insurance play in this mitigation strategy?;
- Does the current cyber insurance sector possess the necessary capacity to address potential catastrophic cyber incidents adequately?;
- If the current cyber insurance sector lacks sufficient capacity, what roles and methods of support can the government employ to enhance this capacity, and how can these governmental support mechanisms be effectively implemented?

3. Methodology

In this paper, we conducted a mixed-method content analysis of 56 unique comments submitted in response to the Treasury Department's RFI to answer our four research questions. These comments came from stakeholders belonging to entities in critical infrastructure sectors, the insurance industry, and the cybersecurity industry. Private citizens also provided comments.

3.1. Research Design and Approach

Content analysis is a methodological approach associated with the study of inscription contained in varying forms of documentation, such as newspapers, journal articles, and other types of written records [53]. It is a mixed-method analysis that combines qualitative steps (grouping and assignment of text segments to categories) and quantitative steps (analyzing the frequency of concepts and terms) [54]. By systematically analyzing textual material (the comments), we identified the presence and frequency of certain concepts and themes. As a method that allows researchers "to make valid inferences from text" [55] (p. 9), we used this approach to answer our research questions, including how to define and mitigate catastrophic cyber incidents, whether the cyber insurance sector has the capacity to address these incidents adequately, and if the federal government should step in. Figure 1 provides a step-by-step view of conducting content analysis.

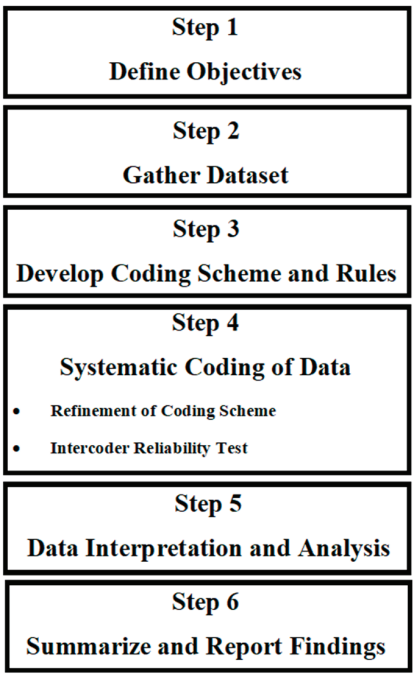


Figure 1. Step-by-step view of content analysis method.

Content analysis has been utilized recently to investigate aspects of the cyber insurance market, including the underwriting process and cyber liability risk. Romanosky et al. (2019) used this method to examine state cyber insurance policies to determine the kinds of losses covered by these policies, how insurers assess risk, and how premiums are determined [13]. Wrede et al. (2020) used this method in conjunction with conducting interviews to examine affirmative and silent coverage in traditional insurance policies for select product lines on the German market [56].

3.2. Data Selection

For the data selection process, we utilized the sample of 60 comments that were submitted in response to the RFI. The dataset comprises comments from major stakeholders in the cybersecurity and insurance industries. The dataset also allows for greater validity and reliability as all comments were written for the same purpose and in the same context, and the views conveyed within the comments are less than two years old. Having this sample also contributed to the trustworthiness of our results as the findings are based on rich, appropriate, and well-saturated data [57] in the comments.

During their review of the comments, coders recognized that some comments were submitted more than once. After removing duplicates, the final sample consisted of 56 unique comments.

3.3. Conducting the Conceptual Analysis

3.3.1. Coding Scheme

In our review of the comments, we used a deductive category structuring. As shown in Figure 2, this meant that coders started with a defined set of categories before reading the comments, with a slight revision to the categories after 20% of the comments were reviewed. We created most of the categories using the questions provided in the RFI, as each RFI question dealt with a different concept related to catastrophic cyber incidents and risk. For example, one of the categories aimed to capture commenter answers to the RFI question about how to define catastrophic cyber incidents. Another category aimed to capture answers to which cybersecurity measures commenters consider most useful in mitigating catastrophic cyber incidents.

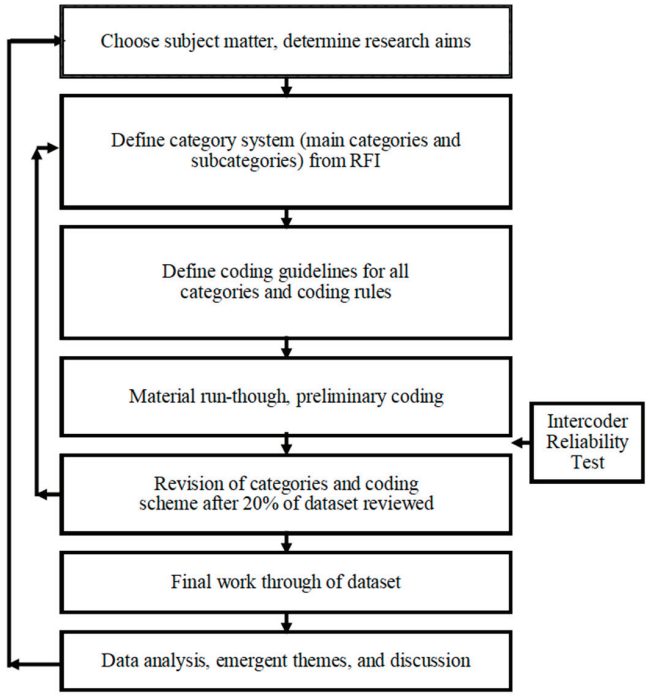


Figure 2. The utilized process model of deductive category application (adapted from [54] p. 319).

In addition to the categories derived from the RFI questions, we also added two categories to code for concerns and regulations. The "concern" category was created to capture any mention of concern or worry about federal government intervention that was conveyed in the comments. The "regulation" category was used to capture any discussion including regulatory language or support of regulation such as the implementation of a federal backstop to improve the capacity of the cyber insurance sector. In total, we outlined 19 different categories to capture opinions conveyed in the comments. Figure 3 shows the 19 distinct categories in this content analysis and indicates which categories we used to answer each research question.

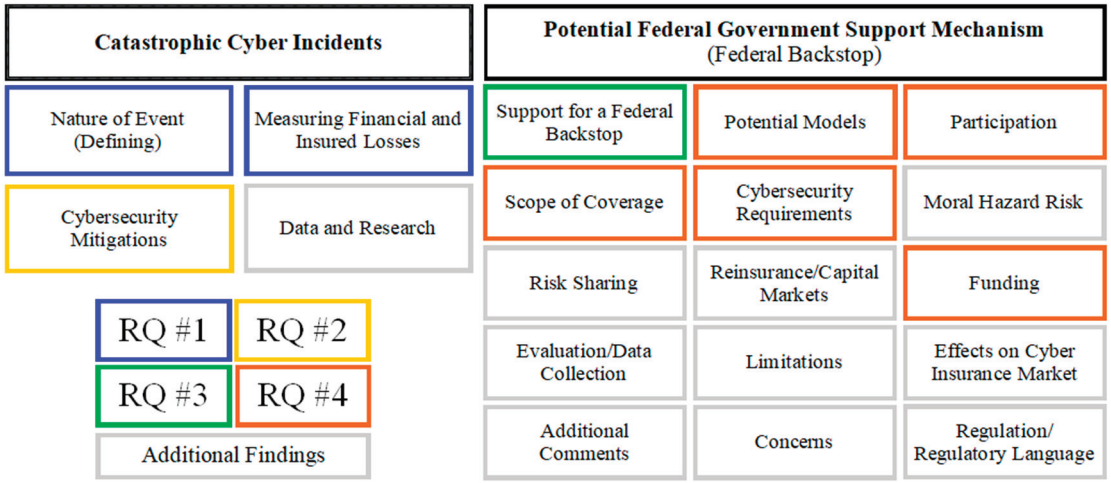


Figure 3. Conceptual mapping.

3.3.2. Intercoder Reliability

To ensure the reliability of the analysis, several tests such as percentage agreement, Holsti's method, Scott's pi (p), Cohen's kappa (k), and Krippendorff's alpha can be used [58]. In this paper, we chose to calculate Cohen's kappa (k) as it measures agreement by considering the agreement observed between coders and what would be expected by chance. Kappa ranges between -1 and 1, with 1 indicating perfect agreement beyond chance, 0 indicating agreement equivalent to chance, and -1 indicating agreement worse than chance. To perform an intercoder reliability test, a minimum of two coders are required, and it is suggested that the second coder code for at least 10% of the full sample [58]. The second coder in our analysis coded 20% of the comments. Our Cohen's kappa (i.e., the ratio of the observed agreement above chance to the maximum possible agreement above chance) was 0.85. This indicates a substantial level of agreement beyond chance between the two coders, and high intercoder reliability. The slight disagreement between coders can likely be attributed to individual coder biases or variations in interpretation due to different backgrounds and knowledge [59].

4. Results

In this section, we discuss the most prominent findings from the content analysis. First, we will start with a review of how we organized our sample and the response rate of each group; then, we answer our four research questions using the responses to specific questions in the Request for Information (RFI).

4.1. Organizing the Sample

The 56 comments were placed into one of five groups—Insurance Providers, Insurance-Affiliated Entities, Critical Infrastructure, Cybersecurity, and Private Citizens—based on the person or entity that submitted the comment. "Insurance Providers" included all comments submitted by insurers, reinsurers, and insurance-related trade associations. "Insurance-Affiliated Entities" included all comments submitted by insurance risk modeling

analytics organizations and insurance think tanks. "Critical infrastructure" included all comments submitted by an entity that worked in or represented one of the sixteen critical infrastructure sectors identified by the Critical Infrastructure and Security Agency (CISA). "Cybersecurity" consists of all comments submitted by entities with a cybersecurity or cyber risk management focus. Finally, "Private Citizens" included all comments submitted by individuals, academic experts, or anonymous sources.

4.1.1. Comment Distribution across Groups

Table 1 shows the breakdown of our sample by group.

Table 1. Comment totals by group.

Group	# of Comments
Insurance Providers	15
Critical Infrastructure	15
Cybersecurity	10
Insurance-Affiliated Entities	8
Private Citizens	8
Total Sample	56

Unsurprisingly, insurers, reinsurers, brokers, and insurance-related trade associations made up one of the two groups with the most comments. Critical Infrastructure also had 15 comments, followed by Cybersecurity, Insurance-Affiliated Entities, and Private Citizens, with 10, 8, and 8 comments, respectively.

Of the comments that fell within the Critical Infrastructure group, 40% were submitted by entities within the healthcare and public health (HPH) sector. Other sectors represented with at least one comment include financial services, energy, communications, government facilities, food and agriculture, and emergency services. There are several reasons why several HPH entities could have submitted comments, but their interest in the topic of catastrophic cyber incidents is likely due to the unprecedented rate of cyberattacks that healthcare entities have endured in recent years. These attacks cost HPH entities millions of dollars and, in some cases, threaten to cause injury and death to patients [60,61].

There was considerable variation in the Private Citizens group as this group included comments submitted anonymously and from academic experts with differing backgrounds. The disciplines and research interests of the experts included cybersecurity policy and research, cyber risk modeling, economics of information security, and cyber insurance. One respondent has over ten years of experience working at a cybersecurity insurance firm that performs risk analysis and modeling, while another has over 20 years of experience working with corporate and military organizations as a cybersecurity management consultant. Most respondents in this group appear to be from the U.S., though at least one of the respondents hails from the United Kingdom.

4.1.2. Variety of Questions Answered and Response Rate

There was considerable variation between the groups regarding which questions they answered. Surprisingly, while Insurance-Affiliated Entities only had eight comments, that group had the highest response rate and answered the widest variety of questions. This is likely due to entities in this group having a combination of specialties, providing services such as insurance risk modeling and cyber insurance analytics, making them familiar with both the insurance and cybersecurity sides of the issue. Table 2 shows the average response rate for each group. The higher the percentage, the more questions that commenters in that group responded to on average. Entities that fell in the Insurance Providers and Insurance-Affiliated Entities groups answered more questions, filling more categories in our coding than comments from the Critical Infrastructure, Cybersecurity, and Private Citizens groups.

Table 2. Average group response rate.

Group	Average Response Rate
Insurance-Affiliated Entities	61%
Insurance Providers	52%
Critical Infrastructure	28%
Cybersecurity	25%
Private Citizens	25%

It should be noted that while the Cybersecurity group had a low response rate overall, it did have the highest rate for the two cybersecurity-focused RFI questions, including the question asking about cybersecurity measures to prevent and mitigate catastrophic cyber incidents.

4.2. Findings

From our coding of the 56 comments, we collected responses from multiple stakeholders to answer our four research questions. Table 3 shows the RFI question categories used predominantly to answer each research question.

Table 3. Corresponding RFI question categories to research questions.

Research Question No.	Request for Information Question Category
1	1, 2
2	3
3	6
4	7

4.2.1. Defining Catastrophic Cyber Incidents (CCIs)

To answer our first research question, we analyzed how commenters responded to questions 1 and 2 of the Request for Information. Question 1 asked, "What type of cyber incidents could have a catastrophic effect on U.S. critical infrastructure? How likely are such incidents? Are particular sectors of U.S. critical infrastructure more susceptible to such incidents? How should the federal government and/or the insurance industry address the potential for cascading, cross-sector impacts from a cyber incident? What type of potential "catastrophic" cyber incident could justify the creation of a federal insurance response?" Question 2 asked, among other things, "What amount of financial losses should be deemed "catastrophic" for purposes of any potential federal insurance response?" Questions 1 and 2 had total average response rates of 55% and 30%, respectively. Question 1 was the fourth most answered question, with most comments coming from the Insurance-Affiliated Entities, Insurance Providers, and Private Citizens groups with response rates of 88%, 73%, and 63%, respectively.

Several comments discussed the difficulty in creating a definition and recognized that being able to estimate the likelihood of a CCI without a clear definition would be nearly impossible. As one commenter put it, "[t]he definition of a catastrophic event in cyber is one of the challenges the industry faces, and it has become one of those 'we will know it when we see it' scenarios" [62] (p. 2). Another commenter went as far as to say that a CCI definition would not be possible without clarification being made to other language, such as "war-like action" and "infrastructure". Nevertheless, several comments did provide terms and definitions that they believed could work for CCIs, including "widespread cyber events", "cyber warfare", and even "war-like actions". It was also recommended that the final definition for "catastrophic" be "consistent with Cyber Incident Reporting for Critical Infrastructure Act of 2022 (CIRCIA) notification and reporting requirements" [63] (p. 1).

Another interesting point was made in regard to which sectors are most at risk. The healthcare and public health sector, financial services sector, and electric grid were all refer-

enced more than once in the responses. Yet, while there appeared to be support for using the term critical infrastructure in the definition, one commenter warned that CCIs should not respond to only critical infrastructure attacks. Additionally, one respondent added that "[w]hile all sectors are potentially vulnerable, those whose legacy infrastructure, 'tech debt', and inability to attract top security talent tend to be more susceptible to successful cyberattacks and incidents" [64] (p. 4).

It was the opinion of one commenter that CCIs be defined through the lens of the effect of a cyber incident rather than its cause "[a] cyber incident resulting in significant, far-reaching insured (and/or uninsured) losses could have a catastrophic impact on the U.S. economy. It is about scale and scope, more than type of loss" [65] (p. 3). In terms of types of cyberattacks that could become CCIs, the attack types that commenters were most worried about included mass outages, distributed denial-of-service attacks, widespread malware attacks, attacks targeting PLCs and SCADA systems, a mass cloud outage, and attacks that specially target the energy grid or supply chain.

Many commenters feel that there are not sufficient data available to estimate the amount of losses that make a cyber incident "catastrophic". Commenters that did suggest trigger points varied significantly in their estimates, ranging from USD 200 million to USD 100 billion. For example, one commenter suggested implementing a trigger amount of aggregate insured losses similar to the threshold for the TRIP—about USD 200 million—so that it is "high enough to encourage the insurance market to manage all but worst-case events" [66] (p. 4). Some commenters also suggested that since "insurers transfer risk to reinsurance in the context of 1 in 200/250 year return periods... anything in excess of this level likely would be 'catastrophic'" [65] (p. 8).

Based on these responses, there appear to be several criteria for defining a catastrophic cyber incident. Consistent with the Treasury Department's definition, commenters stated that a cyber incident could be deemed catastrophic based on the amount of monetary losses, its dispersion across multiple entities and industries, and the degree of critical services affected. In terms of losses, commenters considered looking at both insured loss and overall business loss. Additionally, commenters considered looking at the type of threat actor, that is, if the threat actor is a state actor and the incident is a war-like action. According to some commenters, whether critical infrastructure is impacted and if the orchestrator is a state actor just represent potential indicators and not requirements for an incident to be deemed a catastrophe.

4.2.2. Mitigating Catastrophic Cyber Incidents

To answer our second research question, we analyzed how commenters answered question 3 of the Request for Information which asked in part "What cybersecurity measures would most effectively reduce the likelihood or magnitude of catastrophic cyber incidents?" Question 3 had the third highest response rate of 59%, with responses coming mostly from the Cybersecurity, Insurance Provider, and Insurance-Affiliated Entities groups with response rates of 100%, 73%, and 50%, respectively.

Several different security controls and best practices were discussed in answering how entities could reduce the likelihood and magnitude of a catastrophic cyber incident, including multifactor authentication, mandated employee cyber training, anti-data exfiltration software, endpoint detection and response software, robust patch management, use of backups, and network segmentation. Strategies to improve business resiliency and redundancy were also mentioned, such as supply chain risk management and operational technology (OT) management [67]. Essentially, most commenters felt that implementing a minimum of basic cybersecurity controls was the best way to improve resilience and reduce the magnitude of catastrophic cyber incidents [67,68]. An important point was also made that "...adoption [of cybersecurity measures] is a constant process... Incentives need to be built to encourage regular engagement instead of once and done implementation. Unlike fire, cyber risk is not static" [69] (p. 4).

In answer to the second part of our research question, multiple comments seemed to support using the cyber insurance sector to encourage the adoption of basic cyber hygiene. Already, insurers are requiring their clients to implement and maintain certain security controls in order to receive coverage. For example, Marsh McLennan promotes to their clients a list of 12 key controls that are "most necessary for successful cyber insurance renewals" [64] (p. 7). One commenter even suggested that entities be required to demonstrate compliance with chosen security procedures and practices through periodic third-party assessment [70]. This kind of requirement can motivate organizations who have been slow to adopt such controls as the failure to do so would mean no longer having cyber insurance coverage. Commenters suggested using NIST and ISO/IEC frameworks as resources for organizations looking to achieve high levels of protection.

4.2.3. Capacity of the Cyber Insurance Industry

To answer our third research question, we analyzed how commenters answered question 6 of the Request for Information, which asked "Is a federal insurance response for catastrophic cyber incidents warranted? Why or why not?" Question 6 had the second highest response rate in our analysis at 61%. Of those who provided a comment, 62% said yes, definitively, that they support the federal government's effort to create a federal backstop due to the failing capacity of the private insurance market. A total of 12% stated that they did not support the creation of a federal backstop. The rest of the comments (26%) seemed to offer "cautious support" to the idea of a federal backstop. These comments did not definitively state that they supported the creation of a backstop, but they did posit that a federal government response may be warranted or that a federal response could be the right measure. Table 4 provides a look at the opinions of each group. The percentages shown were created based on the total number of comments from each group that answered the question—not every commenter in each group provided an answer to this question.

Table 4. Support for a federal backstop by group.

Q: Do You Support the Creation of a Federal Backstop?			
Group	**Yes**	**No**	**Cautious Support**
Insurance-Affiliated Entities	87.5%	-	12.5%
Private Citizens	66.7%	-	33.3%
Cybersecurity	66.7%	33.3%	-
Critical Infrastructure	50.0%	-	50.0%
Insurance-Providers	50.0%	25.0%	25.0%

While question 6 may not seem directly related to our research question, responses typically described the current conditions of the private cyber insurance market to explain why or why not a federal backstop was warranted, which does directly answer our third research question. For example, one comment indicated support for a federal backstop because in their view "[i]nsurance companies and reinsurance companies cannot provide sufficient coverage while making a profit" [71] (p. 2). Another commenter said "there could be extraordinary cyber incidents that rise to the level of a catastrophic exposure that could exhaust the capacity of the private market and may warrant a Federal response" [72] (p. 4).

In addition to assessing the industry's capacity to address this risk, we must also consider the industry's intent to cover it. For example, Lloyd's "acknowledges that there are two clear areas in which the cyber insurance market presently has little to no appetite to provide cover—namely, broad infrastructure outages and state backed cyber-attacks" [73] (p. 3). It was clear that many of the commenters felt that the cyber insurance market does not have sufficient capacity or the intent to bear the costs of a catastrophic cyber incident.

Answers to other questions examined during this analysis also provided useful responses to our research question. In response to insurance coverage availability, one commenter stated that "some insurers have sought to reduce coverage for incidents deemed catastrophic, either through sub-limits or co-insurance" and the use of exclusions or narrow

definitions [74] (p. 12). Another part of the problem, according to one commenter, is the fact that "[t]here has also been little movement in establishing a retro reinsurance market for cyber catastrophe risk", leaving catastrophic cyber incident coverage to the traditional reinsurance market [75] (p. 11). Overall, the ability to "provide the coverage required to keep U.S. companies afloat in the time of a catastrophic event is critical but difficult for the private insurance market" [76] (p. 4).

4.2.4. Potential Governmental Support Mechanism

To answer our fourth and final research question, we analyzed how commenters addressed questions related to the structuring of a federal backstop for catastrophic cyber incidents, including sub-questions 7.1 and 7.4. Creating a federal backstop was widely supported by multiple stakeholders as a government method to enhance the cyber insurance sector's capacity. By establishing a backstop, the government, as it has done in the past with the FCIP, NFIP, and TRIP, can take on the risk that is causing uncertainty and hardening of the cyber insurance market and create a mechanism for mandating certain cybersecurity controls and best practices.

Question 7.1 asked, "Should an existing federal insurance program (e.g., NFIP or TRIP) or other U.S. or international public-private insurance mechanism serve as a model for, or be modified to address catastrophic cyber incidents?" There appeared to be broad agreement that the TRIP could serve as a model for a backstop for CCIs. More specifically, of those comments that answered the question (63%), 43% of comments either supported using the TRIP as a model for a new backstop (31%) or recommended that the TRIP be extended to include CCIs (17%). Only 20% of comments recommended against using the TRIP as a model and cited several reasons why. Reasons included issues with the "mandatory make available" aspect of the TRIP structure, the uncertainty around the attribution of cyberattacks, covering cyberattacks under the definition of terrorism, and the inability of the TRIP to "to financially cover the projected catastrophic financial losses arising from certified cyber acts of terrorism" [77] (p. 3). The rest of the comments (31%) that provided a response took the opportunity to recommend several different models other than the TRIP that could be good reference points when the time comes to create a federal backstop for CCIs. These models include the NFIP, the United Kingdom's Pool Re, the Federal Deposit Insurance Corporation (U.S.), and a Capital Markets solution. Table 5 shows how each group responded to this question.

Table 5. Potential models for a federal backstop by group.

Q: Do You Support Using TRIP as a Model for a New Backstop for CCIs?				
Group	**Yes**	**Extend TRIP to CCIs**	**No**	**Mentions/Endorses Some Other Model**
Critical Infrastructure	55.6%	-	22.2%	22.2%
Insurance-Affiliated Entities	37.5%	25.0%	12.5%	25.0%
Private Citizens	33.3%	33.3%	-	33.3%
Insurance Providers	16.7%	16.7%	25.0%	41.7%
Cybersecurity	-	33.3%	33.3%	33.3%

Question 7.4 asked "Should cybersecurity and/or cyber hygiene measures be required of policyholders under the structure? If so, which measures should be required?" This question had a response rate of 59%. Of those who answered, the majority agreed that cyber measures should be required (66%) for policyholders. Only 7% (two comments) said cybersecurity measures should not be required, while the rest of the commenters (28%) did not make a clear choice about requiring the measures; they still discussed their use. Table 6 shows responses to question 7.4 by group.

Table 6. Opinion on cyber hygiene eligibility requirement by group.

Q: Do You Support Using TRIP as a Model for a New Backstop for CCIs?			
Group	**Yes**	**No**	**Does Not Explicitly Say**
Private Citizens	100%	-	-
Cybersecurity	87.5%	-	12.5%
Insurance-Affiliated Entities	66.7%	-	33.3%
Critical Infrastructure	50.0%	-	50.0%
Insurance Providers	42.8%	28.6%	28.6%

Many of these commenters supported baseline cybersecurity requirements or a "back to basics" approach like that discussed in Section 4.2.2. Again, some recommended measures that were referenced in the comments included mandatory multifactor authentication, regular patching, backups, mobile device security, employee cyber training, anti-data exfiltration, complex passwords, developed incident response plans, and endpoint detection and response software.

According to one commenter, "[t]he most efficient and effective means for the federal government to establish and enforce cybersecurity measures would be to do so directly" [78] (p. 9). This could be achieved by implementing an eligibility requirement for any federal insurance program the government creates. Implementing and maintaining certain best practices would then be an expectation for access to any federal program or solution [65].

Only two comments, both in the Insurance Providers group, stated that cybersecurity measures should not be required through the program. They reasoned that the threat landscape is changing so fast that requiring control measures would be impractical and very challenging. One comment further stated that "requiring particular cyber hygiene measures to be met, could, in some instances quickly prove to be out of date and become burdensome to amend" [64] (p. 17).

5. Concerns and Benefits of a Federal Backstop

Here we summarize the concerns and benefits shared by multiple stakeholders regarding the federal government potentially stepping in to solve the hardening of the cyber insurance sector due to the risk of catastrophic cyber incidents.

5.1. Concerns Surrounding a Federal Backstop

In addition to capturing respondents' answers to the RFI questions, we coded for respondents' concerns for the backstop. Some of these were captured within the RFI questions asking about moral hazard and limitations to a backstop, while the rest were captured in our "concerns" and "regulation" categories. It was clear throughout our analysis that many commenters have worries about the backstop and how it may or may not be structured. The top concerns related to the backstop include the moral hazard problem, the coverage of physical losses, and the fear that a backstop would result in restrictive regulation. Fortunately, there are solutions available to ease these concerns.

The moral hazard problem is the phenomenon in which those who are insured take advantage of the fact that they have insurance and engage in dangerous or risky behavior that they otherwise would not. In this circumstance, some commenters fear that because insured entities would be covered from financial losses through the backstop, they will not maintain the appropriate level of cybersecurity and resilience. Fortunately, there are ways to ensure that insured entities do not have an undue reliance on the backstop. One strategy that many commenters support is deciding if an entity is eligible to participate in the federal insurance program or backstop based on their adoption and maintenance of certain cybersecurity controls and risk management practices. Another strategy discussed by a commenter was to utilize "public transparency" with respect to the backstop. If all products backed by a federal program were required to be filed with state regulators, then public review would help to prevent any exploitation by those participating [78].

Another concern surrounding the backstop was whether physical losses resulting from a CCI could or would be covered by the backstop. Often, cyberattacks, especially those that are widespread, result in physical damage, whether direct damage to computer hardware and systems or indirect damage to infrastructure due to the loss of communication and internet services. For example, in the 2010 Stuxnet attack, a malicious computer worm caused substantial damage to an Iranian uranium enrichment plant, as the worm was able to change the centrifuges' rotor speed so excessively that the change in force destroyed the centrifuges [79]. The traditional cyber insurance market often delineates between physical and cyber damages and, in some cases, may not cover financial losses from physical damage. Yet, some commenters were hopeful that a program for CCIs could in fact provide a backstop to events with both kinds of damages, essentially "bridging the gap" between cyber insurance and traditional insurance.

A final concern conveyed by the commenters was the potential for a federal insurance program to add more regulatory requirements to the sector that prove to be burdensome. That is, if the backstop is not focused where it is needed most, it could end up hindering the competitive nature of the private market [80]. These views were unsurprising as past discussions of the backstop by major insurers also conveyed a worry that potential strings could be attached to the federal program, such as the backstop enabling the federal government to dictate insurers' capacity. This includes Jeremy Gittler's remarks shared at the 2023 Professional Liability Underwriting Society's cyber symposium. Gittler, who is head of cyber and technology-Americas at AXA XL, stated that "[w]e don't want a situation where they're saying . . . you can only deploy this much capacity, you have to get this much reinsurance" [81] (para. 3). Though Gittler later goes on to support the idea of a "true" federal backstop (one where the government only comes in and takes over after losses hit a specific number so that the insurance industry can survive), his comments and some of the comments examined in this analysis show that there is fear that a federal backstop will enable the government to step in and stifle innovation, impose costly burdens, and otherwise lessen the power the private market currently has [81]. To alleviate this fear, the federal government will have to engage multiple stakeholders, ranging from private insurers to sector representatives (e.g., sector risk management agencies), to comprehensively grasp and accommodate the significant distinctions existing among different industry sectors. An inclusive approach can help enable the development of strategies that are specifically attuned to the diverse cybersecurity and insurance requirements within each sector.

5.2. Potential Benefits of a Federal Backstop

The benefits of a federal backstop were discussed throughout most of the comments in this analysis, particularly in responses to RFI question 8, which asked about the potential effects of a federal backstop on the cyber insurance market. A federal backstop could raise the level of cybersecurity of participating entities, help insurers avoid litigation, and remove significant uncertainty in the market.

Arguably, the most significant advantage that can arise from the creation of a new backstop for CCIs is the potential to raise the overall cybersecurity and cyber resiliency of U.S. entities to a higher level. If a minimum level of cybersecurity was required for participation in the program, then more entities would take the time and resources to ensure their level of security is adequate enough to meet eligibility requirements. With greater cyber hygiene, the severity of future CCIs is likely lowered, and fewer entities would fall victim to attacks.

With the creation of a federal backstop, U.S. entities, insurance, and reinsurance companies will also be relieved of the burden of assuming CCI risk. One commenter indicated that if a CCI were to happen today, the attack would not only cause devasting losses due to a lack of cybersecurity posture in the U.S. but many insurers would find themselves dealing with significant litigation [82]. As we saw following the NotPetya malware attack, Mondelez International engaged in a four-year legal battle with Zurich American Insurance after the insurer denied covering Mondelez's NotPetya losses. With a

backstop in place to cover losses from CCIs, insurers will be relieved of carrying the risk and companies can avoid this kind of time-consuming and costly litigation.

Finally, a federal backstop for CCIs would have positive effects on the availability and affordability of the cyber insurance market as it could increase insurers' willingness to provide cyber coverage. By reducing the uncertainty that exists today in the cyber insurance market, insurers will be more willing to reduce prices and premiums and increase the size of the limits that those insured can access. Stabilizing the cyber insurance market could help insurers provide better and more comprehensive coverage to policyholders overall.

6. Picturing the Backstop

Likely, the most important aspect of this analysis is that it can tell us what a federal backstop for catastrophic cyber incidents could look like using the attitudes captured in the comments. Based on the aspects of the structure that had the greatest consensus, the base structure of a backstop for CCIs would start out looking similar to the TRIP. The backstop will most likely have cybersecurity requirements associated with participation, whether or not participation is mandatory or voluntary. Finally, while there is not yet a finalized definition for CCIs, attacks on critical infrastructure sectors, entities domiciled in the U.S., and the supply chain appear to be targets that would fall within the scope of coverage of the backstop as attacks on these targets would have a widespread impact on multiple entities and industries. Additionally, these types of attacks would result in significant losses, which, according to commenters, is another important criterion for deciding if a cyber incident reaches the level of a catastrophe.

While there was some variation in the comments regarding the need for a federal insurance response, the considerable potential benefits accompanying the creation of a federal backstop appear sufficient to overcome any current disagreement and there are strategies to combat concerns. For example, making the adoption and maintenance of cybersecurity measures and practices mandatory for participants will help to combat the moral hazard problem. When it comes to the coverage of physical losses and fear of restrictive regulation, there is plenty of time for public and private stakeholders to work with the government to structure the program in such a way that all parties are satisfied. By taking into consideration the answers and suggestions left in the comments submitted in response to the Treasury's RFI, the government will be well on its way to creating and implementing a federal backstop for catastrophic cyber incidents that not only protects and stabilizes the insurance industry but increases the level of cyber resilience in the United States.

7. Conclusions

By conducting a multistakeholder analysis using the comments submitted in response to the Treasury Department's Request for Information, we were able to capture the thoughts and opinions of 56 different stakeholders to help answer our four research questions related to defining and mitigating catastrophic cyber incidents, assessing the capacity of the cyber insurance sector to address catastrophic cyber incidents and how the government could enhance the capacity of the private cyber insurance sector. Table 7 summarizes the research questions explored along with the corresponding findings.

It was clear from our analysis of the comments that there is considerable variation in how commenters would define, measure, and mitigate catastrophic cyber incidents, though their responses provided helpful insights to use as starting points. It was also made clear by commenters that the cyber insurance market does not have the capacity to bear the losses associated with a catastrophic cyber incident. For this reason, most commenters support the government stepping in to create a backstop for catastrophic cyber incidents to enhance the industry's capacity. Yet, before the government can begin creating a backstop, finalized definitions for a "catastrophic cyber incident" and other related terms must be developed to remove uncertainty from the market and give the stakeholders a clearer path forward. Without a clear definition for "catastrophic", it is difficult to say definitively what

the best strategies for mitigating catastrophic cyber risk are and what level of losses should trigger a government response.

Table 7. Research questions and findings on catastrophic cyber incidents and insurance sector responses.

Problem: Catastrophic cyber incidents—events of low probability but high impact, with the potential to incur billions in damages—are prompting insurers to elevate premiums, create higher barriers for potential buyers, and tighten policies with exclusions. These responses have led to a notable gap in market protection.	
Method: Content analysis of 56 unique comments submitted in response to the Treasury Department Request for Information. The answers commenters provided to the RFI questions were used to answer our research questions.	
Research Questions	Findings
What specific factors and conditions elevate a cyber incident to the level of a catastrophe, impacting businesses and society at large?	• Suggested definition criteria: commenters stated that a cyber incident could be deemed catastrophic based on the total amount of business or insured losses, its dispersion across multiple entities and industries, the degree of critical services affected, and the type of threat actor responsible.
How can the impact of catastrophic cyber incidents be effectively mitigated, and what role does insurance play in this mitigation strategy?	• Suggested mitigations: Multifactor authentication, employee cyber training, anti-data exfiltration software, endpoint detection and response software, robust patch management, data backups, and network segmentation. • The insurance industry can continue to require their policyholders meet a base level of cybersecurity to maintain their eligibility.
Does the current cyber insurance sector possess the necessary capacity to address potential catastrophic cyber incidents adequately?	• It was the opinion of several commenters that the private insurance industry does not have the want or the capacity to address catastrophic cyber risk adequately. • Of those who answered the RFI question, 62% said yes, definitively, that they support the federal government stepping in to create a federal backstop for catastrophic cyber incidents.
If the current cyber insurance sector lacks sufficient capacity, what roles and methods of support can the government employ to enhance this capacity, and how can these governmental support mechanisms be effectively implemented?	• Commenters largely support the federal government stepping in to create a federal backstop for catastrophic cyber incidents. • There was broad agreement that the Terrorism Risk Insurance Program could serve as a model for the backstop. • Most commenters felt that a base level of cybersecurity should be required of entities looking to be eligible for the backstop. Adding this requirement would help to combat a potential moral hazard problem.

This study holds significance given the dynamic cyber threat landscape and increased risk of catastrophic cyber incidents in an unstable geopolitical climate. The government's potential involvement in establishing a backstop mechanism for cyber insurance carries costs and benefits. While it can boost economic resilience and cybersecurity of the nation, the backstop will need to be carefully structured to avoid the moral hazard problem and burdensome cybersecurity requirements and other restrictions that could impede innovation and competition in the cyber insurance sector. Additionally, the research findings, which represent one of the initial data-driven studies on the subject, offer insights from diverse stakeholders, including the public and private sectors. These insights serve as crucial input for policymakers in addressing the issue effectively. The federal government should continue to engage with stakeholders from different industries and sectors, including private insurance entities and cybersecurity experts to study this issue.

Author Contributions: Conceptualization, U.T.; methodology, B.B., E.D. and U.T.; investigation, B.B. and U.T.; formal analysis, B.B.; writing—original draft preparation, B.B., E.D. and U.T.; writing—review and editing, B.B. and U.T.; supervision, U.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Acknowledgments: We would like to express our deepest gratitude to the reviewers for their invaluable feedback and insightful comments, which greatly enhanced the quality of this work.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Beer, J. "WannaCry" Ransomware Attack Losses Could Reach $4 Billion. *CBS News*. 2017. Available online: https://www.cbsnews.com/news/wannacry-ransomware-attacks-wannacry-virus-losses/ (accessed on 6 September 2023).
2. Greenberg, A. The Untold Story of NotPetya, The Most Devastating Cyberattack in History. *Wired*. 2018. Available online: https://www.wired.com/story/notpetya-cyberattack-ukraine-russia-code-crashed-the-world/ (accessed on 6 September 2023).
3. Kreese, B. The Insurance Market Is Hardening: What Does That Mean for Your Business? *Buffalo Business First*. 2021. Available online: https://www.bizjournals.com/buffalo/news/2021/01/25/the-insurance-market-is-hardening-what-does-that.html (accessed on 6 September 2023).
4. GAO (U.S. Government Accountability Office). Cyber Insurance: Action Needed to Assess Potential Federal Response to Catastrophic Attacks. (GAO-22-104256). 2022. Available online: https://www.gao.gov/products/gao-22-104256 (accessed on 6 September 2023).
5. U.S. Treasury Department. Potential Federal Insurance Response to Catastrophic Cyber Incidents, 87 FR 59161. *Federal Register*. 2022. Available online: https://www.federalregister.gov/documents/2022/11/09/2022-24476/potential-federal-insurance-response-to-catastrophic-cyber-incidents (accessed on 6 September 2023).
6. Marotta, A.; Martinelli, F.; Nanni, S.; Orlando, A.; Yautsiukhin, A. Cyber-insurance survey. *Comput. Sci. Rev.* **2017**, *24*, 35–61. [CrossRef]
7. FIO (Federal Insurance Office). Report on the Effectiveness of the Terrorism Risk Insurance Program. U.S. Department of the Treasury. 2022. Available online: https://home.treasury.gov/system/files/311/2022%20Program%20Effectiveness%20Report%20(FINAL).pdf (accessed on 6 September 2023).
8. Woods, D.; Simpson, A. Policy measures and cyber insurance: A framework. *J. Cyber Policy* **2017**, *2*, 209–226. [CrossRef]
9. Xie, X.; Lee, C.; Eling, M. Cyber insurance offering and performance: An analysis of the U.S. cyber insurance market. *Geneva Pap. Risk Insur. Issues Pract.* **2020**, *45*, 690–736. [CrossRef]
10. Baker, T.; Shortland, A. Insurance and enterprise: Cyber insurance for ransomware. *Geneva Pap. Risk Insur. Issues Pract.* **2022**, *48*, 275–299. [CrossRef]
11. Sophos. The Critical Role of Frontline Cyber Defenses in Cyber Insurance Adoption. [Whitepaper]. 2023. Available online: https://assets.sophos.com/X24WTUEQ/at/qmqkh63jxfbpmtpfftrxsnq/sophos-cyber-insurance-adoption-survey-2023-wp.pdf (accessed on 14 November 2023).
12. Falco, G.; Eling, M.; Jablanski, D.; Miller, V.; Gordon, L.A.; Wang, S.S.; Schmit, J.; Thomas, R.; Elvedi, M.; Maillart, T.; et al. A Research Agenda for Cyber Risk and Cyber Insurance. In Proceedings of the Workshop on the Economics of Information Security (WEIS), Boston, MA, USA, 3–4 June 2019. Available online: https://cyber.fsi.stanford.edu/publication/research-agenda-cyber-risk-and-cyber-insurance (accessed on 6 September 2023).
13. Romanosky, S.; Ablon, L.; Kuehn, A.; Jones, T. Content analysis of cyber insurance policies: How do carriers price cyber risk? *J. Cybersecur.* **2019**, *5*, tyz002. [CrossRef]
14. Tsohou, A.; Diamantopoulou, V.; Gritzalis, S.; Lambrinoudakis, C. Cyber insurance: State of the art, trends and future directions. *Int. J. Inf. Secur.* **2023**, *22*, 737–748. [CrossRef] [PubMed]
15. Cowbell Cyber. Survey Results: The Economic Impact of Cyber Insurance (Small and Mid-Size Enterprises in the U.S.). 2020. Available online: https://cowbell.insure/wp-content/uploads/2020/06/Cowbell-Cyber-data-report.pdf (accessed on 14 November 2023).
16. Morgan, S. Cybercrime to Cost the World $10.5 Trillion Annually by 2025. *Cybercrime Magazine*. 2020. Available online: https://cybersecurityventures.com/cybercrime-damage-costs-10-trillion-by-2025/ (accessed on 6 September 2023).
17. Biener, C.; Eling, M.; Wirfs, J.H. Insurability of Cyber Risk: An Empirical Analysis. *Geneva Pap. Risk Insur. Issues Pract.* **2015**, *40*, 131–158. [CrossRef]
18. Kshetri, N. The evolution of cyber-insurance industry and market: An institutional analysis. *Telecommun. Policy* **2020**, *44*, 102007. [CrossRef]

19. Mondelez, V. Zurich (Mondelez International Inc. v. Zurich American Insurance Co.), No. 2018L11008 (Circuit Court of Cook County, Illinois). 2018. Available online: https://regmedia.co.uk/2022/11/02/pacer_mondelez_zurich_complaint.pdf (accessed on 6 September 2023).
20. Ferland, J. Cyber insurance—What coverage in case of an alleged act of War? Questions raised by the Mondelez v. Zurich case. *Comput. Law Secur. Rev.* **2019**, *35*, 369–376. [CrossRef]
21. Tatar, U.; Nussbaum, B.; Gokce, Y.; Keskin, O.F. Digital force majeure: The Mondelez case, insurance, and the (un)certainty of attribution in cyberattacks. *Bus. Horiz.* **2021**, *64*, 775–785. [CrossRef]
22. Department of Financial Services. Insurance Circular Letter No. 2. *New York State*. 2021. Available online: https://www.dfs.ny.gov/industry_guidance/circular_letters/cl2021_02 (accessed on 7 September 2023).
23. Tatar, U.; Nussbaum, B.; Keskin, O.F.; Dubois, E.; Foti, D.; Setting the Scene: Framing Catastrophic Cyber Risk An Expert Panel Discussion Part 1. The Society of Actuaries Research Institute. 2023. Available online: https://www.soa.org/resources/research-reports/2023/cat-cyber-risk/ (accessed on 6 September 2023).
24. The Geneva Association. Cyber Risk Accumulation: Fully Tackling the Insurability Challenge. 2023. Available online: https://www.genevaassociation.org/sites/default/files/2023-11/cyber_accumulation_report_91123.pdf (accessed on 14 November 2023).
25. CyberCube. Designing a Cyber Catastrophe: A Guide to the Thought Process behind Creating Cyber Disaster Scenarios. 2020. Available online: https://insights.cybcube.com/en/a-guide-to-designing-scenario-narratives-for-cyber-catastrophe (accessed on 7 September 2023).
26. Dubois, E.V.; Keskin, O.F.; Tatar, U. Cyber Risk Modeling Methods and Data Sets. SOA. 2022. Available online: https://www.soa.org/4a81c2/globalassets/assets/files/resources/research-report/2022/cyber-risk-modeling.pdf (accessed on 19 June 2024).
27. Cremer, F.; Sheehan, B.; Fortmann, M.; Kia, A.N.; Mullins, M.; Murphy, F.; Materne, S. Cyber risk and cybersecurity: A systematic review of data availability. *Geneva Pap. Risk Insur. Issues Pract.* **2022**, *47*, 698–736. [CrossRef] [PubMed]
28. Sylvester, J. Two Years Later: An Analysis of SolarWinds and the Impact on the Cyber Insurance Industry. Gallagher USA. 2022. Available online: https://www.ajg.com/us/news-and-insights/2022/aug/two-years-later-an-analysis-of-solarwinds-and-the-impact-on-the-cyber-insurance-industry/ (accessed on 25 September 2023).
29. CISA (The Cybersecurity and Infrastructure Security Agency). The Attack on Colonial Pipeline: What We've Learned & What We've Done over the Past Two Years [Blog]. 2023. Available online: https://www.cisa.gov/news-events/news/attack-colonial-pipeline-what-weve-learned-what-weve-done-over-past-two-years (accessed on 25 September 2023).
30. BBC. Cyber-Attack on Irish Health Service 'Catastrophic'. 2021. Available online: https://www.bbc.com/news/world-europe-57184977 (accessed on 19 June 2024).
31. PwC. Conti Cyber Attack on the HSE: Independent Post Incident Review. 2021. Available online: https://www.hse.ie/eng/services/publications/conti-cyber-attack-on-the-hse-full-report.pdf (accessed on 19 June 2024).
32. European Parliament. Cyber Security Strategy for the Energy Sector [Study]. 2016. Available online: https://www.europarl.europa.eu/RegData/etudes/STUD/2016/587333/IPOL_STU(2016)587333_EN.pdf (accessed on 19 June 2024).
33. World Economic Forum. Global Cybersecurity Outlook 2023. 2023. Available online: https://www3.weforum.org/docs/WEF_Global_Security_Outlook_Report_2023.pdf (accessed on 19 June 2024).
34. ENISA; Robinson, N.; RAND Europe. Incentives and Barriers of the Cyber Insurance Market in Europe. 2012. Available online: https://www.enisa.europa.eu/publications/incentives-and-barriers-of-the-cyber-insurance-market-in-europe (accessed on 19 June 2024).
35. ENISA. Cyber Insurance—Models and Methods and the Use of AI. 2024. Available online: https://www.enisa.europa.eu/publications/cyber-insurance-models-and-methods-and-the-use-of-ai (accessed on 19 June 2024).
36. ENISA. Commonality of Risk Assessment Language in Cyber Insurance. 2017. Available online: https://www.enisa.europa.eu/publications/commonality-of-risk-assessment-language-in-cyber-insurance (accessed on 19 June 2024).
37. ENISA. Demand Side of Cyber Insurance in the EU. 2023. Available online: https://www.enisa.europa.eu/publications/demand-side-of-cyber-insurance-in-the-eu (accessed on 19 June 2024).
38. Lloyd's of London. Business Blackout: The Insurance Implications of a Cyber Attack on the U.S. Power Grid. 2015. Available online: https://www.lloyds.com/news-and-insights/risk-reports/library/business-blackout/ (accessed on 7 September 2023).
39. Congressional Research Service. A Brief Introduction to the National Flood Insurance Program. 2023. Available online: https://crsreports.congress.gov/product/pdf/IF/IF10988 (accessed on 6 September 2023).
40. Congressional Research Service. Federal Crop Insurance: A Primer. 2021. Available online: https://crsreports.congress.gov/product/pdf/R/R46686 (accessed on 6 September 2023).
41. Congressional Research Service. Farm Bill Primer: Federal Crop Insurance Program. 2022. Available online: https://crsreports.congress.gov/product/pdf/IF/IF12201 (accessed on 6 September 2023).
42. Vicevich, D.L. The Case for a Federal Cyber Insurance Program. *Neb. L. Rev.* **2018**, *97*, 555. Available online: https://digitalcommons.unl.edu/nlr/vol97/iss2/7 (accessed on 19 June 2024).
43. Bace, B. The Insurer of Last Resort: Investigating a Federal Insurance Backstop for Catastrophic Cyber Incidents. 2023. Available online: https://scholarsarchive.library.albany.edu/honorscollege_pos/43/ (accessed on 7 September 2023).
44. Cunningham, B.; Talesh, S.A. Uncle Sam RE: Improving Cyber Hygiene and Increasing Confidence in the Cyber Insurance Ecosystem via Government Backstopping. *Conn. Insur. Law J* **2021**, *28*, 1–84. Available online: https://cilj.law.uconn.edu/wp-content/uploads/sites/2520/2022/10/CILJ-Vol.-28.1.pdf (accessed on 6 September 2023).

45. Pal, P.; Huang, Z.; Yin, X.; Liu, M.; Lototsky, S.; Crowcroft, J. Sustainable Catastrophic Cyber-Risk Management in IoT Societies. In Proceedings of the 2020 Winter Simulation Conference (WSC), Orlando, FL, USA, 14–18 December 2020; pp. 3105–3116. [CrossRef]
46. Pal, R.; Huang, Z.; Lototsky, S.; Yin, X.; Liu, M.; Crowcroft, J.; Sastry, N.; De, S.; Nag, B. Will Catastrophic Cyber-Risk Aggregation Thrive in the IoT Age? A Cautionary Economics Tale for (Re-)Insurers and Likes. *ACM Trans. Manag. Inf. Syst.* **2021**, *12*, 17. [CrossRef]
47. Cremer, F.; Sheehan, B.; Mullins, M.; Fortmann, M.; Ryan, B.J.; Materne, S. On the insurability of cyber warfare: An investigation into the German cyber insurance market. *Comput. Secur.* **2024**, *142*, 103886. [CrossRef]
48. Bateman, J. War, Terrorism, and Catastrophe in Cyber Insurance: Understanding and Reforming Exclusions. Carnegie Endowment. 2020. Available online: https://carnegieendowment.org/research/2020/10/war-terrorism-and-catastrophe-in-cyber-insurance-understanding-and-reforming-exclusions?lang=en (accessed on 25 June 2024).
49. Baker, T.; Shortland, A. The government behind insurance governance: Lessons for ransomware. *Regul. Gov.* **2023**, *17*, 1000–1020. [CrossRef]
50. Eling, M.; Elvedi, M.; Falco, G. The Economic Impact of Extreme Cyber Risk Scenarios. *N. Am. Actuar. J.* **2023**, *27*, 429–443. [CrossRef]
51. Knake, R.K. Creating a Federally Sponsored Cyber Insurance Program. Council on Foreign Relations. 2016. Available online: https://www.cfr.org/report/creating-federally-sponsored-cyber-insurance-program (accessed on 25 June 2024).
52. Sayre, M. Impossible Math: The Need for Government-Backed Cyber Insurance. *Tort Trial Insur. Pract. Law J.* **2022**. Available online: https://ssrn.com/abstract=4699671 (accessed on 25 June 2024).
53. Prior, L. Content Analyses. In *The Oxford Handbook of Qualitative Research*, 1st ed.; Leavy, P., Ed.; Oxford University Press: Oxford, UK, 2014; p. 359, ISBN 9780199811755.
54. Mayring, P.A.E. Qualitative content analysis. In *International Encyclopedia of Education*, 4th ed.; Tierney, R.J., Rizvi, F., Ercikan, K., Eds.; Elsevier Inc.: Amsterdam, The Netherlands, 2023; pp. 314–322. [CrossRef]
55. Weber, R. *Basic Content Analysis*, 2nd ed.; Sage: Newbury Park, CA, USA, 1990.
56. Wrede, D.; Stegen, T.; Von der Schulenburg, J.M.G. Affirmative and silent cyber coverage in traditional insurance policies: Qualitative content analysis of selected insurance products from the German insurance market. *Geneva Pap. Risk Insur. Issues Pract.* **2020**, *45*, 657–689. Available online: https://link.springer.com/article/10.1057/s41288-020-00183-6 (accessed on 8 September 2023). [CrossRef]
57. Elo, S.; Kääriäinen, M.; Kanste, O.; Pölkki, T.; Utriainen, K.; Kyngäs, H. Qualitative Content Analysis: A Focus on Trustworthiness. *SAGE Open* **2014**, 4. [CrossRef]
58. Lombard, M.; Snyder-Duch, J.; Bracken, C.C. Practical Resources for Assessing and Reporting Intercoder Reliability in Content Analysis Research Projects. 2005. Available online: https://www.researchgate.net/publication/242785900 (accessed on 25 June 2024).
59. Drisko, J.W.; Maschi, T. *Content Analysis*; Oxford University Press: Oxford, UK, 2016; pp. 81–121.
60. IBM Security & Ponemon Institute. Cost of Data Breach Report 2023. IBM. 2023. Available online: https://www.ibm.com/downloads/cas/E3G5JMBP (accessed on 8 September 2023).
61. Miller, M. The Mounting Death Toll of Hospital Cyberattacks. Politico. 2022. Available online: https://www.politico.com/news/2022/12/28/cyberattacks-u-s-hospitals-00075638 (accessed on 8 September 2023).
62. Association of Bermuda Insurers and Reinsurers. Comment from Association of Bermuda Insurers and Reinsurers. TREAS-DO-2022-0019-0042. 2022. Available online: https://www.regulations.gov/comment/TREAS-DO-2022-0019-0042 (accessed on 15 November 2023).
63. Berger, M. Comment from Berger, Mitchell. TREAS-DO-2022-0019-0010. Regulations.gov. 2022. Available online: https://www.regulations.gov/comment/TREAS-DO-2022-0019-0010 (accessed on 15 November 2023).
64. McLennan, M. Comment from Marsh McLennan. TREAS-DO-2022-0019-0025. Regulations.gov. 2022. Available online: https://www.regulations.gov/comment/TREAS-DO-2022-0019-0025 (accessed on 15 November 2023).
65. Reinsurance Association of America. Comment from Reinsurance Association of America. TREAS-DO-2022-0019-0028. Regulations.gov. 2022. Available online: https://www.regulations.gov/comment/TREAS-DO-2022-0019-0028 (accessed on 15 November 2023).
66. Axio. Comment from Axio. TREAS-DO-2022-0019-0017. Regulations.gov. 2022. Available online: https://www.regulations.gov/comment/TREAS-DO-2022-0019-0017 (accessed on 15 November 2023).
67. American Property Casualty Insurance Association. Comment from American Property Casualty Insurance Association (APCIA). TREAS-DO-2022-0019-0050. Regulations.gov. 2022. Available online: https://www.regulations.gov/comment/TREAS-DO-2022-0019-0050 (accessed on 15 November 2023).
68. CyberCube Analytics Inc. Comment from CyberCube Analytics Inc. TREAS-DO-2022-0019-0029. Regulations.gov. 2022. Available online: https://www.regulations.gov/comment/TREAS-DO-2022-0019-0029 (accessed on 15 November 2023).
69. Zurich North America. Comment from Zurich North America. TREAS-DO-2022-0019-0047. Regulations.gov. 2022. Available online: https://www.regulations.gov/comment/TREAS-DO-2022-0019-0047 (accessed on 15 November 2023).
70. HITRUST. Comment from HITRUST. TREAS-DO-2022-0019-0062. Regulations.gov. 2022. Available online: https://www.regulations.gov/comment/TREAS-DO-2022-0019-0062 (accessed on 15 November 2023).

71. Rasmussen, G.T. Comment from Rasmussen—Federal Cyber Insurance Feedback. TREAS-DO-2022-0019-0005 Regulations.gov. 2022. Available online: https://www.regulations.gov/comment/TREAS-DO-2022-0019-0005 (accessed on 15 November 2023).
72. Institute of International Finance. Comment from Institute of International Finance. TREAS-DO-2022-0019-0031 Regulations.gov. 2022. Available online: https://www.regulations.gov/comment/TREAS-DO-2022-0019-0031 (accessed on 15 November 2023).
73. Underwriters at Lloyd's London. Comment from Underwriters at Lloyd's, London. TREAS-DO-2022-0019-0026. Regulations.gov. 2022. Available online: https://www.regulations.gov/comment/TREAS-DO-2022-0019-0026 (accessed on 15 November 2023).
74. Aon. Comment from Aon. TREAS-DO-2022-0019-0040. Regulations.gov. 2022. Available online: https://www.regulations.gov/comment/TREAS-DO-2022-0019-0040 (accessed on 15 November 2023).
75. Cowbell. Comment from Cowbell. TREAS-DO-2022-0019-0022. Regulations.gov. 2022. Available online: https://www.regulations.gov/comment/TREAS-DO-2022-0019-0022 (accessed on 15 November 2023).
76. Converge Inc. Comment from Converge Inc. TREAS-DO-2022-0019-0006. Regulations.gov. 2022. Available online: https://www.regulations.gov/comment/TREAS-DO-2022-0019-0006 (accessed on 15 November 2023).
77. Fedtribe. Comment from Fedtribe. TREAS-DO-2022-0019-0014. Regulations.gov. 2022. Available online: https://www.regulations.gov/comment/TREAS-DO-2022-0019-0014 (accessed on 15 November 2023).
78. Centers for Better Insurance LLC. Comment from Centers for Better Insurance, LLC. TREAS-DO-2022-0019-0024. Regulations.gov. 2022. Available online: https://www.regulations.gov/comment/TREAS-DO-2022-0019-0024 (accessed on 15 November 2023).
79. Roscini, M. Cyber Operations and the jus ad bellum. In *Cyber Operations and the Use of Force in International Law*; Oxford University Press: Oxford, UK, 2014; pp. 43–116.
80. Gallagher, R.; Comment from Gallagher Re. TREAS-DO-2022-0019-0048. Regulations.gov. 2022. Available online: https://www.regulations.gov/comment/TREAS-DO-2022-0019-0048 (accessed on 15 November 2023).
81. Greenwald, J. Experts Weigh Cyber Risks, Need for Federal Backstop. Business Insurance. 2023. Available online: https://www.businessinsurance.com/article/20230302/NEWS06/912355809/Experts-eigh-cyber-risks,-need-for-federal-backstop (accessed on 15 November 2023).
82. Marius Insurance. Comment from Marius Insurance. TREAS-DO-2022-0019-0052. Regulations.gov. 2022. Available online: https://www.regulations.gov/comment/TREAS-DO-2022-0019-0052 (accessed on 15 November 2023).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

electronics

MDPI

Article

Security and Trust in the 6G Era: Risks and Mitigations

Giulio Tripi, Antonio Iacobelli, Lorenzo Rinieri and Marco Prandini *

Department of Computer Science and Engineering (DISI), Alma Mater Studiorum—Università di Bologna, 40136 Bologna, Italy; giulio.tripi@unibo.it (G.T.); antonio.iacobelli@unibo.it (A.I.); lorenzo.rinieri@unibo.it (L.R.)
* Correspondence: marco.prandini@unibo.it

Abstract: The ubiquitous diffusion of connected devices in every context of the daily life of citizens, public bodies, and companies is stimulating the creation of new applications that require very high wireless communication performances. To fulfill this need, the sixth generation of communication standards (6G) is planned to roll out by 2030. While structuring this new standard, it is crucial to take into account the security aspects given the impact of the technologies that will rely on its reliability and resiliency. In this paper, we provide an overview of the technologies that will be used in 6G to achieve the required functional goals for the development of key applications. Then, we proceed to discuss the threats and the solutions to make the communications infrastructure secure and reliable, and finally, we elaborate on the concept of how to achieve trust in this scenario.

Keywords: 6G; security; trust; wireless networks; blockchains

1. Introduction

The advent of 6G technology heralds a new era in mobile networking, promising unparalleled speed, connectivity, and potential for innovation. As the successor to 5G, 6G is not merely an incremental upgrade but a quantum leap forward, aiming to address the escalating demands of network traffic driven by the proliferation of mobile devices, Internet of Things (IoT) applications, smart city initiatives, and the burgeoning requirements of emerging technologies.

Notwithstanding the slow adoption of 5G, which has not yet fully replaced 4G [1], the first studies concerning a sixth-generation wireless network are beginning to take shape. First conceptualized as soon as 5G was standardized, the vision for 6G was born out of the recognition that existing infrastructure would soon face limitations in managing the exponential growth in data traffic and the need for ultra-low latency and broader network coverage demanded by future technologies. To satisfy these new requirements, new technologies should be developed in order to address issues spread over the full stack, ranging from the physical problems of the wireless frequencies to the adoption of highly automated management paradigms.

The potential applications of 6G are as vast as they are revolutionary, promising to reshape industries and redefine human interaction with technology. From multisensory eXtended Reality (XR) experiences to the seamless coordination of Connected Robotics and Autonomous Systems (CRAS), and from Wireless Brain–Computer Interactions (BCI) to the secure deployment of Blockchain and Distributed Ledger Technologies (DLT), there are lots of implementations that could benefit from 6G technology.

At the heart of 6G lies a trifecta of transformative technologies: Reconfigurable Intelligent Surfaces (RIS), Visible Light Communication (VLC), and the integration with Artificial Intelligence (AI). RIS offers a solution to the limitations of high-frequency transmissions, enhancing signal propagation and stability by utilizing surfaces that intelligently reflect and redirect signals. Meanwhile, VLC leverages visible light to achieve high-speed data transmission, promising seamless integration with existing infrastructure and negligible impact on human health.

The fusion of 6G with Artificial Intelligence marks a paradigm shift in network management, enabling the real-time optimization of resources, dynamic spectrum management, and robust cybersecurity measures. The role of AI extends beyond mere optimization, empowering networks to predict and adapt to traffic patterns, to seamlessly integrate application demands with network management, and to preempt cyber threats against communications.

In detail, recognizing the imperatives of security, trust, and privacy, the 6G architecture envisages a novel security paradigm, reimagining traditional approaches to safeguard against evolving threats. Divided into distinct layers—the physical layer, connection layer, and application layer—6G's security architecture adopts a multi-pronged approach to address vulnerabilities and mitigate risks.

In the physical layer, where the hardware foundation of 6G systems resides, innovative solutions such as Reconfigurable Intelligent Surfaces and Friendly Jamming offer robust defenses against eavesdropping, jamming, and spoofing attacks. These technologies not only enhance signal integrity but also render interception virtually impossible, ensuring the confidentiality of communications.

Meanwhile, at the connection layer, technologies such as Quantum Key Distribution (QKD), Network Slicing, and Intrusion Detection Systems (IDS) bolster defenses against a spectrum of threats, from Denial of Service (DoS) attacks to Man-in-the-Middle (MitM) and Replay attacks. These technologies not only fortify network integrity but also lay the groundwork for dynamic, adaptive security measures capable of thwarting emerging threats in real time.

At the top of the stack, in the application layer, Quantum Homomorphic Cryptography (QHC) and Authentication and Key Agreement (AKA) could be used to protect from threats like social engineering that take advantage of the victim's naivety.

As the 6G landscape continues to evolve, our approach to security, trust, and privacy must evolve alongside it. Concepts such as Zero Trust Architecture (ZTA), Distributed Ledger Technologies (DLT), and Trust Anchors offer a glimpse into the future of network security, promising a paradigm shift from perimeter-based defenses to a dynamic, risk-based approach that scrutinizes every interaction, device, and user within the network.

In the journey toward 6G, security is not merely a feature—it is a fundamental prerequisite for realizing the transformative potential of next-generation telecommunications. As we embark on this journey, the fusion of innovation and security will pave the way for a future where connectivity is not just fast and ubiquitous but also resilient, trustworthy, and secure.

The structure of this paper is outlined as follows: the next section gives the motivation for our work with respect for related ones, then in Section 3 we delve into the challenges, technologies, and application scenarios pertinent to 6G. Subsequently, Sections 4–6 scrutinize security and trust considerations across the physical, connection, and application layers, respectively. Lastly, in Section 7, we offer a comprehensive conclusion, summarizing the key findings and contributions of this study.

2. Motivations and Related Works

This section presents a summary of existing surveys on the subject and highlights our motivation for this work. The first contribution regards breadth: various surveys have dealt with key issues about 6G, including security, trust, and privacy. However, these surveys often focus on specific 6G areas, such as quantum security technologies [2], AI-driven security [3,4] and trusted networks [5]. These are considered to be the main 6G priorities. The second contribution instead tackles depth: other studies have conducted analyses of only one specific layer. For instance, in [6–8], the security measures concerning the physical layer are treated in detail. In [9], the potential security and privacy challenges in various 6G technologies and applications are highlighted but without focusing on the trust aspect. Finally, Nguyen et al. [10] propose a review of the security and privacy of 6G, which is divided into levels. However, the paper does not delve deeply into the trust aspects. Consequently, there is no comprehensive survey that provides a holistic view of

6G security, privacy, and trust issues in the context of the overall security architecture. Our work therefore aims to fill this gap, as we believe that treating security, privacy, and trust issues in different layers allows network and security practitioners and researchers to apply the concept of defense in depth to security design. This approach permits the design of network security in an end-to-end manner as opposed to a piecemeal approach to securing the network against attacks. Furthermore, the layered approach to security can assist in addressing conflicting objectives, such as the prioritization of security through encryption and network intelligence.

3. The Dawn of 6G

The term 6G refers to the sixth generation of mobile networking, not yet implemented, which was first discussed in [11], right before the roll-out of 5G. The idea behind 6G is the need to manage a rapid growth of network traffic due to the proliferation of mobile devices (not only mobile phones but also IoT, smart city, smart manufacturing, and smart mobility) and the request for a fast and stable connection. In addition, new technologies are rising which require a latency unreachable with today's 5G implementation [12]. In the following sections, we will describe the technologies that 6G will use to overcome the limitations of 5G, the application scenarios that require the performance enhancements that 6G will provide, and the security and privacy concerns in 6G, without forgetting an analysis on trust between the actors on the network.

3.1. Technologies

The 6G technology aims to improve wireless connectivity overcoming actual and future challenges relying on new technologies that are only theorized at the moment. The new technologies that will be used to provide the performance required by 6G are clearly described in [13]; for the sake of readability, we summarize them here:

- Reconfigurable Intelligent Surfaces (RIS): very high frequencies, beyond 100 GHz and up to a few THz [14], will be used by 6G to achieve high-speed data transmission. However, this comes with some limitations from the coverage point of view. In fact, high frequencies have a lower range and they cannot penetrate obstructions as well as lower frequencies can. Thus, to address these limitations, reflective surfaces known as RIS could be employed. These surfaces reflect signals from the sender circumventing obstacles, redirecting them to the receiver intact. As depicted in Figure 1, RIS could be Reflective, meaning that they reflect the signal in a specific direction to improve the route of the signal, or Transmissive, that is they transmit the signal by modulating and controlling it with its constituent elements. Furthermore, thanks to their versatility and scalability, they would fit seamlessly into existing networks without changing any protocols or hardware components and would cost very little for their functionality. RIS can also be useful in trying to counteract the Doppler effect [15], thus allowing a stable connection even to fast-moving devices.
- Visible Light Communication (VLC): VLC is a high-speed wireless communication technology based on the employment of visible light to transmit data. In particular, it uses LED lights to modulate visible light using a very high bandwidth with a frequency of 400–800 THz, which is much higher than the radio frequencies. VLC is a good alternative to standard wireless communication because it allows transfering large amounts of data with existing lighting systems with negligible costs and no harm to human health [16].
- Artificial Intelligence applied to 6G: The exponential development of AI in recent years will inevitably affect 6G technologies, bringing significant advantages to communication networks. In particular, Artificial Intelligence can be used to improve and automate various aspects of IT operations (AIOps-enabled) [17]. These solutions are designed to cope with the complexity and scale of modern IT environments, facilitating the management of infrastructure, applications, and services and allowing the detection of any anomalies. Furthermore, the application of Artificial Intelligence can

be employed for real-time resource management and optimization within networks (Self-Optimizing Networks). This includes enhancing traffic flow, connection speed, and forecasting traffic peaks to proactively manage them. Additionally, it will enable dynamic frequency allocation planning (Dynamic Spectrum Management) [13] and mitigate interference issues. Finally, the application of AI on a large scale will result in a notable increase in the costs associated with training and the collection of large-scale datasets. Consequently, it becomes imperative to utilize specialized AI approaches such as federated learning [18], which facilitate the coordination of the learning process across millions of distributed devices to enhance the quality of the centralized learning model (global federated learning model).

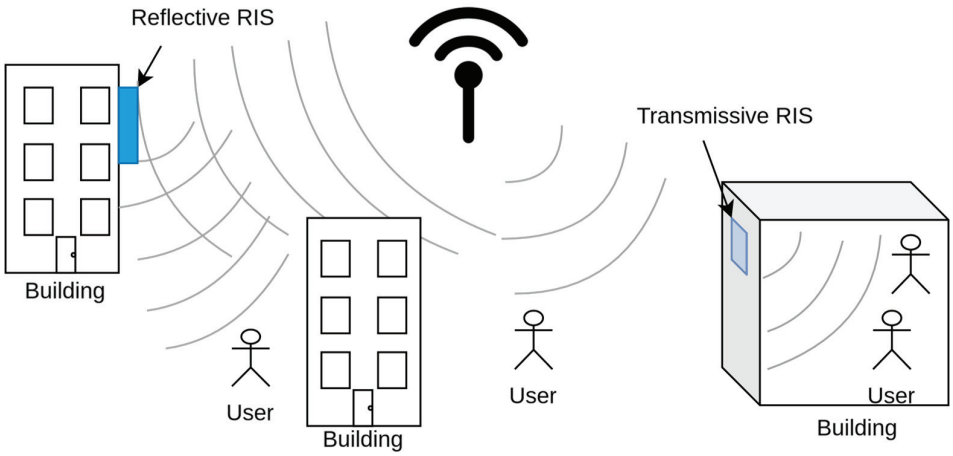


Figure 1. Examples of reflecting and transmitting Reconfigurable Intelligent Surfaces.

3.2. Application Scenarios

The 6G features enable a wide range of innovative applications, as illustrated in [19], which can be summarized as follows:

- Multisensory XR: eXtended Reality (XR) is a generic term that is used to the combination of virtual reality (VR), augmented reality (AR) and mixed reality (MR). XR is applicable to rethink how many activities of real life can happen, starting with education, healthcare, and extending to entertainment, combining the physical and digital worlds and allowing users to immerse and interact with objects in a sort of hybrid universe between virtual reality and the real world [20].
- Connected Robotics and Autonomous Systems (CRAS): One of the main objectives of 6G is certainly to enable the maximum efficiency of CRAS. The term refers to a wide range of robotic and autonomous systems connected to each other and to a common network, enabling them to share data, collaborate, and perform tasks independently. CRAS will be crucial for the operation of new services such as autonomous vehicles, remote surgery, environmental and pollution monitoring, industrial automation, and remote rescue operations in disaster-stricken locations.
- Wireless Brain–Computer Interactions (BCI): BCIs are a disruptive technology with respect to how we interact with computers and other devices. Specifically, they are based on certain devices (worn or implanted) that transmit wireless signals corresponding to the user's brain activity, thus allowing them to control local or remote devices without the need for physical contact and taking advantage of 6G connectivity. BCI was previously limited to healthcare contexts, but with 6G, its potential will increase to the point where it will be applicable in many sectors [21].
- Blockchain and Distributed Ledger Technologies (DLT): DLT is a type of decentralized technology that allows data to be recorded and shared among all participants in the network, guaranteeing authenticity and integrity without the need for a central, trusted authority. DLTs can benefit from 6G systems to provide highly scalable and secure networks that are capable of supporting the growing number of devices and

data-intensive applications. In turn, DLTs would provide 6G systems with much more security and transparency in the transfer of data and interactions between devices [22].

3.3. Security and Privacy in 6G

Security, trust, and privacy stand as cornerstones in the foundation of 6G networks. As we embark on the journey toward 6G, it is becoming increasingly evident that the existing security architecture outlined by the 3rd Generation Partnership Project (3GPP) standard [23] will require substantial enhancements to accommodate the myriad innovations poised to redefine the landscape of wireless communication. While the 3GPP standard has been instrumental in shaping the security framework of preceding generations, the advent of 6G necessitates significant changes to address the evolving threat landscape and the unique demands of futuristic applications. In response to this imperative, 6G proposes new technical solutions that build upon the foundational structure laid out by its predecessor, 5G, while incorporating significant adaptations to meet the distinctive requirements of the forthcoming systems. We will analyze those solutions by dividing them into three layers: the physical, connection, and application layers. This hierarchical division aims to provide a comprehensive and layered approach to security, effectively addressing vulnerabilities at each stratum of the network stack.

The physical layer serves as the bedrock upon which the entire network infrastructure is built, encompassing the tangible hardware components and transmission mediums that facilitate the exchange of data. Inherent vulnerabilities at this layer include physical tampering, signal interception, and electromagnetic interference. To mitigate these risks, robust encryption protocols, tamper-evident hardware, and physical security measures are employed to safeguard the integrity and confidentiality of data transmissions.

Moving up the hierarchy, the connection layer serves as the conduit through which data traverses between network nodes, encompassing protocols and procedures governing the establishment and maintenance of communication links. Vulnerabilities at this layer comprise a broad spectrum of threats, including man-in-the-middle attacks, protocol vulnerabilities, and unauthorized access to network resources. Countermeasures such as secure authentication mechanisms, intrusion detection systems, and protocol-hardening techniques are deployed to fortify the resilience of the connection layer against malicious intrusions.

At the apex of the architectural pyramid lies the application layer, the interface through which end-users interact with network services and applications. This layer represents the front-line defense against a diverse array of threats, including malware, social engineering attacks, and data breaches. To safeguard the confidentiality, integrity, and availability of user data, a multifaceted approach to security is employed, encompassing robust authentication mechanisms, data encryption, and behavioral analysis to detect anomalous activities.

In addition to delineating the vulnerabilities inherent in each layer, it is imperative to analyze the types of attacks that may exploit these vulnerabilities and the corresponding countermeasures that can be implemented to mitigate their impact. From sophisticated cyberattacks orchestrated by nation-states to opportunistic exploits perpetrated by cyber criminals, the threat landscape facing 6G networks is multifaceted and constantly evolving. By adopting a proactive stance toward security, leveraging cutting-edge technologies such as Artificial Intelligence, blockchain, and quantum cryptography, and fostering collaboration across industry stakeholders, the vision of a secure, trusted, and privacy-respecting 6G ecosystem can be realized.

3.4. Trust in 6G

Trust in telecommunication systems is a crucial requirement that is intrinsically linked to the security of the information exchanged between devices and users. In particular, trust refers to the relationship between user and machine and could be of two kinds:

- Direct: based on the past interactions between the user and the system;
- Indirect: based on third parties' opinions or recommendations.

Both types of trust play an important role in understanding the level of trustworthiness of a given system; however, it is important to consider that trust is a dynamic concept and can change over time. In 6G, due to the increase in network complexity and the exponential development of the number of interconnections, a significant improvement in trust management is required compared to today's networks, as new 6G systems will lead to the emergence of new threats and vulnerabilities that will have to be effectively countered. For this reason, trust management in 6G will be different from the traditional one, with the implementation of new models and mechanisms, analyzed in the following paragraphs, to guarantee reliability in communications.

The security, trust, and privacy vulnerabilities of each layer are analyzed below with a focus on the types of attacks that can be used and the countermeasures that can be implemented. Figure 2 shows an overview of the aforementioned elements and their division in layers, with each vulnerability linked to its countermeasure.

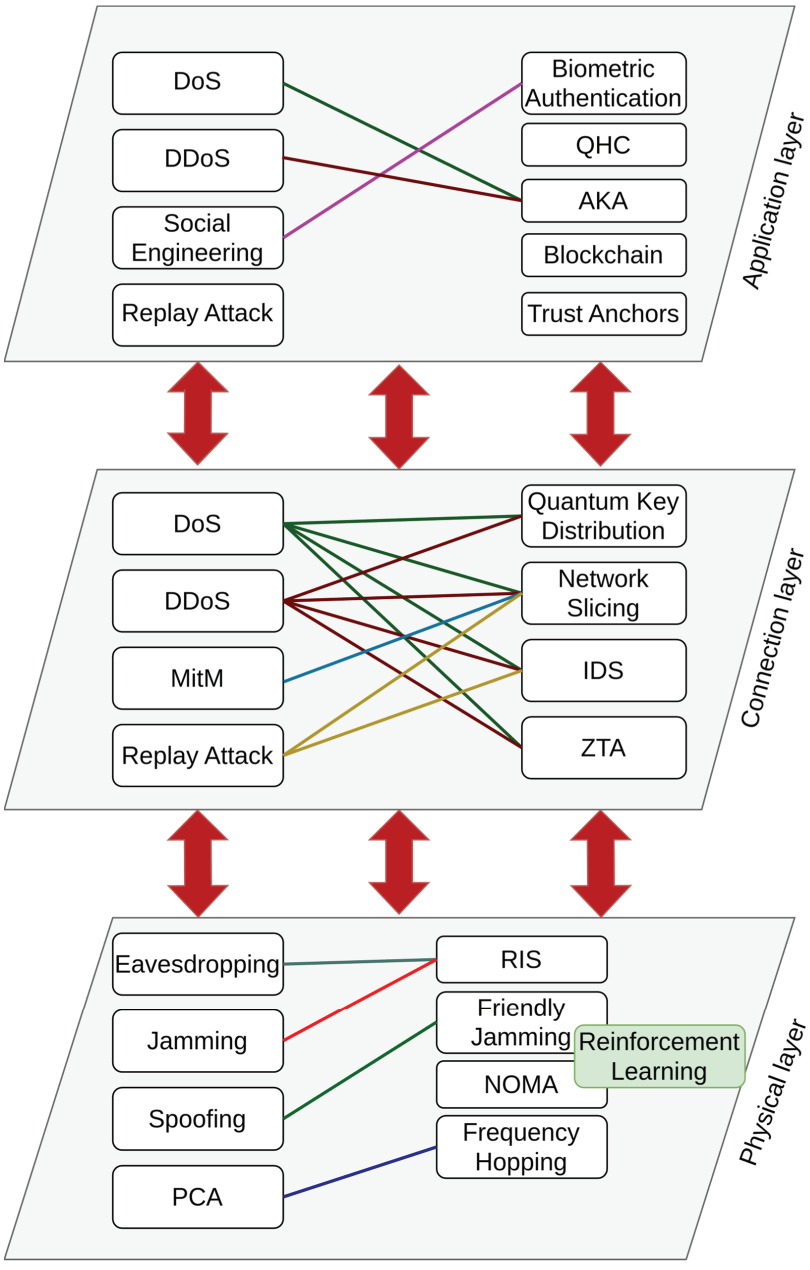


Figure 2. Overview of vulnerabilities and solutions layer by layer (link colors added for readability only—no specific meaning).

4. Physical Layer

The physical layer is the hardware foundation of 6G systems. Its critical importance is reflected in vulnerabilities that could have fatal consequences for the operation and security of the systems. In particular, the security implemented in this layer must be simple and low cost, so that it can work on devices that do not have high computing power [24]. Figure 3 shows the relations between the mitigations proposed and the security and trust threats.

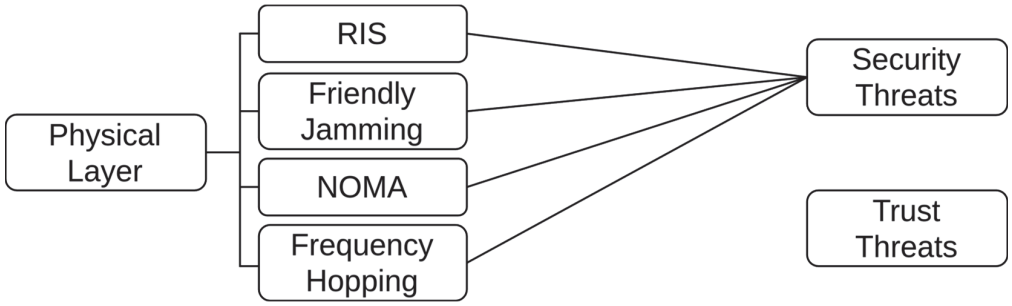


Figure 3. Diagram of the mitigations in the physical layer divided by security and trust threats.

4.1. Vulnerabilities and Threats

The most dangerous and widespread attack types in the physical layer of 6G systems will be the following:

- Eavesdropping: this is a type of attack where the attacker eavesdrops on conversations between two parties without authorization and intercepts sensitive information [25].
- Jamming: this attack generates interference signals or noise in the communication channel, making the transmission of information unreliable and disrupting communication [26].
- Pilot Contamination Attack (PCA): this attack "contaminates" the pilot signal from a base station, compromising communications and network performance [27].
- Spoofing: this attack involves spoofing the identity of a person or device to deceive other users or systems and gain unauthorized access [26].

4.2. Mitigations

4.2.1. Reconfigurable Intelligent Surfaces

Reconfigurable Intelligent Surfaces are crucial in avoiding eavesdropping attacks due to their ability to reflect the signal without degrading it. In particular, they can be configured to reflect the signal itself in unpredictable ways, making it almost impossible for the attacker to identify the direction of the signal and intercept it. The effectiveness of this countermeasure increases proportionally to the scale of RIS deployment: the more possible paths for the signal being sent, the more difficult it is for the attacker to identify and intercept the communication. Figure 4 illustrates an example of how RIS, by reflecting the signal in an indeterminate manner, is capable of evading eavesdropping attacks. This is achieved by allowing the communication to traverse two distinct paths, thereby preventing the attacker from identifying the correct one. In addition, active RIS could be used to amplify only the signal directed to the legitimate user [28].

4.2.2. Friendly Jamming

Friendly Jamming, also known as artificial noise, consists of the insertion of jamming signals into the communications channel from sources that are considered "friendly", such as other base stations in the network, and which do not compromise the integrity of the communications. In fact, these jamming signals are designed to disrupt a potential attacker, making it difficult for them to decipher the message or attempt to interfere. Despite the jamming, the receiver manages to receive the signal undisturbed. This is achieved by modulation techniques as described in [29]. In this article, Friendly Jamming

is applied to MIMO-OFDM (Multiple-Input, Multiple-Output Orthogonal Frequency-Division Multiplexing) systems or by using specific frequencies to filter out artificial noise. Friendly Jamming has some limitations: first, it could be subject to spoofing attacks, i.e., an attacker could spoof a friendly signal to make it indistinguishable from the real one for the detection system and thus gain access to the communication channel; a second weakness is the lack of sufficient resources to ensure reliable Friendly Jamming: this is more common in systems with limited processing capacity, such as small devices.

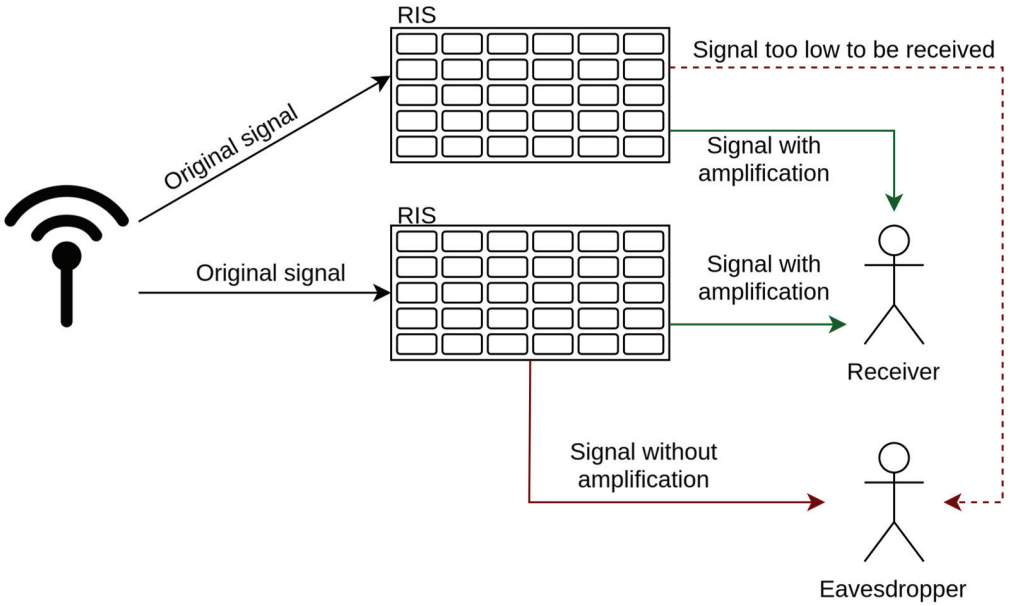


Figure 4. Example of an RIS-based countermeasure to the eavesdropping attack.

4.2.3. NOMA

NOMA (Non-Orthogonal Multiple Access) is an access technique for multiple users to communicate over the same data spectrum [30]. It achieves non-orthogonal access by spatially overlapping the signals transmitted, allowing multiple users to share time and frequency resources. This gives higher power to data with higher priority and removes it from data with lower priority, making the former more difficult to intercept and increasing their security. To receive the distinct signal, users must decode and demodulate the overlapping signals, exploiting differences in signal power levels and using advanced modulation techniques. Due to its flexibility and optimization, NOMA is particularly suitable for systems such as 6G where a large number of devices are connected and there is a need to ensure numerous simultaneous communications. Despite the high security of NOMA, some vulnerabilities could jeopardize the integrity and security of communication. One of these is the wiretapping of transmission devices through which attackers passively intercept communications. Attackers who gain control of multiple transmission devices can spy on communications passively and try to discover the modulated message by simulating the modulation technique of the intercepted channel. A further vulnerability could arise if attackers send numerous messages with high priority to allocate enough power to make it impossible for other communications to have protection. To ensure greater security and reduce the number of vulnerabilities, the literature [30] proposes several solutions. These include the implementation of Friendly Jamming and the use of RIS in the NOMA itself, both of which are particularly optimized for 6G systems.

4.2.4. Frequency Hopping

Frequency Hopping [31] is a technique used against PCA attacks. It is based on changing the frequency of the communication channel in an uncoordinated way. This makes it difficult for an attacker attempting to contaminate the pilot signal of the communication to identify the path of the communication and intercept it. Although an attacker may be able

to detect the signal on a specific frequency, it would be impossible to completely reconstruct it due to the unknown jumps. Additionally, Zhang et al. [31] have developed an algorithm to detect a PCA, while an attacker attempts to emulate communication frequency hopping.

4.2.5. Reinforcement Learning

Reinforcement Learning (RL) is a machine learning technique where an agent interacts with the environment, learning an optimal policy, by trial and error [32]. Reinforcement Learning can be a cornerstone for securing the 6G systems' physical layer. It enables the automatic optimization of security criteria by analyzing the environment through an algorithm and adjusting individual parameters. This eliminates the need to rely solely on pre-established attack patterns. Furthermore, it could be utilized to optimize previously discussed physical layer techniques, such as NOMA and Friendly Jamming. The first could be optimized through the use of RL, with an algorithm capable of selecting the appropriate transmission power for each signal, setting the optimal ratio between secrecy rate and interception probability, as demonstrated in [33]. Friendly Jamming can be optimized through beamforming, which is a technique that concentrates the signal beam in a specific direction instead of spreading it in all directions, based on RL. Through Reinforcement Learning, beamforming parameters such as orientation and intensity can be dynamically modified to adapt to the attacker's tactics and behaviors [34].

5. Connection Layer

The connection layer is the one responsible for managing and establishing connections between devices. It is also responsible for authenticating and discovering devices on the network and managing traffic to avoid congestion and improve QoS. Figure 5 shows the relations between the mitigations proposed and the security and trust threats.

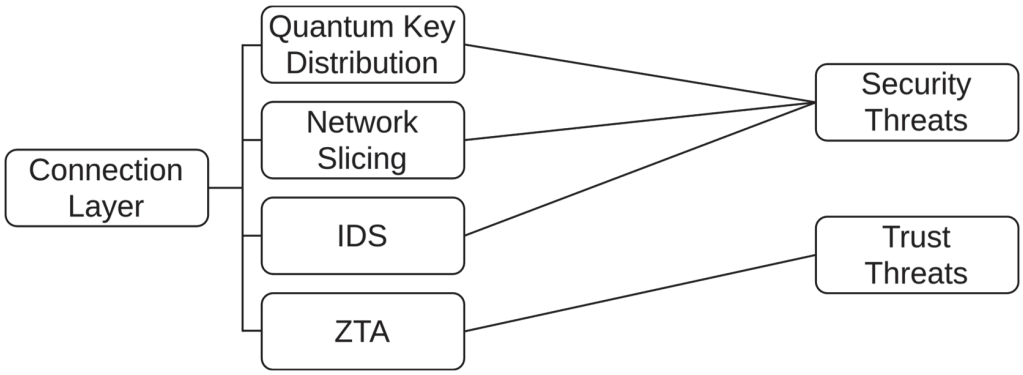


Figure 5. Diagram of the mitigations in the connection layer divided by security and trust threats.

5.1. Vulnerabilities and Threats

The connection layer can fall victim to other types of attacks than those of the physical layer:

- DoS (Denial of Service): this attack aims to make a service, resource, or network inaccessible to legitimate users by creating an excessive volume of requests, saturating the available traffic [35].
- Distributed Denial of Service (DDoS): this is a similar type of attack to DoS but uses multiple devices to achieve higher impact by increasing the number of requests and hampering defenses based on isolating the attacker's address [36].
- MitM (Man-in-the-Middle): this is an active attack in which attackers are able to position themselves on the path between legitimate endpoints to intercept and alter communication without their knowledge [37].
- Replay Attack: in this attack, the attacker intercepts and stores communication data and then later reuses it to gain unauthorized access or to perform malicious operations [38].

5.2. Mitigations

5.2.1. Quantum Key Distribution

Quantum Key Distribution (QKD) uses the fundamental principles of quantum physics to generate keys that are impossible to intercept. In a QKD system, there are two key figures, the transmitter and the receiver, connected via a dual link, one classical and one quantum. In the classic channel, information is transmitted with traditional IT techniques, while in the quantum channel, data are transmitted through quantum states, e.g., the instantaneous microscopic state of individual photons that can be used to transmit information [39]. If an attacker attempts to intercept a communication on the quantum channel, the quantum states of the photons will be altered by the laws of quantum physics, consequently altering the communication. Therefore, QKD allows a more secure distribution of symmetric keys between the transmitter and receiver compared to 5G's symmetric encryption mechanism. However, this technology still has evident limitations due to the necessity of utilizing specialized equipment. Quantum Key Distribution (QKD) is based on physical principles for its security, which is derived from the distinctive characteristics of the physical layer in communications. This necessitates that users either lease dedicated fiber connections or manually manage transmitters in open space. Its implementation is not feasible via software or as a network service, and it presents challenges for seamless integration into the current network infrastructure.

5.2.2. Network Slicing

Network Slicing is a technology that allows for dividing the network into multiple virtual networks each with its own resources and characteristics. The division is based on the functionality and performance of each service so that each service has specific requirements for the particular use case to which it refers. From a security point of view, Network Slicing can be used to ensure the confidentiality and control of resource consumption in a 6G system, as cross-slice communication is not allowed. Although this may seem like a limitation, it guarantees a greater level of security because any attack will be limited to a single slice and will not spread throughout the entire system. A further security mechanism consists of authenticating the managers of two network slices belonging to two different endpoints before allowing the single slice to process data, as described in [40]. Figure 6 shows a representation of how network slicing separates different kinds of traffic, adapting the transmission based on their different needs.

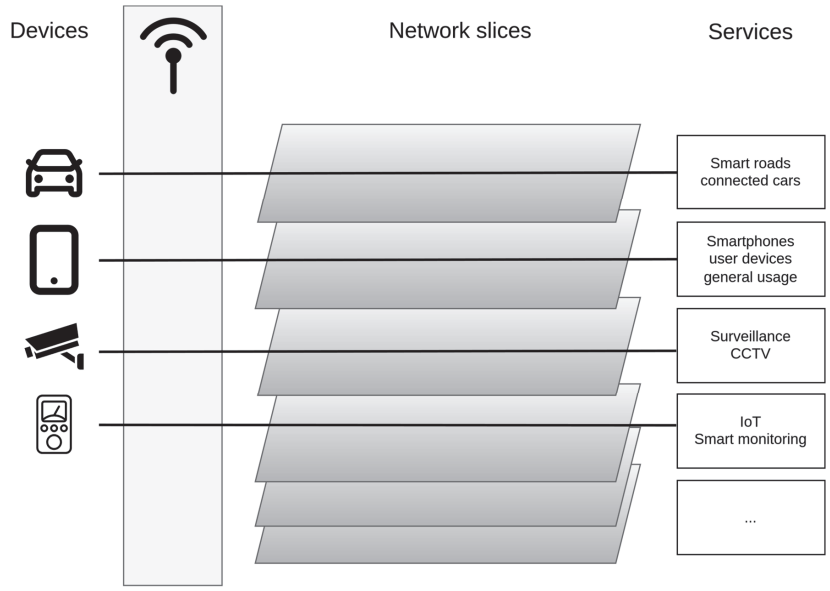


Figure 6. Example of how Network Slicing works.

5.2.3. IDS

The IDS is a system that automates the event monitoring process, analyzing and researching security threats [41]. An IDS observes event data and detects potentially dangerous activity, and it can also prevent activities (IDPS). In 6G systems, IDS will have a fundamental role. In particular, it will focus on detection methods based on Deep Learning (DL), which is a type of machine learning that uses learning algorithms to represent processed data as a hierarchy of nested concepts [42]. Deep Learning can be applied to the IDS of 6G systems to better handle the increasing complexity and size of transmitted data. Its ability to automatically learn from complex data makes it a valuable tool. DL can be used for dimensionality reduction, which reduces the size of data while preserving its meaningful properties, and for data classification, which assigns labels to new data based on patterns learned from training data. The complexity of an Intrusion Detection System (IDS) based on Deep Learning is high due to the numerous mathematical operations involved, which require significant computing power. This can be challenging to process even with AI accelerators. Recent studies [43–45] have shown promising results, and their potential application on a large scale is becoming increasingly feasible.

5.3. Zero Trust Architecture (ZTA)

Another approach to the concept of trust in 6G is that of Zero Trust Architecture. Zero Trust Architecture is a strategy based on the idea that any user, device, or system within a network can be compromised at any time and is not to be trusted. Thus placing itself on the defensive, the ZTA views the network as a continuously hostile environment and its geographic location as not relevant enough to deem it trustworthy and credible. From this assumption, a series of techniques are implemented based on tight control of network access with an analysis of the risk associated with each access request; an authentication system for each user based on location, behavior, and device to verify his or her identity with the utmost precision; and an algorithm based on machine learning that analyzes all the information transmitted in the network and modifies parameters and criteria to regulate access and authentication [46]. From the perspective of logical architecture, a ZTA consists of three basic components [47]:

- Policy Engine: evaluates the validity of a user's access request using a Trust Algorithm, which is a machine learning algorithm that greatly optimizes this process;
- Policy Administrator: cooperating with the Policy Engine, decides whether to allow or deny the access using a Trust Algorithm, automating access control with a minimal error rate;
- Policy Enforcement Point: monitors and manages links between users and resources.

As of today, the foundations of the Zero Trust Architecture have been laid, but the technologies to be implemented to achieve the requirements in a system such as 6G are still far from being fully realized. The biggest problem concerns the large-scale deployment of this architecture: it requires considerable computing power that cannot be guaranteed in every device in a heterogeneous network such as 6G. In addition, access control cannot be uniform for each network; therefore, it is necessary to implement a dynamic one: the literature proposes solutions [48] but they are still immature and not fully applicable.

6. Application Layer

The application or service layer acts as a middleware platform that manages the system's interactions with third-party services, such as edge computing, to ensure efficiency and responsiveness. Figure 7 shows the relations between the mitigations proposed and the security and trust threats.

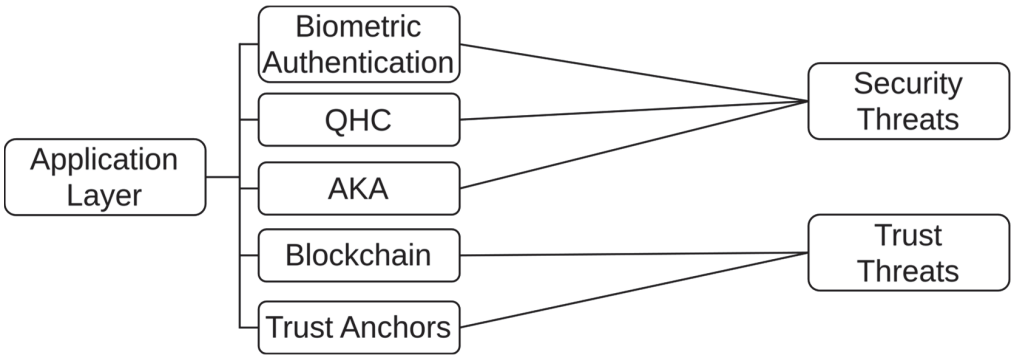


Figure 7. Diagram of the mitigations in the application layer divided by security and trust threats.

6.1. Vulnerabilities and Threats

In addition to the sheer volume of data processed, which makes the application layer sensitive to the aforementioned DoS and DDoS attacks, the wide variety of interaction models exposes this layer to some threats that are even more difficult to counter. A peculiar threat targeting this layer, in fact, is likely to be not a purely technical one, but social engineering [49]. It consists of a series of methods used by the attacker to obtain the user's personal information through deception. There is no real method to counter social engineering, as the attacker often takes advantage of the victim's ignorance and naivety, but there are countermeasures that can reduce the ease of access to sensitive data, even in the new 6G systems, where the various connected devices will allow access to information and devices that are also particularly critical to human health.

6.2. Mitigations

6.2.1. Biometric Authentication

Biometric authentication [50] uses physical and behavioral properties to verify an individual's identity. There are three categories:

- Biologic, based on genetic factors of the individual;
- Morphological, based on physical traits of the individual;
- Behavioral, based on the individual's behavioral patterns.

Although biometrics have been a widely used authentication system for years, the exponential growth in computing power and advances in AI are making biometric spoofing increasingly sophisticated. As a result, to prevent potential spoofing, biometric authentication needs to ensure a higher level of authenticity. On the other hand, 6G systems are expected to significantly improve this technology, which is particularly due to the power of THz waves. These waves have high transmission capacity, enabling an unprecedented level of detail and accuracy in biometric authentication. They can facilitate the identification of individuals based on unique characteristics such as heartbeat [51] or brain waves [52].

6.2.2. Quantum Homomorphic Cryptography (Qhc)

Homomorphic cryptography is an advanced cryptographic technique that allows operations to be performed on encrypted data without the need for decryption, providing extremely high security in centralized edge computing nodes. Figure 8 illustrates the phases of homomorphic cryptography in a simplified manner. A recent paper by Jonas Zeuner et al. [53] discusses a specific variant of homomorphic encryption based on the principles of quantum physics. The paper describes the use of the quantum properties of particles, such as photons, to perform operations on encrypted quantum data. Instead of conventional encryption techniques, the direction of polarization of certain photons is randomly changed to encrypt the data. The encrypted data are then transmitted to the desired service, which performs operations on it without decrypting it. Finally, the processed data are returned to the user by performing the inverse transformation. The quantum variant involves operations on particles and interpreting data through quantum

states. There is little doubt about the security of QHC, and the direction for application layer cryptography is certainly geared toward it. However, there are still numerous problems to be addressed, which are particularly due to the practical implementation of quantum operations, similar to what we already noted in Section 5.2.1. This technology is still far from being fully realized.

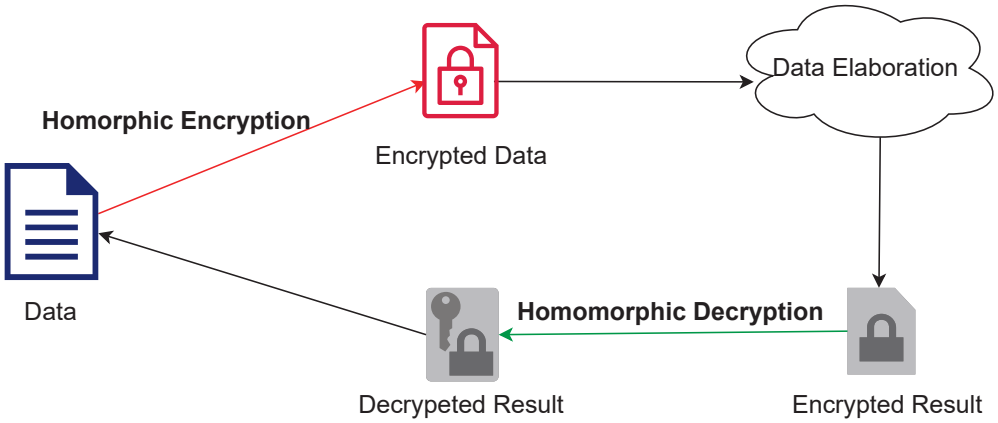


Figure 8. Example of homomorphic cryptography.

6.2.3. Authentication and Key Agreement (AKA)

Authentication and Key Agreement (AKA) is a security protocol used for mutual authentication between a device and a network provider. AKA negotiates session keys to encrypt communications between the two parties. The protocol was first used in 3G networks and has since evolved to subsequent ones, remaining the standard for mutual authentication. In 6G, the Authentication and Key Agreement protocol will need to be faster, more reliable, and provide more secure authentication. This is particularly important given the vulnerabilities that have been exposed in 5G, such as Distributed Denial of Service (DDoS) attacks [54]. Furthermore, the AKA protocol will face new challenges compared to previous network systems due to the large number of connected devices and their heterogeneous security systems. Ensuring a consistent authentication model across systems and devices with varying security capabilities presents a significant challenge. An attacker may deceive the network provider into believing that they possess a lower security capability, thereby obtaining a weaker authentication level and gaining access to sensitive data. To ensure a unique and secure authentication model, quantum algorithms will be implemented. Section 5.2.1 analyzes the quantum key distribution technique.

6.3. Blockchain

A blockchain is a particular type of DLT structured as a chain of blocks containing data. The concatenation of the blocks is guaranteed by cryptography. Blockchain technology is characterized by two key principles: the Consensus Algorithm and the Smart Contract. The Consensus Algorithm is a protocol used by blockchains to establish which transactions (or blocks) to add or keep in the Distributed Ledger. Every block or transaction added to the chain must be validated by all participants (also called miners), guaranteeing integrity and creating an unbreakable bond with the other blocks or transactions in the blockchain. Thanks to this process, in the case of a large-scale blockchain, it would be almost impossible to change one part of the chain retroactively, since it would be necessary to change all the blocks subsequent to the one desired. In addition, each block keeps track of the party, making an action on the blockchain and where it came from, thus ensuring traceability on each individual transaction. There is no single Consensus Algorithm, but there are different types, each with its own advantages and disadvantages. The main ones are Proof Of Work, in which miners compete by solving cryptographic puzzles and the first to reach the answer has the right to create the new block, and Proof Of Stake, in which the one who has the most resources has the right of precedence in the creation of a new block [55].

The Smart Contract, on the other hand, is a computer program that outlines the conditions of a "contract" and in turn manages its implementation when those conditions are met. It is a useful tool because it allows specific transactions to be handled automatically when certain conditions are met and is particularly effective in areas such as programmatic banking functions and decentralized markets [56]. Blockchain technology, thanks to its many functionalities, is a crucial tool for ensuring security and reliability in modern network systems, and its application in 6G networks will be the backbone of trust in any domain: in fact, the purpose of blockchain is for every user to trust the blockchain itself so that they do not need to trust each other. As far as communications are concerned, this technology will have the dual task of ensuring the credibility of each entity on the network, assigning a trust value to each entity based on certain indicators, possibly detecting and expelling malicious ones from the network [57]; and improving the authenticity of transmitted data, using its own technologies to augment data encryption in edge computing [58].

6.4. Trust Anchors

Trust Anchors are entities whose trust is predetermined and not derived from other sources. In an environment like 6G, it is crucial to have a universally recognized trusted entity among all involved parties. The architecture of a Trust Anchor can vary from a centralized to a distributed model, adapting to the needs of the network in which it is located. In order to be valid, Trust Anchors must meet certain requirements: first, they must guarantee a non-repudiation of actions within the system. That is, once a user has provided information to the Trust Anchor, he cannot go back and deny the action. Another key aspect is the ability to verify that the information saved was not compromised. Ultimately, the mechanism to verify the integrity of the information must be transparent and known to all the external users. In addition, Trust Anchors are often aided by tools called Trust and Traceability Functions (TTFs) [59], which is a particular blockchain-based service that helps establish a connection between the core of a network and its peers by storing via a distributed ledger similar to that of DLTs all communication activity in a secure manner. In addition, this service also provides a trusted logging facility that securely records the path of data exchanged during communication. However, TTFs are independent of Trust Anchors: in fact, the latter does not store user data, which is instead a task entrusted to TTFs, but focuses on verifying consensus among TTFs in distributed registries, acting as guarantors of the trustworthiness of the registries and ensuring that the information stored is authentic and verified. This division of labor provides an additional level of security and reliability in the area of data management and verification within the system. The Trust Anchors model has excellent potential for application, but the debate still remains about how to assign the label of "trustworthy" to a given Trust Anchor.

7. Conclusions

Every new generation of telecommunications technology is expected to reach ever-increasing speed and capacity. The 6G technology makes no exception, but the raw performance is only the tip of the iceberg. The real game changer will be a deep integration between the network and the applications that need it. While 5G started considering the services as drivers for network configuration but still placed the burden of configuration on the network operators on request of the service providers, 6G aspires to achieve a seamless interplay between users' needs and network reaction to satisfy them. The impact of malicious behavior in such an integrated environment could be hard to contain in terms of spread and gravity. Consequently, security issues shall be taken into account at every phase of development, starting from the fundamental concepts, through the design of standards, and clearly ending in implementation. In light of this, it is of the utmost importance to conduct in-depth research into the countermeasures to be implemented in order to ensure that this new standard is as secure as possible. In our work, we undertook an analysis of the security, privacy, and trust aspects in different layers, namely the physical layer, connection layer, and application layer. This was completed in order to

gain a comprehensive understanding of the security issues present in each layer and to facilitate the design of an end-to-end network security solution. Specifically, for each layer, we identified vulnerabilities, such as DoS, jamming, and pilot contamination attacks, and proposed potential mitigations, such as Friendly Jamming, Network Slicing, and QKC. It is evident that meeting real-time protection requirements and energy efficiency remain significant challenges for these technologies. Without these features, it is unlikely that many 6G security services will be able to achieve their intended objectives. In future works, it would be beneficial to examine each level of the OSI model in order to gain a more comprehensive understanding of potential security issues.

Author Contributions: Conceptualization, G.T., A.I. and L.R.; Methodology, G.T., A.I. and L.R.; Validation G.T., A.I. and L.R.; Writing—original draft, G.T. and A.I.; Writing—review and editing, L.R. and M.P.; Supervision, M.P. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partially supported by project SERICS (PE00000014) under the MUR National Recovery and Resilience Plan funded by the European Union—NextGenerationEU and by the IGNITE5.0 project funded by the Emilia-Romagna region in the framework of the POR-FESR 2021-2027 program (the EU program for regional development).

Data Availability Statement: Data is contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

3GPP	3rd Generation Partnership Project
AI	Artificial Intelligence
AKA	Authentication Key Agreement
BCI	Wireless Brain–Computer Interactions
CRAS	Connected Robotics and Autonomous Systems
DDoS	Distributed Denial of Service
DL	Deep Learning
DLT	Distributed Ledger Technologies
DoS	Denial of Service
IDPS	Intrusion Detection and Prevention System
IDS	Intrusion Detection System
IoT	Internet of Things
MIMO-OFDM	Multiple-Input, Multiple-Output Orthogonal Frequency-Division Multiplexing
MitM	Man-in-the-Middle
NOMA	Non-Orthogonal Multiple Access
PCA	Pilot Contamination Attacks
QHC	Quantum Homomorphic Cryptography
QKD	Quantum Key Distribution
QoS	Quality of Service
RIS	Reconfigurable Intelligent Surfaces
RL	Reinforcement Learning
TTFs	Trust and Traceability Functions
VLC	Visible Light Communication
XR	eXtended Reality
ZTA	Zero Trust Architecture

References

1. GSMA. The Mobile Economy 2024. 2024. Available online: https://www.gsma.com/solutions-and-impact/connectivity-for-good/mobile-economy/wp-content/uploads/2024/02/260224-The-Mobile-Economy-2024.pdf (accessed on 17 April 2024).
2. Pirandola, S.; Andersen, U.L.; Banchi, L.; Berta, M.; Bunandar, D.; Colbeck, R.; Englund, D.; Gehring, T.; Lupo, C.; Ottaviani, C.; et al. Advances in quantum cryptography. *Adv. Opt. Photonics* **2020**, *12*, 1012–1236. [CrossRef]

3. Sun, Y.; Liu, J.; Wang, J.; Cao, Y.; Kato, N. When machine learning meets privacy in 6G: A survey. *IEEE Commun. Surv. Tutorials* **2020**, *22*, 2694–2724. [CrossRef]
4. Sheth, K.; Patel, K.; Shah, H.; Tanwar, S.; Gupta, R.; Kumar, N. A taxonomy of AI techniques for 6G communication networks. *Comput. Commun.* **2020**, *161*, 279–303. [CrossRef]
5. Ylianttila, M.; Kantola, R.; Gurtov, A.; Mucchi, L.; Oppermann, I.; Yan, Z.; Nguyen, T.H.; Liu, F.; Hewa, T.; Liyanage, M.; et al. 6G white paper: Research challenges for trust, security and privacy. *arXiv* **2020**, arXiv:2004.11665.
6. Mitev, M.; Chorti, A.; Poor, H.V.; Fettweis, G.P. What Physical Layer Security Can Do for 6G Security. *IEEE Open J. Veh. Technol.* **2023**, *4*, 375–388. [CrossRef]
7. Mucchi, L.; Jayousi, S.; Caputo, S.; Panayirci, E.; Shahabuddin, S.; Bechtold, J.; Morales, I.; Stoica, R.A.; Abreu, G.; Haas, H. Physical-Layer Security in 6G Networks. *IEEE Open J. Commun. Soc.* **2021**, *2*, 1901–1914. [CrossRef]
8. Xie, N.; Chen, J.; Huang, L. Physical-layer authentication using multiple channel-based features. *IEEE Trans. Inf. Forensics Secur.* **2021**, *16*, 2356–2366. [CrossRef]
9. Porambage, P.; Gür, G.; Osorio, D.P.M.; Liyanage, M.; Gurtov, A.; Ylianttila, M. The roadmap to 6G security and privacy. *IEEE Open J. Commun. Soc.* **2021**, *2*, 1094–1122. [CrossRef]
10. Nguyen, V.L.; Lin, P.C.; Cheng, B.C.; Hwang, R.H.; Lin, Y.D. Security and Privacy for 6G: A Survey on Prospective Technologies and Challenges. *IEEE Commun. Surv. Tutor.* **2021**, *23*, 2384–2428. [CrossRef]
11. David, K.; Berndt, H. 6G Vision and Requirements: Is There Any Need for Beyond 5G? *IEEE Veh. Technol. Mag.* **2018**, *13*, 72–80. [CrossRef]
12. Nguyen, D.C.; Ding, M.; Pathirana, P.N.; Seneviratne, A.; Li, J.; Niyato, D.; Dobre, O.; Poor, H.V. 6G Internet of Things: A Comprehensive Survey. *IEEE Internet Things J.* **2022**, *9*, 359–383. [CrossRef]
13. Jiang, W.; Han, B.; Habibi, M.A.; Schotten, H.D. The Road Towards 6G: A Comprehensive Survey. *IEEE Open J. Commun. Soc.* **2021**, *2*, 334–366. [CrossRef]
14. Rappaport, T.S.; Xing, Y.; Kanhere, O.; Ju, S.; Madanayake, A.; Mandal, S.; Alkhateeb, A.; Trichopoulos, G.C. Wireless Communications and Applications Above 100 GHz: Opportunities and Challenges for 6G and Beyond. *IEEE Access* **2019**, *7*, 78729–78757. [CrossRef]
15. Bilotti, F.; Barbuto, M.; Hamzavi-Zarghani, Z.; Karamirad, M.; Longhi, M.; Monti, A.; Ramaccia, D.; Stefanini, L.; Toscano, A.; Vellucci, S. Reconfigurable intelligent surfaces as the key-enabling technology for smart electromagnetic environments. *Adv. Phys. X* **2024**, *9*, 2299543. [CrossRef]
16. Chi, N.; Zhou, Y.; Wei, Y.; Hu, F. Visible Light Communication in 6G: Advances, Challenges, and Prospects. *IEEE Veh. Technol. Mag.* **2020**, *15*, 93–102. [CrossRef]
17. Ziegler, V.; Schneider, P.; Viswanathan, H.; Montag, M.; Kanugovi, S.; Rezaki, A. Security and Trust in the 6G Era. *IEEE Access* **2021**, *9*, 142314–142327. [CrossRef]
18. Liu, Y.; Yuan, X.; Xiong, Z.; Kang, J.; Wang, X.; Niyato, D. Federated learning for 6G communications: Challenges, methods, and future directions. *China Commun.* **2020**, *17*, 105–118. [CrossRef]
19. Saad, W.; Bennis, M.; Chen, M. A vision of 6G wireless systems: Applications, trends, technologies, and open research problems. *IEEE Net.* **2019**, *34*, 134–142. [CrossRef]
20. Minopoulos, G.; Psannis, K.E. Opportunities and Challenges of Tangible XR Applications for 5G Networks and Beyond. *IEEE Consum. Electron. Mag.* **2023**, *12*, 9–19. [CrossRef]
21. Hu, H.; Chen, X.; Jiang, T. Guest editorial: Brain-computer-interface inspired communications. *China Commun.* **2022**, *19*, iii–v. [CrossRef]
22. Baskaran, S.B.M.; Faisal, T.; Wang, C.; Lopez, D.R.; Ordonez-Lucena, J.; Arribas, I. The Role of DLT for Beyond 5G Systems and Services: A Vision. *IEEE Commun. Stand. Mag.* **2023**, *7*, 32–38. [CrossRef]
23. Cao, J.; Ma, M.; Li, H.; Ma, R.; Sun, Y.; Yu, P.; Xiong, L. A Survey on Security Aspects for 3GPP 5G Networks. *IEEE Commun. Surv. Tutor.* **2020**, *22*, 170–195. [CrossRef]
24. López, O.L.A.; Alves, H.; Souza, R.D.; Montejo-Sánchez, S.; Fernández, E.M.G.; Latva-Aho, M. Massive Wireless Energy Transfer: Enabling Sustainable IoT Toward 6G Era. *IEEE Internet Things J.* **2021**, *8*, 8816–8835. [CrossRef]
25. Wei, Z.; Liu, F.; Masouros, C.; Su, N.; Petropulu, A.P. Toward Multi-Functional 6G Wireless Networks: Integrating Sensing, Communication, and Security. *IEEE Commun. Mag.* **2022**, *60*, 65–71. [CrossRef]
26. Vaishnavi, K.N.; Khorvi, S.D.; Kishore, R.; Gurugopinath, S. A Survey on Jamming Techniques in Physical Layer Security and Anti-Jamming Strategies for 6G. In Proceedings of the 2021 28th International Conference on Telecommunications (ICT), London, UK, 1–3 June 2021; pp. 174–179. [CrossRef]
27. Akgun, B.; Krunz, M.; Koyluoglu, O.O. Pilot contamination attacks in massive MIMO systems. In Proceedings of the 2017 IEEE Conference on Communications and Network Security (CNS), Las Vegas, NV, USA, 9–11 October 2017. [CrossRef]
28. Zhang, S.; Gao, H.; Su, Y.; Cheng, J.; Jo, M. Intelligent Mixed Reflecting/Relaying Surface-Aided Secure Wireless Communications. *IEEE Trans. Veh. Technol.* **2024**, *73*, 532–543. [CrossRef]
29. Cepheli, Ö.; Kurt, G.K. Efficient PHY layer security in MIMO-OFDM: Spatiotemporal selective artificial noise. In Proceedings of the 2013 IEEE 14th International Symposium on "A World of Wireless, Mobile and Multimedia Networks" (WoWMoM), Madrid, Spain, 4–7 June 2013; pp. 1–6. [CrossRef]

30. Furqan, H.M.; Hamamreh, J.; Arslan, H. Physical layer security for NOMA: Requirements, merits, challenges, and recommendations. *arXiv* **2019**, arXiv:1905.05064.
31. Zhang, W.; Lin, H.; Zhang, R. Detection of pilot contamination attack based on uncoordinated frequency shifts. *IEEE Trans. Commun.* **2018**, *66*, 2658–2670. [CrossRef]
32. Kaelbling, L.P.; Littman, M.L.; Moore, A.W. Reinforcement learning: A survey. *J. Artif. Intell. Res.* **1996**, *4*, 237–285. [CrossRef]
33. Xiao, L.; Li, Y.; Dai, C.; Dai, H.; Poor, H.V. Reinforcement learning-based NOMA power allocation in the presence of smart jamming. *IEEE Trans. Veh. Technol.* **2017**, *67*, 3377–3389. [CrossRef]
34. Xiao, L.; Sheng, G.; Liu, S.; Dai, H.; Peng, M.; Song, J. Deep reinforcement learning-enabled secure visible light communication against eavesdropping. *IEEE Trans. Commun.* **2019**, *67*, 6994–7005. [CrossRef]
35. Pelechrinis, K.; Iliofotou, M.; Krishnamurthy, S.V. Denial of service attacks in wireless networks: The case of jammers. *IEEE Commun. Surv. Tutor.* **2010**, *13*, 245–257. [CrossRef]
36. Mirkovic, J.; Reiher, P. A taxonomy of DDoS attack and DDoS defense mechanisms. *ACM Sigcomm Comput. Commun. Rev.* **2004**, *34*, 39–53. [CrossRef]
37. Conti, M.; Dragoni, N.; Lesyk, V. A survey of man in the middle attacks. *IEEE Commun. Surv. Tutor.* **2016**, *18*, 2027–2051. [CrossRef]
38. Syverson, P. A taxonomy of replay attacks [cryptographic protocols]. In Proceedings of the Computer Security Foundations Workshop VII, Franconia, NH, USA, 14–16 June 1994; pp. 187–191.
39. Bae, J.; Kwek, L.C. Quantum state discrimination and its applications. *J. Phys. Math. Theor.* **2015**, *48*, 083001. [CrossRef]
40. 5G Americas. *The Evolution of Security in 5G, a Slice of Mobile Threats*; 5G Americas: Bellevue, MA, USA, 2019.
41. Liao, H.J.; Lin, C.H.R.; Lin, Y.C.; Tung, K.Y. Intrusion detection system: A comprehensive review. *J. Netw. Comput. Appl.* **2013**, *36*, 16–24. [CrossRef]
42. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef]
43. Al-Garadi, M.A.; Mohamed, A.; Al-Ali, A.K.; Du, X.; Ali, I.; Guizani, M. A survey of machine and deep learning methods for internet of things (IoT) security. *IEEE Commun. Surv. Tutor.* **2020**, *22*, 1646–1685. [CrossRef]
44. Xin, Y.; Kong, L.; Liu, Z.; Chen, Y.; Li, Y.; Zhu, H.; Gao, M.; Hou, H.; Wang, C. Machine learning and deep learning methods for cybersecurity. *IEEE Access* **2018**, *6*, 35365–35381. [CrossRef]
45. Aldweesh, A.; Derhab, A.; Emam, A.Z. Deep learning approaches for anomaly-based intrusion detection systems: A survey, taxonomy, and open issues. *Knowl. Based Syst.* **2020**, *189*, 105124. [CrossRef]
46. He, Y.; Huang, D.; Chen, L.; Ni, Y.; Ma, X. A survey on zero trust architecture: Challenges and future trends. *Wirel. Commun. Mob. Comput.* **2022**, *2022*, 6476274. [CrossRef]
47. Syed, N.F.; Shah, S.W.; Shaghaghi, A.; Anwar, A.; Baig, Z.; Doss, R. Zero Trust Architecture (ZTA): A Comprehensive Survey. *IEEE Access* **2022**, *10*, 57143–57179. [CrossRef]
48. Vanickis, R.; Jacob, P.; Dehghanzadeh, S.; Lee, B. Access Control Policy Enforcement for Zero-Trust-Networking. In Proceedings of the 2018 29th Irish Signals and Systems Conference (ISSC), Belfast, UK, 21–22 June 2018; pp. 1–6. [CrossRef]
49. Salahdine, F.; Kaabouch, N. Social engineering attacks: A survey. *Future Inter.* **2019**, *11*, 89. [CrossRef]
50. Weaver, A. Biometric authentication. *Computer* **2006**, *39*, 96–97. [CrossRef]
51. Arnau-González, P.; Katsigiannis, S.; Arevalillo-Herráez, M.; Ramzan, N. BED: A new data set for EEG-based biometrics. *IEEE Internet Things J.* **2021**, *8*, 12219–12230. [CrossRef]
52. Arteaga-Falconi, J.S.; Al Osman, H.; El Saddik, A. ECG authentication for mobile devices. *IEEE Trans. Instrum. Meas.* **2015**, *65*, 591–600. [CrossRef]
53. Zeuner, J.; Pitsios, I.; Tan, S.H.; Sharma, A.N.; Fitzsimons, J.F.; Osellame, R.; Walther, P. Experimental quantum homomorphic encryption. *Npj Quantum Inf.* **2021**, *7*, 25. [CrossRef]
54. Fei, T.; Wang, W. The vulnerability and enhancement of AKA protocol for mobile authentication in LTE/5G networks. *Comput. Net.* **2023**, *228*, 109685. [CrossRef]
55. Du, M.; Ma, X.; Zhang, Z.; Wang, X.; Chen, Q. A review on consensus algorithm of blockchain. In Proceedings of the 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Banff, AB, Canada, 5–8 October 2017; pp. 2567–2572. [CrossRef]
56. Schär, F. Decentralized Finance: On Blockchain- and Smart Contract-Based Financial Markets. In Proceedings of the Federal Reserve Bank of St. Louis Review, Second Quarter 2021, 31 March 2021; pp. 153–174. Available online: https://ssrn.com/abstract=3571335 (accessed on 30 May 2024).
57. Cinque, M.; Esposito, C.; Russo, S. Trust Management in Fog/Edge Computing by Means of Blockchain Technologies. In Proceedings of the 2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), Halifax, NS, Canda, 30 July–3 August 2018; pp. 1433–1439. [CrossRef]

58. Zhaofeng, M.; Xiaochang, W.; Jain, D.K.; Khan, H.; Hongmin, G.; Zhen, W. A Blockchain-Based Trusted Data Management Scheme in Edge Computing. *IEEE Trans. Ind. Inform.* **2020**, *16*, 2013–2021. [CrossRef]
59. Krummacker, D.; Veith, B.; Lindenschmitt, D.; Schotten, H.D. DLT architectures for trust anchors in 6G. *Ann. Telecommun.* **2023**, *78*, 551–560. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

electronics

MDPI

Article

LoRa Radio Frequency Fingerprinting with Residual of Variational Mode Decomposition and Hybrid Machine-Learning/Deep-Learning Optimization

Gianmarco Baldini [1,*] and Fausto Bonavitacola [2]

[1] European Commission, Joint Research Centre, 21027 Ispra, Italy
[2] Piksel, 20126 Milan, Italy; fausto.bonavitacola@ext.ec.europa.com
* Correspondence: gianmarco.baldini@ec.europa.eu

Abstract: Radio Frequency Fingerprinting (RFF) refers to the technique for identifying and classifying wireless devices on the basis of their physical characteristics, which appear in the digital signal transmitted in space. Small differences in the radio frequency front-end of the wireless devices are generated across the same wireless device model during the implementation and manufacturing process. These differences create small variations in the transmitted signal, even if the wireless device is still compliant with the wireless standard. By using data analysis and machine-learning algorithms, it is possible to classify different electronic devices on the basis of these variations. This technique has been well proven in the literature, but research is continuing to improve the classification performance, robustness to noise, and computing efficiency. Recently, Deep Learning (DL) has been applied to RFF with considerable success. In particular, the combination of time-frequency representations and Convolutional Neural Networks (CNN) has been particularly effective, but this comes at a great computational cost because of the size of the time-frequency representation and the computing time of CNN. This problem is particularly challenging for wireless standards, where the data to be analyzed is extensive (e.g., long preambles) as in the case of the LoRa (Long Range) wireless standard. This paper proposes a novel approach where two pre-processing steps are adopted to (1) improve the classification performance and (2) to decrease the computing time. The steps are based on the application of Variational Mode Decomposition (VMD) where (in opposition to the known literature) the residual of the VMD application is used instead of the extracted modes. The concept is to remove the modes, which are common among the LoRa devices, and keep with the residuals the unique intrinsic features, which are related to the fingerprints. Then, the spectrogram is applied to the residual component. Even after this step, the computing complexity of applying CNN to the spectrogram is high. This paper proposes a novel step where only segments of the spectrogram are used as input to CNN. The segments are selected using a machine-learning approach applied to the features extracted from the spectrogram using the Local Binary Pattern (LBP). The approach is applied to a recent LoRa radio frequency fingerprinting public data set, where it is shown to significantly outperform the baseline approach based on the full use of the spectrogram of the original signal in terms of both classification performance and computing complexity.

Keywords: security; wireless communication; deep learning; machine learning

1. Introduction

Radio frequency fingerprinting is a technique for the identification and authentication of wireless devices, which has received increased attention from the research community in recent years. As described in [1], the security of wireless networks is a very important aspect to ensure the trust of wireless services, which are now pervasive in our society. In this context, networking and security protocols assume that the identities of users and devices are distinct and unique, but this can be violated by cybersecurity attackers, who

can spoof the identity of a wireless device. To address such attacks, cryptographic-based approaches are proposed in the research literature and communication infrastructures. Such approaches usually have the goal of binding a device or a node in a wireless communication system with a cryptographic identity of a certificate. On the other side, the deployment of cryptographic frameworks can be costly because of the need to implement processes like key distribution and key management, which can be particularly cumbersome in Internet of Things (IoT) systems like the LoRa devices considered in this study because of the computing limitations of such devices [2]. Recently, RFF has been proposed by various researchers as an alternative or complementary means of cryptographic identification for wireless communications [1,3,4].

This technique is based on the exploitation of the intrinsic physical differences among wireless devices supporting the same wireless standard. These physical differences are created during the manufacturing process or the design process due to the use of different materials or different configurations of electronic components, especially in the radio frequency front-end. Even if such differences do not usually hamper compliance to the wireless standard, they alter the signal in space generated by the wireless device in such a way that the device itself can be identified by the analysis of the signal. The advantage of RFF in comparison to cryptographic techniques is that it does not require the distribution and secure storing of cryptographic materials (which is particularly beneficial for IoT environments) and that it is based on the unique physical properties of the wireless device [3]. On the other side, the identification is based on probability derived from the matching of the fingerprints, and research is still progressing in improving such probability in terms of accuracy and required computing time [1,4], which is also the objective of this study.

The analysis of the signal and the identification of the wireless devices can be implemented as a pattern recognition problem, and various techniques can be used, including the application of Machine Lerning (ML), DL algorithms to a feature space created from the application of hand-crafted features (e.g., entropy, variance) to the time or spectral domain representation of the signal. Some studies have implemented a pre-processing step on the signals generated by the emitters to enhance the performance of ML. For example, the authors in [5] applied VMD to decompose the signal in modes to generate a more comprehensive feature space. The approaches based on ML are usually computing efficient, but they rely on the identification of hand-crafted features, which may not be optimal for the specific data set. DL has demonstrated a superior classification performance to the so called 'shallow' ML algorithms in various RFF studies, as described in a recent survey [3]. The advantage of DL is also that features are extracted automatically. On the other side, DL requires significant computing resources, particularly for the training time. Some DL algorithms and architectures like CNN are particularly powerful for the classification of images, and some recent studies have converted the 1D initial signal from the emitter to a 2D image using various techniques, like the application of the spectrogram [6], but they may significantly increase the size of the input data to the DL algorithm, thus increasing the computing time even more.

This paper proposes a hybrid (combination of ML and DL) approach where various elements are brought together to enhance the performance of RFF:

- The VMD is applied as a pre-processing step to highlight the fingerprints from the signal. The residual of the application of VMD to the signals of the emitters to be classified is used for RFF rather than the generated modes, as is commonly done in the literature. This pre-processing step is proposed to remove the signal modes, which are common to all the emitters and highlight the fingerprints for further processing.
- The spectrogram is used to obtain a 2D representation of the signal.
- On the spectrograms, a novel image reduction step is applied, which is based on the application of Shannon entropy and LBP image processing features to select the most discriminating Region Of Interest (ROI) so that they can be assembled to generate a new 2D input to a multi-headed CNN. This step improves classification performance and decreases its computing time.

- A multi-headed CNN is used to perform the classification and implement the final step of RFF.

To the knowledge of the authors, this combination of elements is novel for the problem of RFF. The proposed approach is applied to a recent public data set of LoRa devices [6].

This paper has the following structure: Section 2 summarizes the research literature on RFF, with a particular focus on the use of mode decomposition methods for RFF. Section 3 describes the material and methods used to implement the proposed approach. In particular, this section describes the elements in the methodology: the structural diagram of the procedures, the used public data set, the definition of the LBP and VMD, the computing platform used to perform the experiments, CNN architectures, and the evaluation metrics. Section 4 provides the results of the proposed approach on the described data set and conducts an analysis of the findings. Finally, Section 5 provides the conclusions and outlines future developments.

2. Related Work

The literature on radio frequency fingerprinting is quite vast, and many studies have been published in recent years, as described in recent surveys like [3,7]. While initial studies [8] used handcrafted features in combination with ML to implement the classification, the application of DL to RFF was demonstrated to be highly successful in [9–11]. For this reason, this study also uses DL but in combination with the application of VMD.

The rest of this section focuses on the application of VMD to RFF. The predominant use of VMD in this context is to decompose the original signal in modes on which handcrafted features are applied to generate a feature space. ML algorithms are then used to perform the classification and identification of wireless devices. Two different data sets of DSRC (ITS G5 standard [12] used for vehicular applications) devices and IoT devices were used in the study. This is the approach used in [13], where VMD was used in combination with features like variance, entropy, skewness, and kurtosis to create a feature space where Support Vector Machine (SVM) and K-Nearest Neighbor (KNN) were applied. The results from [13] show that this approach was able to outperform both the use of the original signal representation in the time domain and frequency domain or the use of Empirical Mode Decomposition (EMD), which is another mode decomposition algorithm, which predates VMD and which was used in [14].

Similarly, the authors in [5] used VMD in combination with spectral features and KNN on a data set of emitter devices to implement RFF. The study presented in [5] considered both single-hop emitter devices and multi-hop relaying devices with a comparable performance.

Another study exploiting VMD for RFF was [15], where devices implementing the Bluetooth wireless standard [16] were used. As in the previously cited paper, Higher Order Statistics (HOS), namely variance, skewness, and kurtosis, are calculated from the instantaneous amplitude, frequency, and phase of the modes generated by the application of VMD. The Linear Support Vector Machine (LSVM) was used to perform classification.

A similar work is [17], where handcrafted features were applied to the transient portion of the signal generated by Bluetooth devices to generate a feature space, where the LSVM was used to implement RFF.

Another approach based on VMD, but using a different set of features, was proposed in [18], where the Hilbert transform was applied to the modes originating from the VMD decomposition of signals generated by ZigBee devices. The Hilbert transform is used to generate images from the modes on which Histogram of Oriented Gradient (HOG) features are applied to generate a feature space. Then, KNN is applied to implement RFF and classify the ZigBee devices. As described in the subsequent sections, this study also applied images-related features, but it uses LBP instead of HOG to generate a more compact feature space. In addition, the LBP features are applied to a spectrogram instead of the result of the Hilbert transform, and they are not used for the main classification task as in [18], but

to reduce the size of the images given as input to the DL algorithm. The VMD is also used to generate the residual rather than the modes themselves, and DL was not used in [18].

Another set of features (still in combination with VMD) was used in [19], where the Teager operator was applied and the energy and box dimension were extracted as fingerprint features to be given in input to SVM for the classification of the emitters.

Reviewing the studies presented above, we can conclude that no study has applied the analysis to the residual of the VMD rather than the modes generated by the VMD. In addition, all the papers presented so far did not apply DL in combination with VMD.

The application of DL to RFF is quite recent, but it has shown a higher accuracy in comparison to ML algorithms like SVM or KNN used in the references cited above at the cost of an increased computing complexity and time. In particular, a popular approach in the research literature is to combine time-frequency distributions with DL algorithms, which are particularly effective in the analysis of images like CNN. This approach was used with various time-frequency transforms. In [6], which proposed the same data set used in this study, a spectrogram was used in combination with CNN, but the VMD was not used to pre-process the signals. This step provides a significant advantage, as shown in Section 4 of this paper. A similar approach with the spectrogram was also used in [20], still on LoRa devices without using the VMD as well. The relevance of the application of RFF to LoRa devices is also shown in [21], where CNN was used together with a spectral domain (based on the application of the Fast Fourier Transform (FFT) in the study) representation of the signal as it was shown to outperform the original time domain representation.

One key issue of the application of time-frequency transform together with DL is that the size of the input data can be significantly large (in comparison with the original time representation). It would then be beneficial to reduce the input of images created with the application of the time-frequency transform to a smaller size while preserving the discriminating power. One possibility would be to use computing-efficient features and 'shallow' ML algorithms to identify ROIs, which are more discriminating than others, and then give only the assembled ROIs to the DL algorithm. This approach was applied recently in [22] but only to 1D spectral domain representations (i.e., application of the FFT) and not to time-frequency transforms. In other fields, the use of a hybrid approach, which combines handcrafted features and DL to improve computing efficiency and classification performance, has recently been proposed for generic image processing in [23], though with different features than in this study.

In relation to the cited literature, this paper proposes a novel approach, based on the application of image processing features, to reduce the spectral domain representation given in input to the DL algorithm to enhance classification accuracy and reduce computing time.

3. Materials and Methods

3.1. Methodology

The overall methodology of the proposed approach is shown in Figure 1, where the structural diagram of procedures is pictorially represented. The different color codes are used to show the different paths to generate the results, which are compared and analyzed in Section 4. The paths are, respectively: (1) spectrogram with the original signal (blue), (2) spectrogram with the residual from the VMD (black), (3) reduced spectrogram (selected segments) with the original signal (red), and (4) reduced spectrogram (selected segments) with the residual from the VMD (magenta). The spectrogram is reduced thanks to the technique described in the following steps.

In the first step, the signals of the LoRa devices are synchronized and normalized. A set of $N_D = 10$ LoRa devices are used in this study, as described in Section 3.2. For each LoRa device, $N_B = 500$ bursts were generated for a total of 5000 bursts. Then, the VMD was applied to each burst with a value of $\alpha = 2000$ and the number of modes, N_M, which is a hyper-parameter in this study. The residual was extracted from the signals after the main modes were identified. For this purpose, the MATLAB 2022b *vmd* function was used

from the Signal Processing toolbox, which is in turn based on the theoretical framework defined in [24].

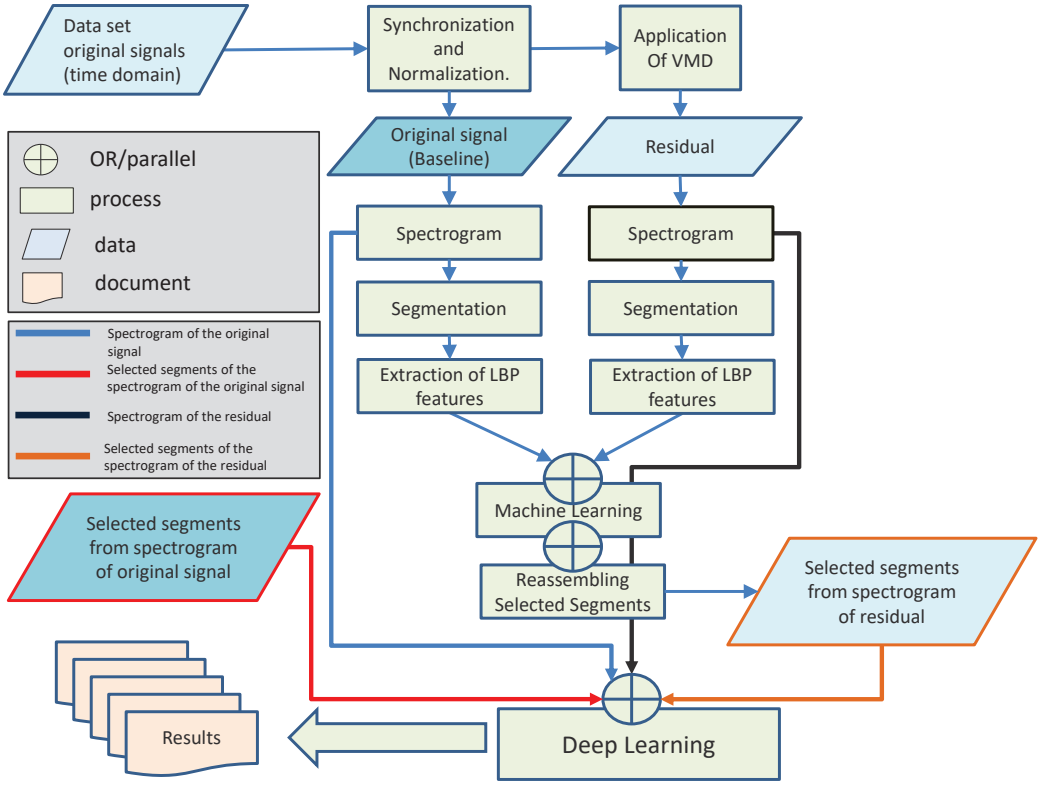
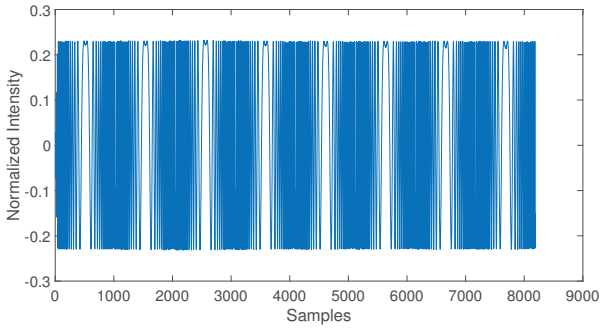


Figure 1. Overall structural diagram of the procedures comprising the proposed approach.

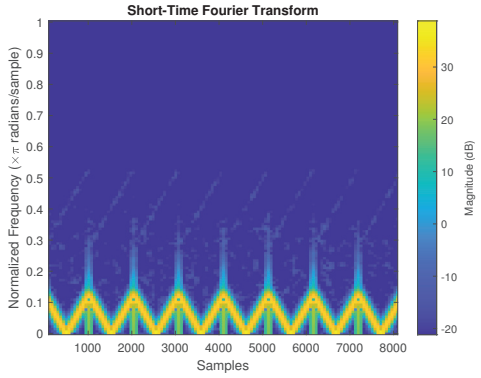
The value of α was identified on the basis of the LoRa modulation and the settings of the data set. Because the study is focused on the residual, the number of modes is kept low within the following range: $N_M = [2, \ldots, 5]$. A large number of modes may include the fingerprints, which are needed to classify the LoRa devices, while a relatively small number of modes would include the common elements of the signals, which we want to remove. Another hyper-parameter in the study is the window size, W_S, used to generate the spectrogram. This value was chosen following the settings defined in [6], where the data set was initially proposed. In [6], a Gaussian window of $W_S = 256$ was used, with an overlapping of 128. In this study, we also added the Gaussian window of $W_S = 128$ with an overlapping of 64.

A representation of the signal is shown in Figure 2, together with one-sided spectrograms with $W_S = 256$. The spectrograms are generated by using the MATLAB 2022b *stft* function. These images are provided only as a reference. As explained in the subsequent sections of this paper, a two-sided centered spectrogram is used as input to the DL algorithm to enhance and center the relevant content of the spectrogram.

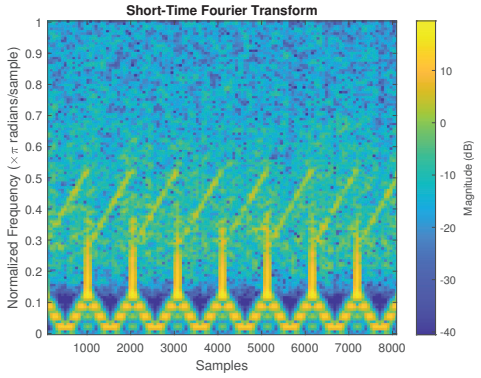
Then, the other element of the proposed approach is applied. Each spectrogram is converted to a grayscale image and sub-divided into 32 rectangles. The initial image is divided by 4 for the frequency dimension and 8 for the time dimension. In each square, the LBP values were calculated in addition to the Shannon entropy of the rectangular image. This creates a feature space of size $N_D * N_B * 60$ (59 features of the LBP plus the Shannon entropy) for each rectangle. For each of the 32 rectangles, the accuracy was calculated using a Decision Tree (DT) algorithm,with the maximum number of splits equal to 12 (this was found to be the optimal value in a range from 1 to 20).



(**a**) Original LoRa signal in the time domain.



(**b**) One-sided spectrogram of the original signal.



(**c**) One-sided spectrogram of the residual obtained from the application of VMD with $N_M = 4$.

Figure 2. One-sided spectrograms of the original signal and its residual from the application of the VMD with $W_S = 256$.

The 8 best results are adopted to represent the optimal set of rectangles, which are reassembled to create a new image, whose size is one-quarter of the original one. This novel step not only decreases the overall classification time, because the DL algorithm has to process smaller images, but it may also enhance the classification performance. This assumption will be proven by the findings shown in the Section 4. Note that the value of 8 was chosen as a compromise between the need to obtain a significant reduction in the input data while still preserving significant data for the classification. This could also be a hyper-parameter to tune, but we preferred to focus the analysis on other hyper-parameters, namely the number of modes, N_M, in the VMD and the window size, W_S, of the spectrogram.

3.2. Data Set

This section describes the data set used in this study. The data set was generated by the authors of [6] and it is based on the identification of wireless devices supporting the LoRa wireless communication standard [25].

LoRa uses Frequency Shift Chirp Modulation (FSCM), which uses chirps for communication [26]. In the context of wireless communications, a chirp can be defined as a signal in which the communication frequency increases (up-chirp) or decreases (down-chirp) with time. The instantaneous frequency of the LoRa signal changes continuously over time, and a basic LoRa symbol (up-chirp) can be written as:

$$u(t) = A^{j\times 2\pi\left(-\frac{B}{2}+\frac{B}{2T}t\right)t}, \tag{1}$$

where A and B represent, respectively, the amplitude and signal bandwidth of the LoRa signal. T is the duration of the LoRa symbol, which can be represented as:

$$T = 2^{SF}/B, \tag{2}$$

where SF is the spreading factor. There are eight repeating up-chirps at the beginning of a LoRa packet called the preamble, which is identical in every LoRa packet regardless of the device type. Then, the preamble can be used for RFF because it is invariant to the data transmitted by the LoRa devices, which may bias the extraction of the RFF.

The signals collected by the authors of [6] need to be pre-processed so that they can be exploited for the RFF implementation. In particular, they need to be synchronized, compensated for Carrier Frequency Offset (CFO), and normalized. The reasons for these pre-processing steps are to avoid the introduction of bias in the implementation of RFF because the DL algorithm used for classification may be confused by the different power levels of the signal (which requires normalization) and misrepresent the wireless devices. In a similar way, the lack of synchronization or CFO may cause similar effects.

The phase information of the LoRa signal in the data set is noisy and can be affected by several issues, such as phase noise, CFO, and sampling time offset. Therefore, only the spectrogram is used in this study, which is consistent with the methodology proposed in [6]. This study uses the fingerprints from 10 LoRa devices from the data set provided by the authors in [6] from the training data set. In the data set, 500 bursts of each LoRa device are provided. Then, a total of 10×500 burst (5000 bursts) are used in the study. Each burst has a size of 8192 samples. This number of devices is consistent with other similar studies using LoRa devices for RF fingerprints, like [27], where 10 devices were also used, or other RFF studies like [13],where 11 devices were used, and [9], where 5 devices were used.

3.3. Local Binary Pattern

This section provides the definition of LBP as it was used in this paper to implement the dimensionality reduction. The LBP is a nonparametric operator for the description of local image features. It has the benefit of being rotation invariant and translation invariant [28]. The LBP of a cell is calculated in the following way. For a center pixel in an image or a cell (in this study, the section of the image identified in Section 3 and the cell size and image size are the same), and the pixel is compared to each of its 8 neighbors. Where the value of the center pixel is greater than the neighbor's value, a value of 0 is recorded. Otherwise, a value of 1 is recorded. This gives an 8-digit binary number as in the following Equation (3), where x_c and y_c are the coordinates of the center pixel and i_n is the gray value of the pixels of its eight neighbors; N is the number of neighbors and the neighbor is identified with the parameter n.

$$LBP(x_c, y_c) = \sum_{n=0}^{7} Ne(i_n - i_c) * 2^n. \tag{3}$$

In the next step of the LBP algorithm, the histogram is computed, over the cell, of the frequency of each "number" occurring. Then, the normalized histograms of all cells are concatenated to give the feature vector for the entire window.

The rationale for using the LBP in this context is that the differences in the spectrogram may indicate the slight variations in time or frequency between the signals generated by the wireless devices that are related to fingerprints. The additional advantage of LBP in the comparison of other image processing features (e.g., Scale-Invariant Feature Transform (SIFT) or HOG) is that the computing complexity is relatively limited. For the sections of the spectrogram images considered in this study, the generated feature space is also limited (i.e., 59 features).

The MATLAB 2022b implementation of LBP provided by the authors of [29] was used in this study.

3.4. Deep-Learning Architecture

This section describes the DL architecture used in this study.

A CNN architecture with two convolutional layers and multi-head attention was used to implement the classification, as shown in Figure 3.

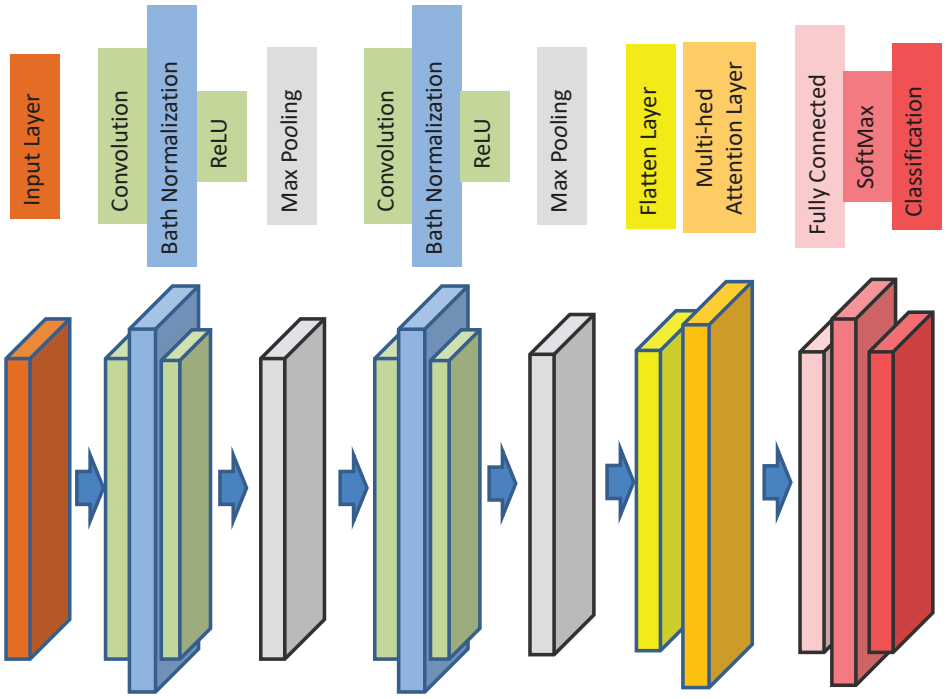


Figure 3. The architecture of the Convolutional Neural Network used in this study, with multi-head attention.

The values of the parameters of the DL architecture used in this study are shown in Table 1. Some values of the CNN parameters were guided by the size of the spectrogram image itself, like the width and the filter size (i.e., 32, as shown in the subsequent table), while other parameters were tuned according to a range of values to achieve the optimal performance for the majority of cases. In particular, the Adaptive moment estimation (Adam) solver was chosen over the Stochastic Gradient Descent with Momentum (SGDM) and the Root mean square propagation (RMSProp). The maximum pooling operation was preferred to the average pooling operation, and the optimal number of heads was chosen in a range of 4 to 10 in steps of 2. Rectified Linear Unit (ReLu) was used as an activation function. Other DL architectures could also be used, including the pre-trained CNN models, but we would like to highlight that the purpose of this study was not to compare DL architectures (which also include specific constraints on the size of the input images and it may not be suitable for this specific context), but to evaluate the performance improvement of the use of the residuals from VMD and the image reduction step with the LBP features and the Shannon entropy feature.

The evaluation metrics used in this study are accuracy, F-score, and the overall computation time. In this study, the computational time of the proposed approach is the sum of the computational times of the different methodological steps described in Section 3.1.

The accuracy is defined by the following equation:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}. \quad (4)$$

The *F-score* is the harmonic mean of precision and recall and it is defined by the following equation:

$$F\text{-}score = \frac{2 * TP}{2 * TP + FP + FN}, \tag{5}$$

where *TP* is the number of True Positives, *TN* is the number of True Negatives, *FP* is the number of False Positives, and *FN* is the number of False Negatives.

The computational time is calculated in a normalized way using as a unitary value the smallest computing time among the various steps described in Section 3.1.

Table 1. List of the CNN parameters used in this study.

CNN Parameter	**Value**
CNN solver algorithm	Adam
Number of Convolutional Layers	2
Width and filter size of the 1st convolutional layer	32,32
Number of filters in the 1st convolutional layer	32,32
With and filter size of the 2nd convolutional layer	12,12
Number of filters in the 2nd convolutional layer	32,32
1st pooling layer	Max pooling (4,4) with Stride (2,2)
2nd pooling layer	Max pooling (4,4) with Stride (2,2)
Activation function	ReLu
Number of heads in the attention layer	8
Number of channels for keys and queries in the attention layer	64
Maximum number of epochs	180

3.5. Variational Mode Decomposition

The VMD technique belongs to the family of mode decomposition algorithms and it was initially proposed in [24]. In comparison to other mode decomposition algorithms like EMD, VMD has a number of significant advantages, including the aspect that it is non-recursive, adaptive, and based on a quasi-orthogonal signal decomposition.

The main goal of the VMD algorithm is to decompose a multi-component signal in a discrete number of sub-signals, which are called modes or Intrinsic Mode Functions (IMF) with limited bandwidth in the spectral domain, assuming each mode to be mostly compact around a center pulsation, ω_k. In the rest of this paper, we use the term mode to indicate the IMF.

The bandwidth of a mode, u_k, is assessed in the following way [24]: (a) using the Hilbert transform, calculate for each mode the associated analytical signal and obtain a unilateral frequency spectrum; (b) for each mode, shift the mode's frequency spectrum to *baseband* by mixing with an exponential, tuned to the respective estimated center frequency; and (c) estimate the bandwidth through the H^1 Gaussian smoothness of the demodulated signal (the squared L^2-norm of the gradient). Therefore, for a given signal, $s(t)$, with k modes, u_k, and related center frequencies, ω_k, the variational problem is expressed as follows:

$$\min_{\{u_k\},\{\omega_k\}} \left\{ \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\}, \tag{6}$$

where δ is the Dirac distribution, $*$ denotes convolution, and s is subject to $\sum_k u_k = s$.

The VMD has two main hyper-parameters to tune. The first hyper-parameter is the α parameter from Equation (15) in [24], which is the balancing parameter of the data-fidelity constraint, and the second hyper-parameter is the number of modes, k.

The hyper-parameter α is usually related to the frequency structure of the signal itself, which is known in this case on the basis of the LoRa standard. [25] A value of $\alpha = 2000$ was set for this study. The number of modes, N_M (which is mapped to k of the previous equation), is instead a parameter to be tuned.

As described in Section 2, VMD has been used in wireless communication to extract features from the identified modes [13] or as a denoising method where the residual is considered noise to be removed as it affects the quality of the signal under analysis [30]. This paper proposes an alternative way to use VMD based on the assumption that the Radio Frequency (RF) fingerprints may not be linked to the first modes extracted by the VMD, which can be common to all devices on the basis of the signal structure and modulation. Then, the residual may contain the discriminating features needed to identify the wireless devices, and classification algorithms (like the DL algorithms used in this study) can be applied only to the residual with more success than the original signal. This assumption is going to be proven for this data set in Section 4.

In this paper, we have used the *vmd* function of the MATLAB Signal Processing toolbox, which is in turn based on the theoretical work by the authors of [24].

3.6. Computing Platform

The computing platform used to conduct the experimental evaluation is a workstation equipped with an Intel I9, 64 Gbytes of Random Access Memory (RAM), and the Graphic Processing Unit (GPU) NVIDIA Quadro RTX4000. MATLAB was used to perform the scientific computations with the Signal Processing toolbox and the Machine-Learning and Deep-Learning toolbox to implement the multi-head attention CNN.

4. Results

This section provides the results from the application of the proposed approach to the data set described in Section 3.2. As described previously, there are two main hyper-parameters to tune, W_S and N_M, with a limited number of range values. For W_S, the same value identified by the authors of the data set in [6] were used (i.e., $W_S = 256$), but the value of $W_S = 128$ was also used for comparison. For both values, an overlap, which was half of W_S, was used in the application of the spectrogram (this is also consistent with the approach used in [6]).

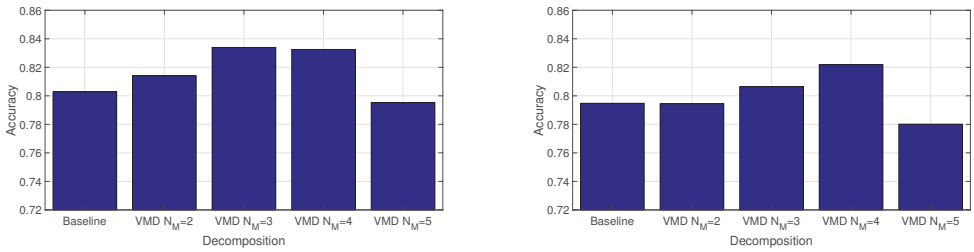
This section is structured in four sub-sections. The first, Section 4.1, provides the results of the application of VMD to calculate the residuals on which the DL is applied. The second, Section 4.2, presents the results for the dimensionality reduction step. Section 4.3 compares the detailed results among the various approaches. Finally, Section 4.4 provides the confusion matrices obtained for some values of the hyper-parameters.

4.1. Application of VMD

In the application of VMD, the range of values for N_M is [2,3,4,5]. The reason for such small values of N_M is because the goal is to remove (with the residual) the common elements of the signal among the wireless devices. Then, a large number of modes may also remove the fingerprints that we would like to preserve.

This assumption is validated by the following Figure 4a,b, respectively for spectrograms at $W_S = 128$ and $W_S = 256$ for different values of N_M. Note that these figures are based on a partial application of the approach where the dimensional reduction with LBP is not implemented. Figure 4a,b show that the use of the residual from the VMD is able to outperform the direct application of the DL algorithm on the spectrogram of the original signal for both values $W_S = 128$ and $W_S = 256$. The potential issue is to select the appropriate value of N_M to achieve the maximum value of the accuracy. For $W_S = 128$, the optimal value of N_M is 3 (even if the value of $N_M = 4$ achieves a similarly high accuracy). For $W_S = 256$, the optimal value of N_M is 4.

In both cases, we notice that a higher value of $N_M = 5$ decreases the accuracy.



(**a**) Based on spectrogram with W_S = 128. (**b**) Based on spectrogram with W_S = 256.

Figure 4. Comparison of the accuracies obtained for different values of N_M (related to the application of VMD) and W_S (related to the application of the spectrogram).

4.2. Dimensionality Reduction

The previous results were obtained without the application of the dimensionality reduction step. We show with the subsequent results that the use of LBP can reduce the size of the input spectrogram images while preserving the discriminating features and decrease the DL classification time.

To enhance the image content and to place the richest content (from the discriminating point of view) in the center of the figure, a two-sided, centered spectrogram is used instead of the one-sided spectrogram shown in the previous Figure 2b,c.

Figure 5 shows the two-sided centered spectrogram of the original signal with $W_S = 256$. The spectrogram was converted to a gray-scale image for the application of LBP.

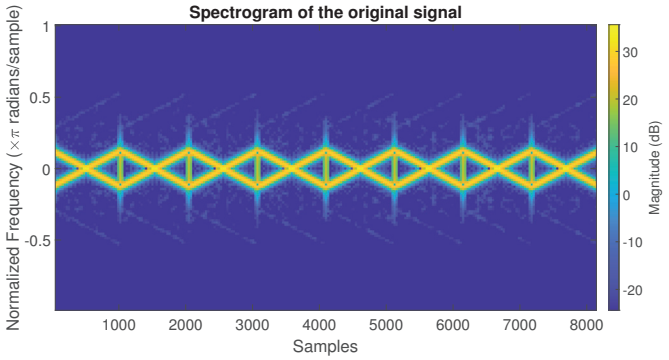
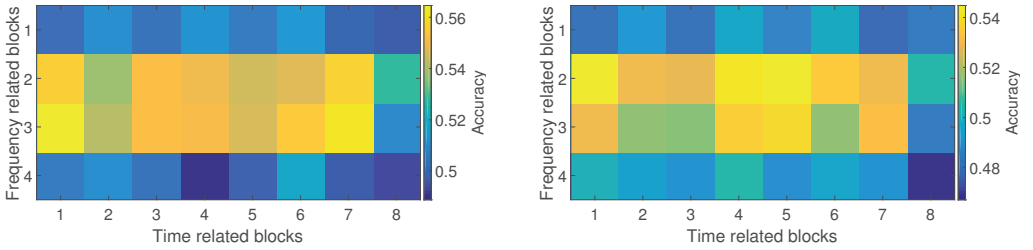


Figure 5. Two-sided centered spectrogram of the original signal used in this study as input to the LBP feature extraction.

As described previously, the spectrogram images are split into areas of the same size on which the LBP features are extracted. Each spectrogram image (both the one originating from the baseline and the one originating from the residual after the application of VMD) was divided into 32 areas: 8 divisions along the time domain and 4 divisions along the frequency domain. These subdivisions were executed in this way because of the structure of the LoRa preamble in the time domain (thus a value of 8) and because most of the energy in the centered two-sided spectrogram is along the frequency center line. Then, a value of 8 was used to distinguish between the areas where most of the energy is concentrated and the areas where it is not. A larger number of areas would go against the rationale that on each area the LBP features have to be extracted, which requires a statistically significant number of points to calculate the LBP. In addition to the features generated by LBP (59 in this case), the Shannon entropy of each area is also calculated for a total of 60 features. Then, the DT algorithm was applied to the feature space generated for each area and the resulting accuracy was used to identify the potential areas, which are more promising for the future steps. The DT was used with a maximal number of decision splits equal to 12 and the split criterion with the Gini's diversity index.

Figure 6a,b below show the accuracy obtained using the LBP, respectively for the spectrogram of the original signal and the spectrogram of the residual using $W_S = 128$. The residual was calculated with N_M = 4.

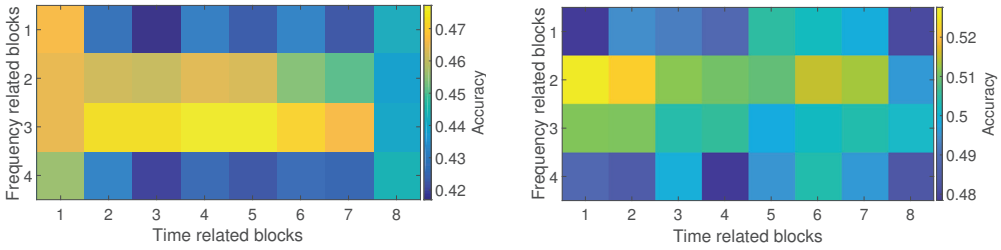
The figures show that the areas with the highest accuracy are along the central frequency region, where most of the energy of the signal is concentrated. Then, the result shown in Figure 6a,b is not surprising, but the proposed approach based on the application of LBP also shows the most promising areas along the time dimension in the central frequency region, which is additional information not provided by the spectral energy density.



(**a**) Original signal. (**b**) Residual of VMD with $N_M = 4$.

Figure 6. Accuracy maps obtained with the DT algorithm and LBP features plus Shannon entropy for each of the spectrogram areas with $W_S = 128$.

Similar results are obtained using $W_S = 256$, and they are shown in Figure 7a,b, respectively for the spectrogram of the original signal and the spectrogram of the residual using N_M=4. The identification of promising areas is even more evident for $W_S = 256$ than for $W_S = 128$. In particular, it is shown that the first areas (first along the time dimension) in Figure 7b have significantly higher accuracy than the other areas. This aspect may point to a more effective selection of the areas for $W_S = 256$. This assumption will be demonstrated in the subsequent results of this section. It is noted that the accuracy shown in the previous four Figures 6a,b and 7a,b is low in absolute terms (highest values around 55% accuracy), but it is not important in the overall context of the approach because the accuracy is only used to select the most promising areas for further processing by the CNN.



(**a**) Original signal. (**b**) Residual of VMD with $N_M = 4$.

Figure 7. Accuracy maps obtained with the DT algorithm and LBP features plus Shannon entropy for each of the spectrogram areas with W_S = 256.

Then, the 8 areas with the highest accuracy over the overall set of 32 areas are selected and assembled to generate reduced images (which are 1/4) of the original size. The 8 areas are selected both for the spectrograms created from the original images and for the spectrogram created from the residuals on the application of VMD. The reduction factor of 4 (8 areas out of 32 areas) is chosen to obtain a significant reduction of the spectrogram images, which can justify the application of the proposed approach as it would decrease the classification time with CNN. As reported in [31], the computational complexity of the CNN convolutional layer is quadratic with the size of the input image. Then, we approximately obtain a 1/16 reduction of the classification time.

4.3. Detailed Results

Table 2 below shows the detailed results for the proposed approach and the baselines (the original signal) for both the window sizes, W_S, considered in this study for the obtained accuracy, F-score, and computing time.

Table 2. Detailed results of the proposed approach.

Method	**Accuracy**	**F-Score**	**Normalized Computation Time**
$W_S = 128$			
Spectrogram of the original signal	0.8028	0.8018	146
Spectrogram of the residual ($N_M = 3$)	0.8339	0.8342	199 (63 VMD, 136 CNN)
Reduced spectrogram of the original signal	0.866	0.865	41 (5 Feature, 1 ML, 35 CNN)
Reduced spectrogram of the residual ($N_M = 4$)	0.8728	0.8732	104 (63 VMD, 5 Feature, 1 ML, 35 CNN)
$W_S = 256$			
Spectrogram of the original signal	0.7947	0.7931	150
Spectrogram of the residual ($N_M = 4$)	0.8218	0.8202	209 (63 VMD, 146 CNN)
Reduced spectrogram of the original signal	0.907	0.89	41 (5 Feature, 1 ML, 35 CNN)
Reduced spectrogram of the residual ($N_M = 4$)	0.919	0.9165	102 (63 VMD, 5 Feature, 1 ML, 33 CNN)

Table 2 shows that the proposed approach is able to significantly outperform the application of CNN directly on the original spectrograms as done in [6] for both window sizes, W_S. For example, for $W_S = 128$ the accuracy obtained using the residual is 0.8339 against the accuracy of 0.8028 obtained using the original signal. The application of the reduction step using LBP significantly increases the accuracy when it is applied to the original signal (accuracy of 0.866) or to the residual (accuracy of 0.8728).

The use of a larger window size (i.e., $W_S = 256$) produces even higher accuracy values than the use of $W_S = 128$, presumably because a large window size generates a sharper frequency resolution in the spectrogram, which can be exploited by the CNN. In fact, we obtain an accuracy of 0.919 using the reduced spectrogram of the residual and an accuracy of 0.907 using the reduced spectrogram of the original image. The F-score results are usually coherent with the accuracy results.

The total computation time is shown in the last column of Table 2. The time is normalized to the smallest computation time (which is the execution of the DT algorithm on the feature space generated with the application of LBP and Shannon entropy). The computation time is the total of the specific computation times needed for each step in the procedural path chosen in the methodology described in Section 3.1. The computation time of each step is shown in parenthesis after the total in the last column of Table 2. ML identifies the application of the DT algorithm.

It can be noted that the reduction of the images significantly reduces the CNN classification time, which is generally quadratic with the size of the images. The computational time spent on the application of LBP and Shannon entropy is also quite small in comparison to the CNN classification time, and it does not significantly impact the overall computation time. Apart from the CNN classification time, the largest contributor is the execution of the VMD algorithm to all the signals of the data set. The VMD execution is relatively long for this specific data set because each sample is a time series of length 8192 samples and the computing complexity of VMD is related to the size of the time series on which to

operate [24]. This is one of the limitations of the proposed approach. Another limitation is that the optimal number of modes is not known a priori.

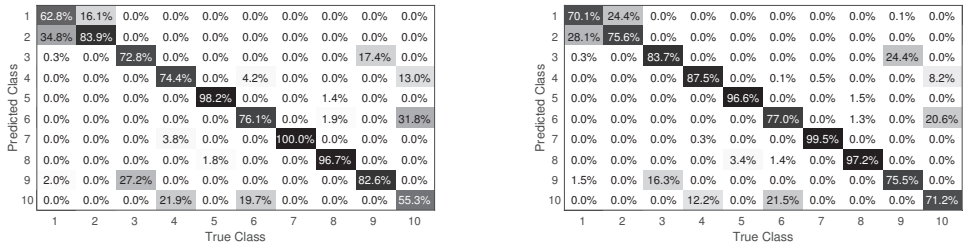
4.4. Confusion Matrices

The objective of this sub-section is to provide confusion matrices for specific sets of values of the hyper-parameters W_S and N_M to complement the accuracy results presented in the previous sub-sections. Confusion matrices provide more detailed information in comparison to the overall accuracy as they compare the predicted values against the true values for each wireless emitter.

Figure 8a,b show, respectively, the confusion matrices for the spectrogram calculated with $W_S = 128$ and applied to the original signal and the residual after the application of VMD.

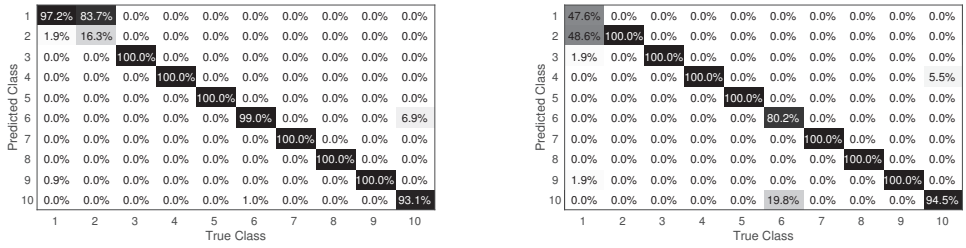
Figure 9a,b show, respectively, the confusion matrices for the spectrogram calculated with $W_S = 256$ and applied to the original signal and the residual after the application of VMD.

It can be seen that the confusion matrices are aligned with the results presented in Table 2 because the accuracy obtained with $W_S = 256$ significantly outperforms the results obtained with $W_S = 128$. In particular, the accuracy obtained for specific devices (devices 3, 4, 8, and 9) is 100%, while it can be noted that the first two devices (devices 1 and 2) are more difficult to distinguish. One possible reason for this is that the first two devices have very similar hardware intrinsic features in comparison to the other devices and the classification algorithm has some challenges (i.e., the classification of many samples is wrongly predicted) to distinguish them.



(**a**) Original signal. (**b**) Residual of VMD with N_M = 4.

Figure 8. Confusion matrices obtained using the reduced spectrogram calculated with $W_S = 128$ on the original signal and the residual. The shades of gray indicate a different accuracy as expressed by the percentage number within the cell.



(**a**) Original signal. (**b**) Residual of VMD with N_M = 3

Figure 9. Confusion matrices obtained using the reduced spectrogram calculated with $W_S = 256$ on the original signal and the residual. The shades of gray indicate a different accuracy as expressed by the percentage number within the cell.

5. Conclusions and Future Developments

This study has proposed a novel approach for RFF, where the VMD is used to highlight the specific differences among the wireless devices (i.e., the fingerprints). The method is based on the extraction of the residuals from the signal generated from the LoRa devices

after the application of VMD with a different number of modes. The method is applied to a public data set of 10 LoRa devices, where the preamble is extracted. On the residual signal (and the original signal for comparison), a spectrogram is applied to create images to be given in input to a CNN with a multi-head attention layer. The potential issue with the use of spectrogram is the large size of the image to be fed to the CNN, which may increase the overall classification time. While it is assumed that the CNN will be able to extract the most discriminating features in the images, this proposal uses a dimensionality reduction method, which combines ML and DL. Specifically, the LBP image processing features are used to single out portions of the spectrogram that have the most discriminating power. Therefore, it is a hybrid approach as it combines the computing efficiency of 'shallow' ML algorithms with the classification performance of DL. The results show that the proposed approach is able to outperform the baseline approach using the spectrogram on the original signal or even the spectrogram on the residual signal both in terms of classification accuracy and computing time.

The strongest limitation of the proposed approach is the computational time of the application of the VMD, which can be significant, especially for long time series as in this data set. In addition, the optimal number of modes is not known a priori and it must be determined empirically, even if the range of values is limited (from 2 to 4).

Future developments may use other image processing features beyond the LBP or other DL architectures. In addition, more recent mode decomposition algorithms could be used as VMD itself has evolved to more sophisticated algorithms from the initial proposal in [24].

Author Contributions: Conceptualization, G.B.; methodology, G.B.; resources, G.B. and F.B.; writing—original draft preparation, G.B.; writing—review and editing, G.B. and F.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Public data provided by the authors of [6].

Conflicts of Interest: Author Fausto Bonavitacola was employed by the company Piksel. The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CFO	Carrier Frequency Offset
CNN	Convolutional Neural Network
DT	Decision Tree
DL	Deep Learning
DSRC	Dedicated Short Range Communications
FFT	Fast Fourier Transform
FSCM	Frequency Shift Chirp Modulation
HOG	Histogram of Oriented Gradient
IMF	Intrinsic Mode Functions
KNN	K-Nearest Neighbor
LBP	Linear Binary Pattern
LoRa	Long Range
LSVM	Linear Support Vector Machine
ML	Machine Learning
RAM	Random Access Memory
RFF	Radio Frequency Fingerprinting
RMSProp	Root mean square propagation
SGDM	Stochastic Gradient Descent with Momentum
SVM	Support Vector Machine
VMD	Variational Mode Decomposition

References

1. Abbas, S.; Abu Talib, M.; Nasir, Q.; Idhis, S.; Alaboudi, M.; Mohamed, A. Radio frequency fingerprinting techniques for device identification: A survey. *Int. J. Inf. Secur.* **2024**, *23*, 1389–1427. [CrossRef]
2. Zhang, J.; Shen, G.; Saad, W.; Chowdhury, K. Radio frequency fingerprint identification for device authentication in the internet of things. *IEEE Commun. Mag.* **2023**, *61*, 110–115. [CrossRef]
3. Soltanieh, N.; Norouzi, Y.; Yang, Y.; Karmakar, N.C. A review of radio frequency fingerprinting techniques. *IEEE J. Radio Freq. Identif.* **2020**, *4*, 222–233. [CrossRef]
4. Zhang, J.; Woods, R.; Sandell, M.; Valkama, M.; Marshall, A.; Cavallaro, J. Radio frequency fingerprint identification for narrowband systems, modelling and classification. *IEEE Trans. Inf. Forensics Secur.* **2021**, *16*, 3974–3987. [CrossRef]
5. Satija, U.; Trivedi, N.; Biswal, G.; Ramkumar, B. Specific emitter identification based on variational mode decomposition and spectral features in single hop and relaying scenarios. *IEEE Trans. Inf. Forensics Secur.* **2018**, *14*, 581–591. [CrossRef]
6. Shen, G.; Zhang, J.; Marshall, A.; Cavallaro, J.R. Towards scalable and channel-robust radio frequency fingerprint identification for LoRa. *IEEE Trans. Inf. Forensics Secur.* **2022**, *17*, 774–787. [CrossRef]
7. Xu, Q.; Zheng, R.; Saad, W.; Han, Z. Device Fingerprinting in Wireless Networks: Challenges and Opportunities. *IEEE Commun. Surv. Tutorials* **2016**, *18*, 94–104. [CrossRef]
8. Reising, D.R.; Temple, M.A.; Jackson, J.A. Authorized and rogue device discrimination using dimensionally reduced RF-DNA fingerprints. *IEEE Trans. Inf. Forensics Secur.* **2015**, *10*, 1180–1192. [CrossRef]
9. Huang, D.; Al-Hourani, A.; Sithamparanathan, K.; Rowe, W.S. Deep Learning Methods for IoT Device Authentication Using Symbols Density Trace Plot. *IEEE Internet Things J.* **2024**, *11*, 18167–18179. [CrossRef]
10. Merchant, K.; Revay, S.; Stantchev, G.; Nousain, B. Deep learning for RF device fingerprinting in cognitive communication networks. *IEEE J. Sel. Top. Signal Process.* **2018**, *12*, 160–167. [CrossRef]
11. Riyaz, S.; Sankhe, K.; Ioannidis, S.; Chowdhury, K. Deep learning convolutional neural networks for radio identification. *IEEE Commun. Mag.* **2018**, *56*, 146–152. [CrossRef]
12. *ETSI EN 302 637-2 V1.3.2 (2014-11)*; Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Part 2: Specification of Cooperative Awareness Basic Service. ETSI: Sophia Antipolis, France, 2011.
13. Baldini, G.; Steri, G.; Giuliani, R.; Dimc, F. Radiometric identification using variational mode decomposition. *Comput. Electr. Eng.* **2019**, *76*, 364–378. [CrossRef]
14. Liang, J.H.; Huang, Z.T.; Li, Z.W. Method of empirical mode decomposition in specific emitter identification. *Wirel. Pers. Commun.* **2017**, *96*, 2447–2461. [CrossRef]
15. Aghnaiya, A.; Ali, A.M.; Kara, A. Variational mode decomposition-based radio frequency fingerprinting of bluetooth devices. *IEEE Access* **2019**, *7*, 144054–144058. [CrossRef]
16. Bisdikian, C. An overview of the Bluetooth wireless technology. *IEEE Commun. Mag.* **2001**, *39*, 86–94. [CrossRef]
17. Aghnaiya, A.; Dalveren, Y.; Kara, A. On the performance of variational mode decomposition-based radio frequency fingerprinting of Bluetooth devices. *Sensors* **2020**, *20*, 1704. [CrossRef] [PubMed]
18. Guo, S.; Xu, Y.; Huang, W.; Liu, B. Specific emitter identification via variational mode decomposition and histogram of oriented gradient. In Proceedings of the 2021 28th International Conference on Telecommunications (ICT), London, UK, 1–3 June 2021; pp. 1–6.
19. Li, Q.; Jin, Y.; Yan, S.; Han, S. Emitter identification based on improved Variational modal decomposition at low SNR. In Proceedings of the 2021 International Conference on Artificial Intelligence and Electromechanical Automation (AIEA), Guangzhou, China, 14–16 May 2021; pp. 152–156.
20. Shen, G.; Zhang, J.; Marshall, A.; Peng, L.; Wang, X. Radio frequency fingerprint identification for LoRa using spectrogram and CNN. In Proceedings of the IEEE INFOCOM 2021-IEEE Conference on Computer Communications, Vancouver, BC, Canada, 10–13 May 2021; pp. 1–10.
21. Hamdaoui, B.; Elmaghbub, A. Deep-learning-based device fingerprinting for increased LoRa-IoT security: Sensitivity to network deployment changes. *IEEE Netw.* **2022**, *36*, 204–210. [CrossRef]
22. Baldini, G.; Amerini, I.; Dimc, F.; Bonavitacola, F. Convolutional neural networks combined with feature selection for radio-frequency fingerprinting. *Comput. Intell.* **2023**, *39*, 734–758. [CrossRef]
23. Shamsipour, G.; Fekri-Ershad, S.; Sharifi, M.; Alaei, A. Improve the efficiency of handcrafted features in image retrieval by adding selected feature generating layers of deep convolutional neural networks. *Signal Image Video Process.* **2024**, *18*, 2607–2620. [CrossRef]
24. Dragomiretskiy, K.; Zosso, D. Variational mode decomposition. *IEEE Trans. Signal Process.* **2013**, *62*, 531–544. [CrossRef]
25. Sornin, N.; Luis, M.; Eirich, T.; Kramp, T.; Hersent, O. *LoRaWAN Specification*; LoRa Alliance, Inc.: San Ramon, CA, USA, 2015.
26. Vangelista, L. Frequency shift chirp modulation: The LoRa modulation. *IEEE Signal Process. Lett.* **2017**, *24*, 1818–1821. [CrossRef]
27. Shen, G.; Zhang, J.; Marshall, A.; Valkama, M.; Cavallaro, J.R. Toward length-versatile and noise-robust radio frequency fingerprint identification. *IEEE Trans. Inf. Forensics Secur.* **2023**, *18*, 2355–2367. [CrossRef]
28. Wei, X.; Yu, X.; Liu, B.; Zhi, L. Convolutional neural networks and local binary patterns for hyperspectral image classification. *Eur. J. Remote Sens.* **2019**, *52*, 448–462. [CrossRef]
29. Tiwari, A.K.; Kanhangad, V.; Pachori, R.B. Histogram refinement for texture descriptor based image retrieval. *Signal Process. Image Commun.* **2017**, *53*, 73–85. [CrossRef]

30. Lu, J.; Yue, J.; Zhu, L.; Li, G. Variational mode decomposition denoising combined with improved Bhattacharyya distance. *Measurement* **2020**, *151*, 107283. [CrossRef]
31. Kiranyaz, S.; Avci, O.; Abdeljaber, O.; Ince, T.; Gabbouj, M.; Inman, D.J. 1D convolutional neural networks and applications: A survey. *Mech. Syst. Signal Process.* **2021**, *151*, 107398. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

electronics

MDPI

Article

Forensic Analysis of File Exfiltrations Using AnyDesk, TeamViewer and Chrome Remote Desktop

Xabiel G. Pañeda, David Melendi *, Víctor Corcoba, Alejandro G. Pañeda, Roberto García and Dan García

Department for Informatics, University of Oviedo, 33203 Xixón, Spain; xabiel@uniovi.es (X.G.P.); corcobavictor@uniovi.es (V.C.); uo197450@uniovi.es (A.G.P.); garciaroberto@uniovi.es (R.G.); garciadan@uniovi.es (D.G.)

* Correspondence: melendi@uniovi.es

Abstract: The use of remote desktop applications has increased greatly in recent years, mainly because of the generalization of telecommuting due to the COVID-19 pandemic. This process has been carried out in a very controlled manner in some companies, but in other organizations it has been introduced in a more anarchic way. The direct use of on-premises company computers and resources from the internet without the necessary protection mechanisms, including VPNs, has increased the risk of data exfiltration. Apart from other types of data exfiltration, there are cases in which employees transfer files using encrypted communications, consciously or unconsciously, producing a leak of information undetected by data loss prevention systems. In this paper we analyse the question of whether a forensic investigation may answer questions about data exfiltrations; questions such as those regarding the when, what and who (or to whom) and the use of application logs and other available tools. The answers to these questions may form the basis of solid digital evidence for legal purposes, though they may only deliver a partial response to said questions. Other complementary sources are necessary to build a complete answer and accurate digital evidence. Nevertheless, we have identified and analysed several use cases that may help to raise an early alarm that can offer warning about certain behaviours in encrypted traffic that may be detected via network monitoring.

Keywords: remote desktop; TeamViewer; AnyDesk; Chrome RD; file exfiltration; evidence; forensic analysis

1. Introduction

In recent years, many company offices have undergone a rapid process of emptying. Telecommuting had already become frequent in certain sectors, but the COVID-19 pandemic caused a tremendous acceleration in the adoption of remote working. Telecommuting has brought many advantages, but it has also created some problems. In fact, it has caused a major paradigm shift in information security [1]. An increasing number of employees of all types of organizations perform tasks outside the office, often using their own devices. Thus, in cases in which its adoption has been carried out in a disorderly fashion, without clear regulation, restrictions or the deployment of access control and monitoring systems, telecommuting has become a major cybersecurity problem. A lack of control can make it easier for outside agents to break into corporate systems and for disloyal employees to exfiltrate sensitive information. While in the world of large companies the adoption of telecommuting has occurred in a very controlled way, in small and medium-sized enterprises it has been carried out in a substantially disordered manner. Previous work has described problems such as the poor competencies required to efficiently use IT solutions, a shortage of equipment, communication and coordination problems, the lack of methodology for organizing distance work, etc. [2,3].

Three applications that have become widely popular under telecommuting scenarios are TeamViewer, AnyDesk and Chrome Remote Desktop. Their adoption is a result of their

ease of use and flexibility. Most internet access services combine both the usage of network address translation (NAT) and a dynamic IP address. This kind of configuration makes the remote access of a computer that is behind a similarly configured router problematic. For instance, special settings need to be used in the router and some dynamic DNS service needs to be used to map a fixed domain name to a variable IP address. To perform these tasks, users need to have some technical knowledge about networking and security. Nevertheless, the chosen applications have been designed to adapt to different network configurations and no changes need to be applied in the face of NAT or routing and security restrictions. They can even set up connections that trespass the usual firewalling controls thanks to the use of HTTPS tunnelling. Beyond their ease of use, all of these products have versions with free licenses.

In the event of an incident of information theft, it is paramount to be able to collect data regarding the activity of the employees. Therefore, in this paper we have analysed the forensic evidence produced when there is the exfiltration of a file from a Windows PC using TeamViewer, AnyDesk or Chrome Remote Desktop. We have chosen TeamViewer and AnyDesk because of their predominant position in the market [4], with TeamViewer having a 55.63% market share in remote support solutions and AnyDesk an 8.99% market share, according to recent reports [5]. Additionally, we have chosen Chrome Remote Desktop because it is a free solution available for most of the operating systems in the market and because it was notably popular during the COVID-19 lockdown [6]. The goal of the paper is twofold. Firstly, in order to be able to initiate a legal process with guarantees, the study aims to determine if the evidence produced by these applications and the underlying systems may answer the following questions: what has been exfiltrated, when has the exfiltration happened and who has exfiltrated the information (or to whom it has been carried out). Being able to respond to these questions is paramount for the said legal process. Secondly, this research seeks to identify aspects in the behaviour of these applications which may allow security officers and appliances to act proactively and stop, or even avoid, data exfiltrations performed with these systems.

The rest of the paper is organized as follows. Section 2 introduces the existing related research. Section 3 presents the base scenarios considered and the analysis methodology that was followed. Section 4 describes our study and the main outcomes of the research. Finally, Section 5 provides the conclusions and future work.

2. Related Work

Telecommuting is not a new concept. However, the COVID 19 pandemic has accelerated its adoption in multiple sectors, including those that are not directly linked with technology [7]. It is now common for companies to use software to access computers remotely, share files or hold meetings. Nevertheless, this software may also be used to illegally extract data and information from companies [8].

There are very few forensic research works analysing remote desktop solutions. Hornyák [9] describes how one may obtain traces of attempts at remote access in Windows using the RDP protocol. Manson [10] analyses the artifacts obtained using TeamViewer and Windows. The author states that it is possible to determine if a machine has used a remote desktop service or has been accessed remotely, if files have been transferred or if that computer has been rebooted remotely. However, this analysis was carried out with old software versions and the study is not focused on file exfiltration. Other common remote desktop applications, such as AnyDesk or Chrome Remote Desktop, were not tested. Kerai and Vekariya [11] conducted a study on evidence of the usage of several remote desktop and teleconferencing applications in Windows. The applications analysed were RealVNC, TightVNC, Cisco WebEx, GoToMeeting and LogMeIn. The authors examined the entries in the Windows registry and the specific log files created by these applications. They found that it is possible to obtain the clients IP addresses and information about transferred files. However, these programs have suffered many changes in recent years and the study is currently outdated. Computers may also be accessed remotely using mobile devices such

as tablets or smartphones. Kerai [12] analysed whether it is possible to obtain evidence of these services on Windows mobile phones. The author found that many of the applications used in this environment did not encrypt connection passwords. Furthermore, it was possible to obtain details such as domain names, IP addresses, transport ports and log-in and log-out times. However, to the best of our knowledge, there are no similar works on current mobile operating systems such as IOS or Android.

Other studies have focused on analysing whether, using side-channel information such as packet lengths and packet arrival times, it is possible to identify the activity being performed by a user accessing a computer via a remote desktop. Altschaffel et al. [13] proposed a method to identify actions made by the user, based on statistical pattern recognition. In their work they used the TeamViewer application. The authors were able to detect file transfers, video conferencing, audio conferencing, text chats and remote sessions. Although the results are interesting, the program version is very old (from 2013) and has undergone many changes. Similarly, Jiang et al. [14] have analysed the possibility of distinguishing between five common activities using machine learning and side-channel information provided by remote desktop software. They analysed AnyDesk, ConnectWise, MicroRDS, RealVNC, TeamViewer and Zoho Assist. The analysed activities included editing documents, reading documents, web browsing, watching videos and installing software. In this study, it was observed that traffic encryption mechanisms are not sufficient to totally hide the activity of the users and that it is possible to identify what they do with great accuracy.

In addition to remote desktop applications, there are other applications that permit file transfers, even though their main use is for meetings or collaborative work. These applications include products such as Skype, Zoom, Slack and Microsoft Teams. Yang et al. [15] analysed the release of Skype available in the official Windows app store. The researchers found valuable information, such as the names of the files transferred and conversation logs. They also observed that traces of these activities remained after the program was uninstalled. Nicoletti and Bernaschi [16] focused their research on the business version of Skype. They analysed the protocols used by the application, its architecture and the information available in Windows logs and the event viewer.

Microsoft Teams is another software that is widely used by companies and educational institutions for telecommuting and e-learning. Paligu and Varol [17] examine the evidence obtained from the IndexedDB storage of Microsoft Teams. In this research, the authors used time-frame analyses and managed to obtain messages from chats, voice mail and several Teams extensions. Additionally, Khalid et al. [18] performed a forensic analysis of the memory, disk and network traffic produced by Microsoft Teams. The authors were able to extract critical information such as email addresses, exchanged messages, deleted messages and transferred files. Google has a comparable tool named Google Meet. Iqbal et al. [19] carried out an exhaustive memory forensic analysis to determine the evidence that Google Meet leaves after it has been used. The authors perform experiments with different browsers (Chrome, Firefox and Mozilla) and different amounts of RAM memory. In their study, they obtained traces of information about the meetings, such as the email addresses and profiles of the participants.

In the literature, we found few studies that examine the traces that may be obtained from the most common applications used for professional communications and to work remotely. However, many of these studies are outdated, do not provide a detailed explanation of how they have obtained the information presented or are not focused on file exfiltrations. In our study, we have analysed artifacts from application logs, operating systems and network traffic to determine if we can answer questions regarding the when, what and who/to whom.

3. Base Scenarios and Methodology

To perform this study, we have analysed two types of situations: (1) when an employee connects from an on-premises computer to a computer outside the company and moves a

file to that external computer (push scenario) and (2) the situation in which a user outside the company connects to an on-premises computer and extracts a file (pull scenario). Figure 1 shows both situations.

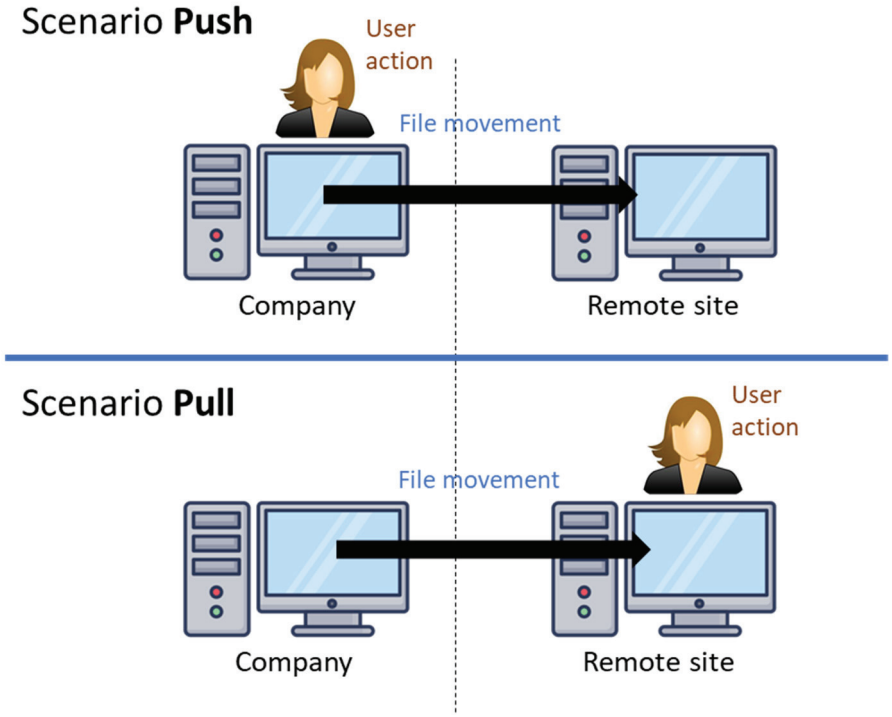


Figure 1. Scenarios considered.

We have created the scenarios using two 64-bit Windows 10 virtual machines. In these machines we have installed AnyDesk version 7.1.6, TeamViewer version 15.31.5 and Chrome Remote Desktop version 2.1. To capture network traffic, we have installed the sniffer WireShark. Additionally, we have created a folder where some image files have been stored. We have activated the Windows Object audit service on the folder to have maximum control over the file actions provided by the Windows operating system. In the virtualization environment, each machine has been connected to a different virtual network. Furthermore, a router allows the communication between both machines and the communication of those machines with the internet by using a static NAT configuration.

To perform the experiments, we have followed a simple methodology. We activated the monitoring tools, performed several exfiltration processes and revised the evidence generated by the different monitoring systems and tools. The exfiltration consists in the transference of three pictures of different sizes to a remote computer: "lapiquera.jpg" (92 KB), "lateyera.jpg" (4.5 MB) and "xixun.jpg" (5 MB). Additionally, to better analyse the traffic generated by the applications, we also performed experiments with large files of sizes greater than 180 MB. Finally, we compared the results with a baseline in which no file transfers were produced.

4. Analyses Performed

4.1. TeamViewer

Distributed and implemented by TeamViewer AG, TeamViewer is an application that provides remote functionalities to computers running a host application. Once a connection has been established, the application provides remote control, file transfer and software deployment capabilities, among others. It is easy to use and does not require end users to perform complex tasks such as modifying firewall rules or configure port forwarding on routers. TeamViewer uses primarily TCP and UDP transmissions over port

5938. Alternatively, it can use ports 443 or 80. It secures data transferences using encryption, based on 4096-bit RSA key exchanges and 256-bit AES session encoding.

TeamViewer offers two different possibilities: to install the software in the computer or to use a portable version. These two possibilities produce different log files in different paths, with interesting information about the activity of the users. The files are either in the "Program Files\TeamViewer" or in the "<user>\AppData\Roaming\TeamViewer" folders. In the "TeamViewer15Logfile.log" file (in "Program Files\TeamViewer") TeamViewer registers details about file transferences. However, this file is not generated by the portable version of the application. Instead, we can find similar information in a file with the same name, located in the "<user>\AppData\Roaming\TeamViewer" folder. In both files we can observe interesting events. We have details about the establishment of the connection with the remote site. Additionally, we can observe the IDs of both the local and remote devices and the timestamp (local time) of the connection. These IDs cannot be directly mapped to a specific IP address. A sample of the records registered in these files follows:

```
2023/01/09 12:18:01.584672 5040 G2              Client connection to
1308133946, server version is 15.37.3, OS = 19 Participant =
[1308133946, 1296448264]
```

File transfers are also detailed in the log file in the on-premises computer. When the connection corresponds to a push scenario, apart from the name of the transferred file, we also have access to the origin and destination paths and to the volume of bytes transmitted (the size of the file), as shown in the following record:

```
Upload from ‘‘C:\fileorigin\lapiquera.jpg’’ to
‘‘C:\Users\smiot\Documents\lapiquera.jpg’’ (92.36 kB)
```

On the other hand, when the connection corresponds to a pull scenario, TeamViewer registers the file transfer with the following line:

```
Send file C:\fileorigin\lapiquera.jpg
```

Moreover, in this scenario the application produces an entry in the "connections_incoming.txt" log file, with the ID of the remote PC and the connection timestamp (GMT). An example of one of these entries follows:

```
1308133946 DESKTOP-HQDP856 09-01-2023 12:43:42 09-01-2023 12:45:31
smiot RemoteControl (e9bef4d9-8fcf-4c78-a956-fd29e8c0856a)
```

However, TeamViewer allows users to configure what is registered in log files, as shown in Figure 2 (in the main menu under "Options → Advanced"). Thus, all of the details may be reduced depending on the configuration available for each user.

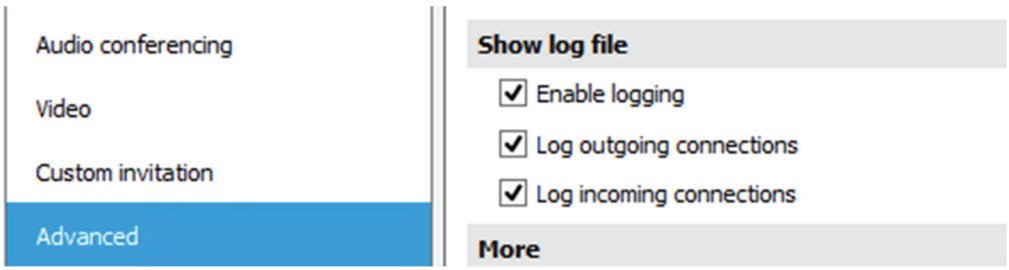


Figure 2. Log options in TeamViewer.

The Windows operating system is also able to register relevant information about the usage of the application. With the default audit, the firewall produces several events (types 5154, 5156 and 5158) that report permitted connections with remote devices. The object audit may complement this information with specific event details (types 4656, 4658 and 4663). This set of events will provide the name of the exfiltrated files.

Network monitoring can also provide interesting information. For instance, although Jiang et al. [14] do not analyse file transfers performed with TeamViewer, the authors have already stated that the traffic encryption mechanisms used by this application cannot completely hide the activity of the users. Remote sessions are usually established with end-to-end full-duplex connections using a proprietary protocol implemented over UDP and variable port numbers. These end-to-end connections are encrypted using 256-bit AES encoding. We have reverse engineered this protocol and have detected that there are different types of messages, each of them with a different type code sent in the 13th byte of the payload of UDP datagrams. Messages containing data have hexadecimal type codes 0x70 and 0x6b. Additionally, they carry a 4-byte sequence number in little endian format in bytes 1 to 4 of the UDP payload, starting in 1. For each direction of the communication, the protocol uses a sliding window controlled with frequent acknowledgement messages identified by type code 0x6f. These acknowledgements carry the number of the data message acknowledged in bytes 5 to 8 of the UDP payload. The frequency of the acknowledgments and the number of acknowledged messages depends on the volume of information delivered in the opposite direction, as shown in Figure 3.

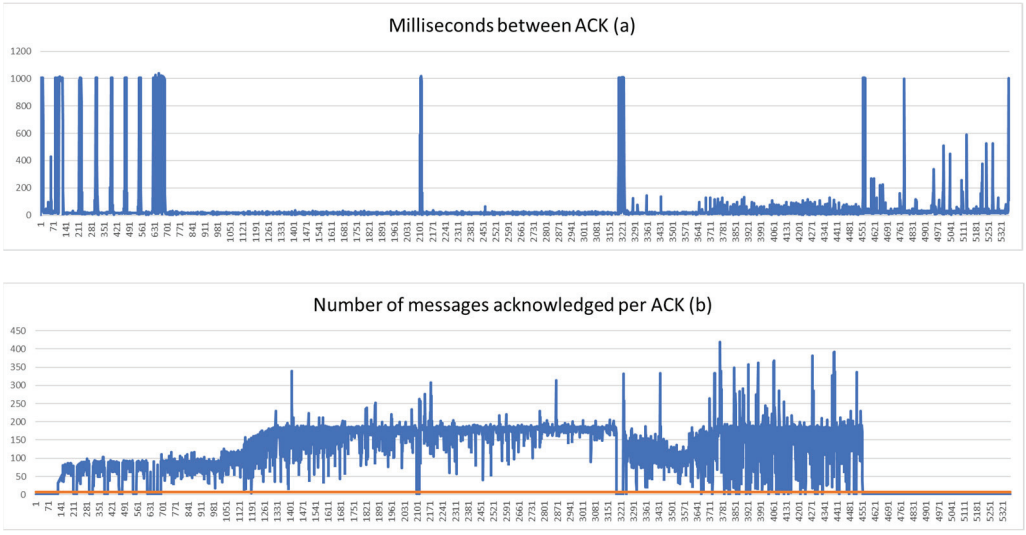


Figure 3. Acknowledgements sent from the remote computer in a push scenario. Milliseconds between acknowledgements (**a**) and number of messages acknowledged in each acknowledgement message (**b**).

Figure 3a shows the frequency of the acknowledgements sent by the remote computer in a push scenario. Low values correspond with acknowledgements sent during a file exfiltration whereas peak (high) values correspond with periods in which no file transfers are being carried out. Moreover, Figure 3b shows the number of messages acknowledged by each acknowledgement message. Values over the orange line correspond to file exfiltrations. We have observed that a network device inspecting this traffic may detect file exfiltrations and the direction they are occurring, even if the information is delivered using encrypted communications.

In a push scenario, the typical traffic of the application produces acknowledgements in the opposite direction of a potential exfiltration every 176 ms on average. Nevertheless, these periods are reduced to approximately 17 ms on average when a file exfiltration is occurring. Additionally, there are differences in the number of messages being acknowledged in those acknowledgements. Acknowledgements of ordinary traffic will be equal to or

lower than 6, as shown with the orange line in Figure 3b. When a file is being transferred, these acknowledgements are greater than 6. On the other hand, in a pull scenario the behaviour of acknowledgments changes because file transfers are carried out in the same direction as the rest of the multimedia elements (e.g., screen information). In this scenario, file exfiltrations can be distinguished from other session activities because the average time between acknowledgements in the opposite direction will decrease. Normal traffic produces acknowledgements every 45 ms on average, but when a file is being transferred, these periods decrease to approximately 17 ms. Additionally, there are differences in the amount of information being acknowledged. Acknowledgements of standard traffic are equal to or lower than 40 and acknowledgements of traffic when there is a file transference are greater than 40.

4.2. AnyDesk

AnyDesk is a remote desktop application implemented and distributed by AnyDesk Software GmbH. It provides a remote access service to personal computers running its host application. The functionalities of this product include remote access, file transfer and remote software installation. Just like its competitors, AnyDesk can run in complex networking scenarios without the need for specific routing rules or firewall changes. AnyDesk uses TCP ports 80, 443 and 6568 to establish connections and TLS 1.2 to secure communications.

AnyDesk has three operating modes. Two of these modes allow users to run the application without installation. The first mode is the portable version and does not require administrative privileges to run. The second mode is called "with elevation" and allows users to interact with Windows user account control (UAC). The third option is to install the program, which permits the application to remain in the background even after all of the program windows have been closed.

From the forensics point of view, AnyDesk has two interesting files in the folder "<user>\AppData\Roaming\AnyDesk", a log file named "ad.trace" and the user configuration file "user.conf". Additionally, in the "<user>\ProgramData\AnyDesk" folder there are another two interesting files named "connection_trace.txt" and "ad_svc.trace".

The "user.conf" file has important information related to the remote PC, including the ID of the remote device, the alias defined by the local user for that remote device and the name of that device if it has been registered by the remote user. Figure 4 shows all of these elements identified in the configuration item.

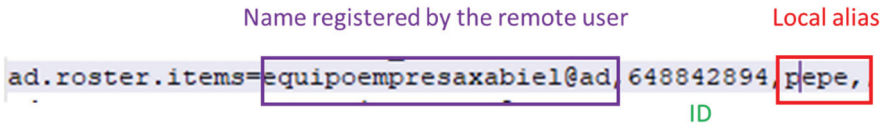


Figure 4. Details of the remote device stored in the "user.conf" file.

The default paths for both sides of the session are also registered in the "user.conf" file, in the "local_browser_start_path" and the "remote_browser_start_path" variables, as shown on the following sample:

```
ad.session.local_browser_start_path=908158965:C*\\origen
ad.session.local_file_sort_order=908158965:33
ad.session.remote_browser_start_path=908158965:d*\\
ad.session.remote_file_sort_order=908158965:33
```

The "ad.trace" file registers the different events that are produced during the sessions. When a remote computer connects with the local computer (pull scenario) the following entry is stored in the file:

```
app.backend_session - Incoming session request:  xabiel (908158965)
```

This log entry includes the ID of the remote PC, in brackets, and the username used to start the connection. In cases when the remote PC has a registered name, the log entry includes this name instead of the ID of the computer as shown, in brackets, in the following sample:

```
Incoming session request:  Alumno (equipoempresaxabiel@ad)
```

On the other hand, when the local computer connects with a remote PC (push scenario), the log entry in the "ad.trace" file includes the ID of the remote PC, as shown in the following sample:

```
win_app.frontend - Sending a connection request for address 908158965.
```

Additionally, both the operating system of the remote device and the version of AnyDesk are registered in the "ad.trace" file, as shown in the following sample:

```
app.frontend_session - Remote OS: Windows
app.backend_session - Remote version:  7.0.14
```

A file transfer would produce several log lines in the "ad.trace" file in the on-premises computer. In one of the lines, we can observe the name of the origin folder in quotation marks. However, the name of the transferred file does not appear. A sample follows:

```
app.prepare_task - Preparing files in 'C:\origen'.
app.local_file_transfer - Preparation of 1 files completed (io_ok).
```

Moreover, the "connection_trace.txt" file registers incoming connections. These log file entries provide the same information in "ad.trace" but in a different format, time and date of the incoming connection and ID or registered names of the involved devices, as shown in the following sample:

```
Incoming 2022-11-15, 08:49 User equipoempresaxabiel@ad 648842894
```

Furthermore, the "ad_svc.trace" file registers incoming connections. However, apart from the device ID, this log file includes the IP address of the remote device. This is very relevant because the destination of the transmission can be discovered. A sample follows:

```
anynet.any_socket - Connecting to 908158965.
anynet.any_socket - Retrieving client information.
anynet.any_socket - Client-ID: 908,158,965 (FPR: 95c686d63e15).
anynet.any_socket - Logged in from 156.35.171.129:63370 on relay
cc24d14a.
net.connection_mgr - making a new connection to client
95c686d63e15355eb7cb
```

Additionally, sessions may have been recorded and the videos would be stored in "<user>\Videos\AnyDesk".

The Windows operating system also registers relevant information. With the default audit configuration, the Windows firewall produces several events (event types 5154 and 5158) indicating some allowed connections with remote devices. The object audit may complement this information with a specific event (event types 4656, 4658 and 4663). This set of events provides the names of exfiltrated files.

Network monitoring may also produce interesting information. AnyDesk transfers all of the multimedia elements over a single end-to-end bidirectional TCP connection using

TLS. In a push scenario there is little information to send from the local computer to the remote device, so most TCP segments will be empty when they are sent and they will only be used to acknowledge the data being received in the opposite direction. Therefore, in this scenario we can clearly identify a file exfiltration because of an increase in the number of TCP segments with payload being sent by the local computer and by the behaviour of the acknowledgements sent back from the remote computer. A sequence of TCP segments with payload sizes equal to the maximum segment size (MSS) in that network segment indicates a file exfiltration. Moreover, a file exfiltration, or the usage of the file transfer tool, will make the remote computer send TCP acknowledgement segments without payload continuously, as piggybacking will not be sufficient to acknowledge all of the data segments being received. On the other hand, in a pull scenario, the delivery of the exfiltrated files takes place in the same direction as the delivery of all of the multimedia elements of the session. Thus, it is more difficult to identify an exfiltration. The delivery of the multimedia elements of the session generates traffic with notably burst-like behaviour. Under a regular use of the system, a sliding window of 1 s used to calculate average times between data segments will obtain values ranging from less than 1 ms to more than one hundred milliseconds. If an exfiltration is performed, this method to calculate average times between data segments will produce sustained average times between data segments below 1 ms. If the connection is monitored in the opposite direction, the same burst-like behaviour is observed for TCP acknowledgements. When an exfiltration is performed, a sliding window of 1 s used to calculate average times between acknowledgement segments will obtain sustained values below 1 ms, whereas under a regular use of the system this value ranges between 1 and more than 100 ms.

In summary, we can ascertain whether a connection with a remote device has taken place and when. Additionally, we can discover if a file transmission has been produced, when it has been produced, the number of transmitted files, the source folder and the ID and IP address of the destination host. If the Windows object audit is activated, we can also determine the names of the exfiltrated files.

4.3. Chrome Remote Desktop

Chrome Remote Desktop (Chrome RD) is an application developed by Google LLC (Menlo Park, CA, USA) that employs a proprietary protocol called Chromoting. This application is compatible with Linux, Windows, MacOS and ChromeOS. Additionally, there is a software client available for mobile devices running Android or IOS. Chrome RD is based on QUIC and the webRTC protocol stack and uses the VP8 and VP9 codecs.

Chrome RD allows two types of connections. The first type is named "Remote Access". It is a pre-authorized and permanent connection that requires both computers to use the same Google account. The second type is named "Remote Support". This type has been designed for occasional connections in which one person provides support to a remote user. In this connection mode, the remote user must generate an access code that should be given to the support person to allow the connection between their devices. Connections are blocked automatically after a certain period, forcing remote users to confirm that they would like to continue sharing their desktop by pressing a "Continue Sharing" button. Our analysis was carried out using Chrome RD version 2.1 for Windows 10, which is currently the latest stable version of the system.

If the "Remote Access" option is used, it is not necessary to install the Chrome RD extension in the on-premises computer in a push scenario. However, traces of the usage of this software may be found in the history of the browser, as shown in Figure 5. The browser records the time and day of the connection, its duration and a counter with the number of times the remote device was accessed. This counter is updated every time a new connection is performed. In addition, an ID is assigned to every new remote computer, to be later used to report future connections. Regarding the transference of files, the Windows object audit permits one to determine the directory and name of exfiltrated files, as shown in Figure 6. There are three events registered for this type of action (event types 4656, 4658 and 4663).

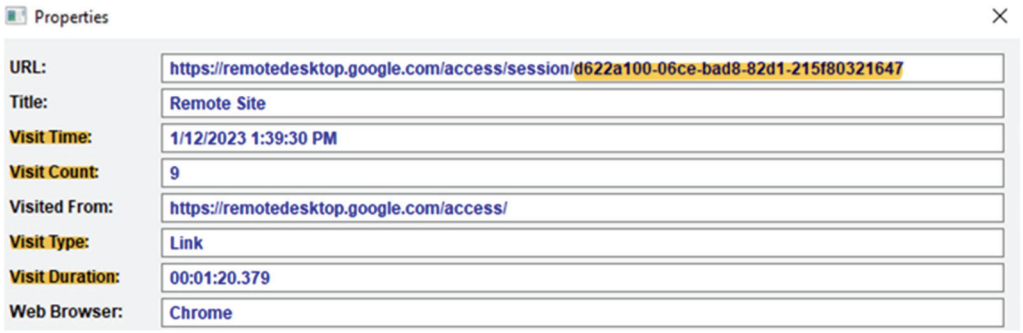


Figure 5. Relevant information found in the browsing history of an on-premises computer in a push scenario, when the "Remote Access" connection mode is used.

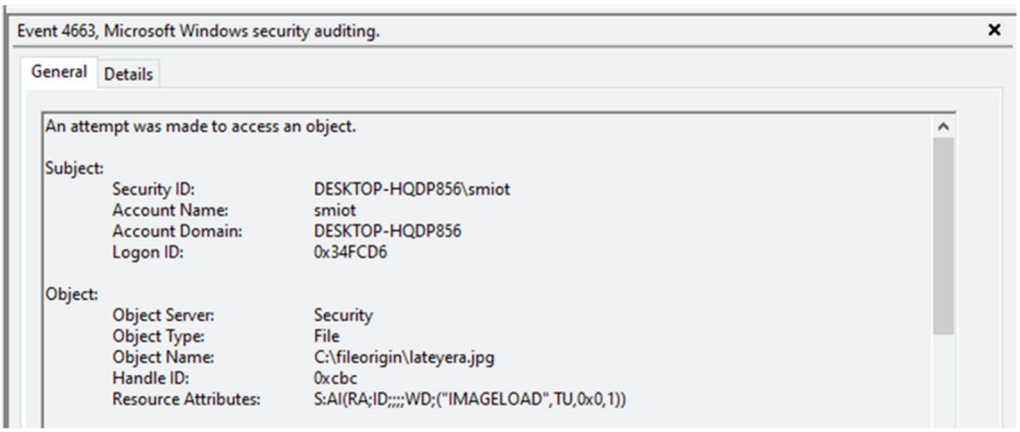


Figure 6. Directory and file names of exfiltrations registered by the Windows object audit.

In a pull scenario, the application running in "Remote Access" mode also produces several artifacts. Two Windows informative events (event types 1 and 4) are generated by Chrome RD when a client connects. Analysing these events, we can obtain the connection date and time, together with the Google email address used for the remote access as shown in Figure 7. The Windows firewall also produces an event (event type 5158). Furthermore, the Windows object audit allows us to determine the name of exfiltrated files (event types 4656, 4658 and 4663).

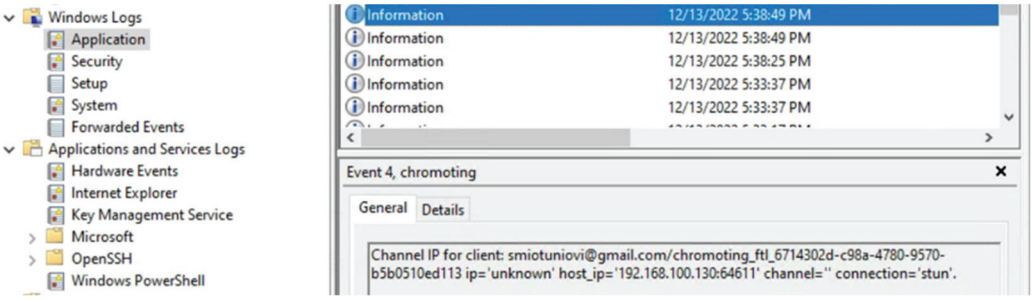


Figure 7. Windows informative event log entries generated by Chrome RD.

Once an exfiltration has been performed, if the perpetrator uninstalls the Chrome RD extension from the on-premises computer to remove evidence, there are several variables related to this software which remain in the Windows registry. Moreover, several artifacts generated by the usage of Chrome RD remain in the cache of Google Chrome, although they are only generated when the program is executed.

If the "Remote Support" connection mode is used, it is only possible to transfer files in the push scenario. Chrome RD allows users to copy files to the machine being assisted, but

it does not allow users to extract files from that device. Evidence from the remote assistance session can be found in the history of the browser, as shown in Figure 8. The information available is the same as that which may be found when the system is used in the "Remote Access" mode.

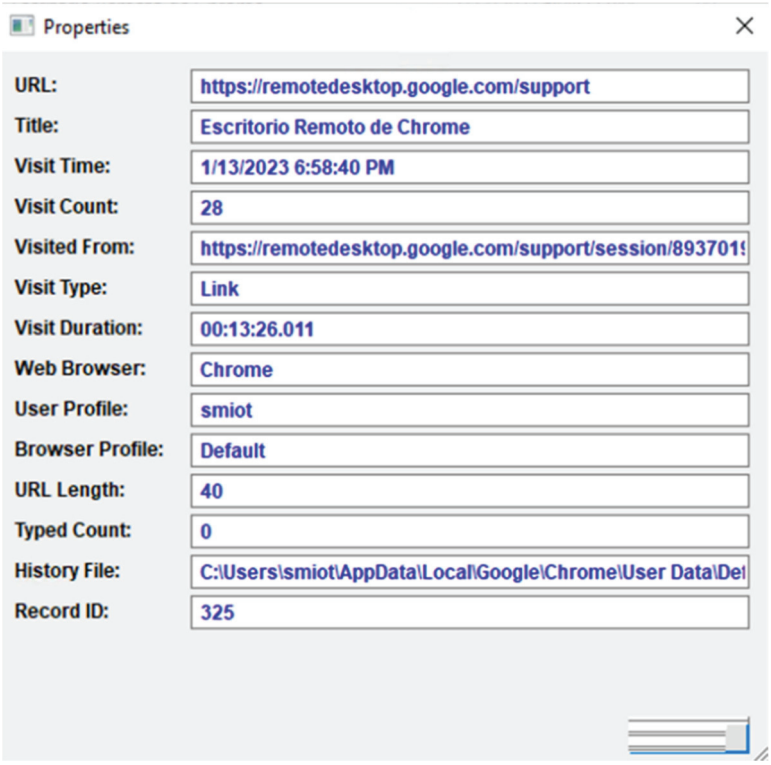


Figure 8. Relevant information found in the browsing history of an on-premises computer in a push scenario, when the "Remote Support" connection mode is used.

Network monitoring may also produce interesting information to detect when a remote session starts and ends using both modes of operation of Chrome RD. During the remote desktop session, QUIC messages are exchanged between the computer and Google servers. However, the exchange of data produced between the involved computers is performed using webRTC data channels. This implies that multimedia data are multiplexed and sent encrypted using DTLS 1.2 over UDP, including the exfiltrated files. As with the rest of the applications, in a push scenario most data are delivered from the remote device to the local computer, so in the opposite direction the majority of UDP datagrams carry control information only, with payloads smaller than 400 bytes and rates under 50 datagrams per second. However, when a file exfiltration occurs, there is a considerable increase in the delivery of UDP datagrams containing DTLS messages. In this scenario, rates over 50 UDP datagrams per second with a DTLS payload from the local computer to the remote computer correspond with file exfiltration operations. Moreover, we also see an increase in the number of DTLS + UDP datagrams being sent in the opposite direction. In a pull scenario, exfiltrated files are sent in the same direction as the rest of the multimedia elements of the session. Though exfiltrations are more difficult to detect in these scenarios, we have seen that rates over 70 UDP datagrams per second with a DTLS payload from the local computer to the remote computer correspond with file exfiltrations.

In summary, we can determine when a connection has been made to a remote computer and its duration by using either the "Remote Access" or the "Remote Support" modes. In addition, it is possible to obtain the Google email account associated with the established session in the pull scenario when the user runs the system in the "Remote Access" mode. If the Windows object audit is activated, we can also determine the name of the files transmitted.

5. Conclusions and Future Work

In this paper we have analysed the behaviour of three applications used for remote desktop operations. The main goal of this research is to verify whether it is possible to detect file exfiltrations using these applications or not. Additionally, we sought to analyse the evidence which may be found in end-systems and the behaviour of their communications to be able to answer a set of interesting questions from a forensics point of view. The answers to these questions may form the basis of solid digital evidence for legal purposes. Nevertheless, our findings only deliver a partial response to said questions, thus, other complementary sources are necessary in order to build a complete answer and strong digital evidence. This includes inventory information, details of connections performed from a certain location or authentication logs in corporate services.

The study has been focused on three popular solutions: TeamViewer, AnyDesk and Chrome Remote Desktop. Moreover, we have studied two types of situations: a push scenario, in which users connect from an on-premises computer to a computer outside the company to exfiltrate files and a pull scenario, in which the user is connecting from a remote location to a on-premises computer. Table 1 shows a summary of the relevant evidence identified in the study.

Table 1. Summary of the relevant evidence identified in the study.

	TeamViewer	**AnyDesk**	**Chrome RD**
When	Remote session initiation time, file transfer time	Remote session initiation time, file transfer time	Remote session initiation time and end time
Who/to whom	ID of remote device	IP address and ID of remote device	Google account
What	Number of files, names, sizes (push scenario), source and destination folders	Number of files, source folder and destination folder	
Additional information	Connection scenario (push, pull)	Connection scenario (push and pull), remote OS and application version, remote device name	Google email, only in pull scenario using remote access
Windows audit	File name	File name	File name
network monitoring	Exfiltration produced, destination IP	Exfiltration produced, destination IP	Exfiltration produced, destination IP

The use of TeamViewer or AnyDesk to exfiltrate files has similarities and differences from the forensics point of view. Both technologies register connections and file transmissions. However, AnyDesk does not include the names of exfiltrated files in log entries, only the source folder. On the other hand, TeamViewer includes not only the name, but also the origin and destination paths. Moreover, AnyDesk registers in one of its log files the IP addresses of remote computers, thus, new investigation lines may be opened thanks to this information (provider, geolocation, etc.).

Chrome Remote Desktop has totally different behaviour. This application is based on a browser plugin and does not generate log files. It has two modes of operation: "Remote Access" and "Remote Support". The latter is more limited and only allows users to transfer files in a push scenario. The activity details need to be extracted from the Chrome history. However, these details do not include relevant information from an investigation point of view. Furthermore, the browser history may be deleted without admin privileges, so evidence may be easily cleaned.

The Windows event log in its default configuration produces limited information for the three analysed applications. Connection events indicate that the application is in use and that connections with master servers have been established. The object audit provides more information about the files transmitted. This is very positive for an investigation, but

it is important to consider that the use of this type of audit is extremely resource consuming and, thus, is rarely activated.

From the point of view of the network traffic generated by the applications, all of them show specific behaviours which may be used to identify an exfiltration, even when encryption is in use. The files being transferred cannot be identified, but a device monitoring the activity in the network may be able to detect file exfiltrations and trigger an alarm. In the case of TeamViewer, a proprietary protocol implemented over UDP is in use. Although contents are sent using encryption, file exfiltrations may be identified by monitoring both the rate at which acknowledgements are being sent and the number of fragments being acknowledged in these messages. In the case of AnyDesk, TLS and TCP are used to deliver multimedia information. File exfiltrations may be detected by monitoring the size of the payload carried in TCP segments and the rate at which segments carrying data are being sent. Finally, Chrome Remote Desktop is based on the webRTC protocol stack. File transferences are performed using DTLS. An exfiltration may be detected by monitoring the rate at which UDP datagrams carrying a DTLS payload are being generated.

If the user behind the remote device would like to be identified, this needs the cooperation of manufacturers in the case of TeamViewer and AnyDesk. Nevertheless, these companies are reluctant to provide this type of information without a warrant, because of personal data protection policies in force in several regions of the world. Thus, it is necessary to produce enough evidence to support a request for this type of information. In the case of Chrome RD, it is possible to obtain the Google account associated with a file exfiltration, when a user from outside the organization connects to an on-premises computer using the "Remote Access" connection mode.

Our study also highlights that Chrome RD is the most dangerous remote access application as the evidence produced for a forensic investigation is very limited and may be easily erased. AnyDesk and TeamViewer produce much more information, but some investigation questions cannot be answered if the information registered in their log files is not combined with other log mechanisms of the operating system in use. Nevertheless, in AnyDesk we do not have the name of the files transferred and in TeamViewer we do not have the destination IP address. Although communications are encrypted for all of the applications, network traffic analysis may be useful as an early warning mechanism, because exfiltrations can be detected. Finally, we must consider that Chrome RD is not a complete application and may be easily blocked. It requires logging into a Google account to function, something that AnyDesk or TeamViewer do not require.

Though our study has produced interesting results and the analyses performed were focused on some of the main solutions that are currently available, technology is constantly evolving. Current applications continuously change and new versions and upgrades are released frequently. Furthermore, new solutions occasionally appear that may use different protocols or communication techniques. Thus, the manner in which exfiltrations may be detected and the registered evidence also change continuously, so this type of study must be constantly repeated.

Author Contributions: Conceptualization, X.G.P. and D.M.; methodology, R.G.; software, D.G.; validation, A.G.P.; formal analysis, R.G.; investigation, X.G.P., D.M. and V.C.; resources, D.M.; data curation, D.G.; writing—original draft preparation, X.G.P., D.M. and V.C.; writing—review and editing, A.G.P.; supervision, X.G.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Dataset available on request from the authors.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Georgiadou, A.; Mouzakitis, S.; Askounis, D. Working from home during COVID-19 crisis: A cyber security culture assessment survey. *Secur. J.* **2022**, *35*, 486–505. [CrossRef]
2. Strakšienė, G.; Ruginė, H.; Šaltytė-Vaisiauskė, L. Characteristics of Distance Work Organization in SMEs During the COVID-19 Lockdown: Case of Western Lithuania Region. *Entrep. Sustain. Issues* **2021**, *8*, 2010–2225. [CrossRef] [PubMed]
3. Islam, M.A.; Igwe, P.A.; Rahman, M.; Saif, A.N.M. Remote working challenges and solutions: Insights from SMEs in Bangladesh during the COVID-19 pandemic. *Int. J. Qual. Innov.* **2021**, *5*, 119–140. [CrossRef]
4. Fortune Business Insights. "Remote Desktop Software Market". Available online: https://www.fortunebusinessinsights.com/remote-desktop-software-market-104278 (accessed on 26 March 2024).
5. 6Sense. "Remote Support", Technographics. 2024. Available online: https://6sense.com/tech/remote-support (accessed on 26 March 2024).
6. TrustRadius. "Remote Desktop Software Statistics and Trends". Available online: https://solutions.trustradius.com/vendor-blog/remote-desktop-buyer-statistics-and-trends (accessed on 26 March 2024).
7. Haider, M.; Anwar, A.I. The prevalence of telework under COVID-19 in Canada. *Inf. Technol. People* **2023**, *36*, 196–223. [CrossRef]
8. Azhar, M.A.H.B.; Timms, J.; Tilley, B. Forensic Investigations of Google Meet and Microsoft Teams—Two Popular Conferencing Tools in the Pandemic. In *Digital Forensics and Cyber Crime*; ICDF2C 2021. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering; Gladyshev, P., Goel, S., James, J., Eds.; Springer: Cham, Switzerland, 2022; Volume 441.
9. Hornyák, O. Protection against remote desktop attacks. *Prod. Syst. Inf. Eng.* **2022**, *10*, 3. [CrossRef]
10. Manson, J. Remote Desktop Software as a forensic resource. *J. Cyber Secur. Technol.* **2022**, *6*, 1–26. [CrossRef]
11. Kerai, P.; Vekariya, V. An exploration of artefacts of remote desktop applications on Windows. In Proceedings of the 14th Australian Digital Forensics Conference, Edith Cowan University, Perth, Australia, 5–6 December 2016; pp. 42–49.
12. Kerai, P. Tracing VNC And RDP Protocol Artefacts on Windows Mobile and Windows Smartphone for Forensic Purpose. In Proceedings of the 1st International Cyber Resilience Conference, Edith Cowan University, Perth, Australia, 23 August 2010; pp. 58–68.
13. Altschaffel, R.; Clausing, R.; Kraetzer, C.; Hoppe, T.; Kiltz, S.; Dittmann, J. Statistical Pattern Recognition Based Content Analysis on Encrypted Network: Traffic for the TeamViewer Application. In Proceedings of the 2013 7th International Conference on IT Security Incident Management and IT Forensics, Nuremberg, Germany, 12–14 March 2013; pp. 113–121.
14. Jiang, M.; Gou, G.; Shi, J.; Xiong, G. I Know What You Are Doing with Remote Desktop. In Proceedings of the 2019 IEEE 38th International Performance Computing and Communications Conference, London, UK, 29–31 October 2010; pp. 1–7.
15. Yang, T.Y.; Dehghantanha, A.; Choo, K.K.R.; Muda, Z. Windows Instant Messaging App Forensics: Facebook and Skype as Case Studies. *PLoS ONE* **2016**, *11*, e0150300. [CrossRef] [PubMed]
16. Nicoletti, M.; Bernaschi, M. Forensic analysis of Microsoft Skype for Business. *Digit. Investig.* **2019**, *29*, 159–179. [CrossRef]
17. Paligu, F.; Varol, C. Microsoft Teams desktop application forensic investigations utilizing IndexedDB storage. *J. Forensic Sci.* **2022**, *67*, 1513–1533. [CrossRef] [PubMed]
18. Khalid, Z.; Iqbal, F.; Al-Hussaeni, K.; MacDermott, A.; Hussain, M. Forensic Analysis of Microsoft Teams: Investigating Memory, Disk and Network. In *Science and Technologies for Smart Cities*; Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering; Springer: Cham, Switzerland, 2022; Volume 442, pp. 583–601.
19. Iqbal, F.; Khalid, Z.; Marrington, A.; Shah, B.; Hung, P.C. Forensic investigation of Google Meet for memory and browser artifacts. *Forensic Sci. Int. Digit. Investig.* **2022**, *43*, 301448. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

MDPI AG
Grosspeteranlage 5
4052 Basel
Switzerland
Tel.: +41 61 683 77 34

Electronics Editorial Office
E-mail: electronics@mdpi.com
www.mdpi.com/journal/electronics

Disclaimer/Publisher's Note: The title and front matter of this reprint are at the discretion of the Guest Editors. The publisher is not responsible for their content or any associated concerns. The statements, opinions and data contained in all individual articles are solely those of the individual Editors and contributors and not of MDPI. MDPI disclaims responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Academic Open
Access Publishing

mdpi.com

ISBN 978-3-7258-4758-7