

Article

Emulation of Digital Substations Communication for Cyber Security Awareness

Filip Holik ^{1,*} , Sule Yildirim Yayilgan ¹  and Guro Bråten Olsborg ² 

¹ Department of Information Security and Communication Technology, Faculty of Information Technology and Electrical Engineering, Norwegian University of Science and Technology, 2815 Gjøvik, Norway; sule.yildirim@ntnu.no

² Department of Information Security and Communication Technology, Faculty of Information Technology and Electrical Engineering, Norwegian University of Science and Technology, 7491 Trondheim, Norway; guro.b.olsborg@ntnu.no

* Correspondence: filip.holik@ntnu.no

Abstract: Increasing power consumption and reliance on non-predictable renewable power generation is pushing the transition from analog to digital power grid substations forward. Grid digitalization helps to reduce substation complexity and therefore costs, and improves observability and management, but introduces new cyber security issues. To make the digital substations secure, cyber security awareness and efficient personnel training is one of the most important research areas as the power grid is a part of critical infrastructure. In our previous work, we have proposed an approach for analyzing cyber security threats and attacks in digital substations based on a case study from Norway. In this article, we present how we developed a tool for emulation of digital substation communication for cyber security awareness based on experiences from the case study. We present technical details of the tool—called the SGSim—so the community can easily replicate the process or only the selected parts. We also freely provide source code on GitHub and distribution in the form of a virtual machine on request. Finally, we validate the tool performance in several scenarios and evaluate its usability on a survey conducted among a wide range of professionals.

Keywords: cyber security awareness; digital substations; communication emulation; impact simulations; digital twin



Citation: Holik, F.; Yayilgan, S.Y.; Olsborg, G.B. Emulation of Digital Substations Communication for Cyber Security Awareness. *Electronics* **2024**, *13*, 2318. <https://doi.org/10.3390/electronics13122318>

Academic Editors: Kutub Thakur and Muath Obaidat

Received: 16 May 2024

Revised: 9 June 2024

Accepted: 12 June 2024

Published: 13 June 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The transition towards digitalization of electrical substation infrastructure is bringing IT (Information Technology) concepts, practices and tools into the OT (Operational Technology) world, making them part of the smart industry. A similar trend is being seen in other domains of the smart city concept including smart buildings, smart transportation, smart manufacturing, etc. This transition greatly increases observability, flexibility and control while reducing the complexity and costs.

On the other hand, introduction of general IT infrastructure and common communication protocols into the OT area breaks the “security through obscurity” paradigm and introduces new security challenges [1]. This can be seen by the increased number of attacks on critical grid infrastructure [2] which has more than doubled between 2020 and 2022 following the Russian invasion of Ukraine [3], with Russia recognized as the top threat to cyber security by the Office of the Director of National Intelligence [4]. Unfortunately, the financial investment in cyber security in the energy sector is significantly smaller than the budgets provided for transportation, health care, banking and general ICT (Information and Communication Technology) according to the ENISA report [5]. Considering the severity of the damage that cyber attacks are capable of causing on critical infrastructures, there is a clear need from a research perspective to clarify the type, reason and timing of these attacks in order to put preventative measures in place.

This is especially problematic for current substation operators with expertise focusing on the OT networks and not on IT cyber security. As digital substations are part of critical infrastructure, it is crucial to provide easy and effective ways to create cyber security awareness. This would cover the following aspects:

- Continuous training and education—an ability to train professionals in realistic scenarios which can easily be adjusted based on changing needs. The tool should provide a realistic environment, employ a simplistic and user friendly approach and keep cost and maintenance requirements low.
- Threat awareness—practical demonstration of threats and their impact on the system so professionals can observe the process and learn to detect and react to threats.
- Infrastructure configuration verification—the use of real hardware and/or software configuration to verify correct settings before production deployment. This might include verification of ACL (Access Control Lists), firewall rules, etc.
- Big data generation—the ability to generate a large amount of real world traffic data which can be used for machine learning or infrastructure testing.

The best approach to provide the aforementioned aspects is to have a physical clone of a digital substation which could be used for training and testing. This would naturally present a significant financial cost and would require complex maintenance with long downtimes before each session. There is also the danger of damaged equipment from cyber attacks. A more sustainable and efficient approach is to utilize the concept of *Digital Twins*.

1.1. Digital Twins

A digital twin is a virtual representation of a physical system—a replica with real world properties such as values, communication and behavior. Digital twins do not only provide a tool for cyber security awareness, but are also invaluable in regards to performance optimization, monitoring, as well as diagnosis and predictive maintenance. On the other hand, a digital twin is “the ultimate, unachievable goal” [6] as there are always certain simplifications.

1.2. Motivation

Cyber security awareness in energy grid networks is an area which has been identified as one of the most important aspects of grid resiliency [7]. The entire process to create cyber security awareness requires a large degree of investment in terms of human resources, time and money. Nevertheless, this is essential due to the following:

- Real grid infrastructures cannot be used for understanding the nature of the attacks as the damage would be costly.
- There is a need for professionals and researchers to gain awareness on the impact and nature of grid attacks in order to understand the attacker’s tactics and objectives and develop preventative strategies.
- Existing simulation tools may fall short of providing the scope and realism needed to understand possible attacks if they focus on lower-level components in the grid. A focus on higher-level components and the communication and messaging among them provides a better platform to simulate the vulnerabilities and attacks that may occur.
- Collecting data from actual grids and strictly controlled lab environments with real hardware limits data collection while ML algorithms used for predictions and forecasting require large amounts of data. Simulations and emulations can provide the data instead.

1.3. Context

In our previous work [8], we defined an approach for analyzing cyber security threats and attacks in digital substations based on a case study from Norway. This approach consists of the following steps:

1. Threat analysis—a process of identification of assets and risks.

2. Threat modeling—a process of identification of the most vulnerable parts of the system by risk quantification composed from effects and likelihoods.
3. Risk assessment—a process of creating a risk assessment model with threats classified into one of the 3 categories: critical, with clear mitigation and can be ignored.
4. Impact simulations—a process of practical implementation and verification of the identified threats classified as critical in a high-detailed simulation model.

This article presents a tool called Smart Grid Simulator (SGSim) which acts as a platform for performing impact simulations for supporting cyber security awareness. It is a digital model (a simplified digital twin without online connection to the real infrastructure) focusing on realistic communication emulation in smart grid networks. The OT assets of the grid such as circuit breakers, merging units, sensors, etc. are only simulated or abstracted from completely. This simplification allowed us to keep the tool accessible on the average laptop. The tool is based on our work on the same case study from Norway in the span of 3+ years and two major projects: ECODIS [9] and InterSecure [10] in which we cooperated with many industrial and research institution partners.

1.4. Contribution and Organization

The main contribution of this article is to present and describe SGSim—a tool for performing impact simulations which are the final step in the comprehensive approach for analyzing cyber security threats and attacks in digital substations. The article describes implementation details, as well as tool validation and usage experience from industrial and research organization partners.

The main contributions of the article are:

- Presenting a platform for performing impact simulations in smart grid scenarios.
- Demonstrating how various aspects of the smart grid infrastructure can be emulated or simulated in a virtual environment.
- Providing results of a survey conducted with field professionals, and researchers including their assessment of the tool and suggestions for future updates.

The rest of the article is organized as follows: Section 2 describes related work in the areas of smart grid networks, digital twins and cyber security awareness; Section 3 introduces the SGSim and its aspects; Section 4 summarizes use case scenarios for SGSim and describes its validation; Section 5 presents the SGSim survey evaluation and discusses its results. Finally, Section 6 concludes the article.

2. Related Work

To the best of our knowledge, there is no similar tool like the one we are presenting in this work. The area of digital models (digital twins not connected to real systems), which could easily be used by any personnel without requiring special hardware, is very limited. The existing work can be divided into theoretical mathematical models and simulations, and systems integrating hardware elements.

The first category is represented by a mathematical model proposed in [11]. The model focuses on IEC 61850 communication and can perform a quantitative risk assessment and assign priorities and probabilities based on set thresholds. The model was tested on a dataset captured in a 24 h window from a real substation in a PCAP (Packet Capture) format and it included four data anomalies: L-DdoS, H-DdoS (Distributed Denial of Service), network storm and network shielding.

The second category focuses on connecting physical systems with digital twins. Paper [12] describes a virtualized experiential learning platform in substation automation area which includes two cyberphysical attacks simulation—MitM (Man in the Middle) and DoS implemented by integrating RTDS HiL.

A physical testbed of a digital station for testing and evaluating cyber attacks was described in [13]. The testbed is based on a hardware-in-the-loop (HiL) architecture and contains real substation equipment. A similar setup was explored in [14], where authors

combined RTDS and RSCAD simulators [15] with PMUs (Phasor Measurement Unit), NS-3 [16] for network simulation and an application for real time voltage stability monitoring.

These cyber-physical setups provide a wide area for detailed experimentations but with the drawback of high costs and difficult operation and maintenance, mostly due to the use of hardware tools such as OPAL-RT [17]. While our solution abstracts from details of the grid equipment and cannot compete in the area of hardware simulators, it excels in emulation of the communication network (unlike simulations used in the aforementioned works). The emulation provides real world communication and an ability to connect the network to real equipment without incurring additional costs.

3. Smart Grid Simulator (SGSim)

SGSim is a tool for cyber security awareness and threat impact simulations. This section describes its architecture and main properties which illustrate how to emulate specific parts of digital substation communication in an environment suitable for any average laptop or computer.

3.1. SGSim Architecture

SGSim is available in the form of a virtual machine (VM) in an .ova file (for access to the file, please contact the correspondence author of this article). This format can be imported on any platform such as VirtualBox or VMWare and on any operating system (Linux, Windows, macOS). The VM uses Linux Lite 6.0 as the guest operating system. The model topology of smart grid networks is emulated using the Mininet network emulator version 2.3.0d6. IEC communication is realized by scripts from two open-source libraries: libIEC61850 version 1.4.2.1 and lib60870 version 2.2.0 [18]. The source files of the SGSim are also available on Github [19]. The architecture of the tool is shown in Figure 1.

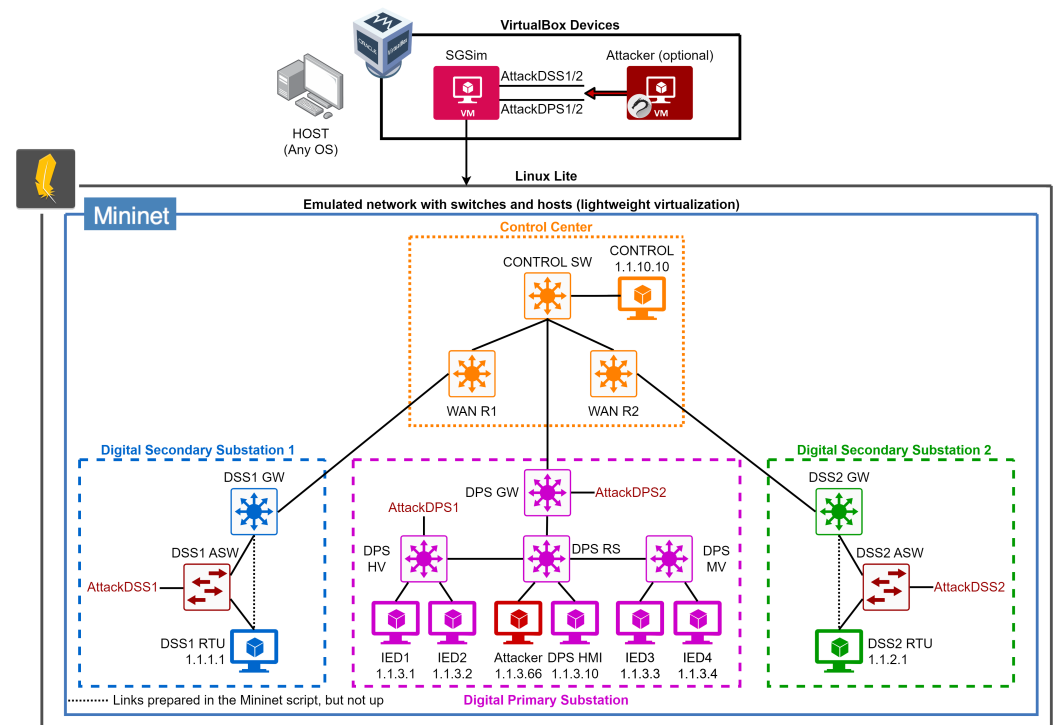


Figure 1. SGSim architecture—an emulated network inside a VM running on any platform.

The architecture is focused on emulation of (digital) communication and it therefore does not contain the grid part. Moreover, it also abstracts from specific network details such as:

- Routers—routers and gateways are replaced by (multilayer) switches and IP (Internet Protocol) routing is not utilized (the entire network topology is in the same network subnet 1.0.0.0/8. The topology however follows realistic traffic patterns—for example, layer 2 traffic in the substation does not leave its boundary (this is achieved by an OpenFlow rule inserted on the DPS GW device).
- VPN (Virtual Private Network)—in a real world scenario, links between substations and control center are encrypted by VPNs. This is omitted in the model.
- IEDs and RTUs—grid equipment is represented by a Linux hosts with corresponding IEC library scripts generating traffic according to predefined values (or values ranges).

The main advantage of this architecture is low resource requirements—the topology can run sufficiently even on an average laptop CPU and with 4 GB assigned RAM. The VM packaging also eliminates the need for a specific operating system or system configuration. Finally, this architecture allows interconnection with real hardware or virtualized devices. This is achieved by creating external links in the model topology (marked in Figure 1 as AttackDSSx and AttackDPSx) which can be easily mapped in the virtualization platform to real (or virtual) interfaces as shown with the Attacker VM. This can then be used for introducing an attacker device or protection system into the network.

3.2. Emulated Topology

SGSim has two prepared smart grid topologies (basic topology and extended topology) which can be started from the desktop and easily modified in corresponding Python scripts. The topologies are composed from the following elements:

- Control center—represents a control room with a device collecting IEC 60870-5-104 (further called IEC104) communication. It contains a simulated SCADA visualization.
- Digital secondary substations (DSS1 and DSS2)—substations responsible for converting medium voltage to low voltage (LV) used in households and small businesses. Their RTUs (Remote Terminal Units) communicate with SCADA via IEC104 messages.
- Digital primary substation (DPS, only in the extended topology)—a substation responsible for converting high voltage (HV) to medium voltage (MV). Substation IEDs (Intelligent Electronic Devices) communicates with GOOSE and SV protocols inside the substation.

The extended topology is shown in Figure 2 with the above-described blocks marked with separate color areas. The only difference in the basic topology is lack of the DPS and therefore the focus is only on the secondary substations, the control center and the IEC104 communication. This can save computing resources if GOOSE and SV communications are not required (as they are more demanding than IEC104 communication). The grid equipment (lower gray area in the figure) is not part of the model except for the circuit breakers which are included within the SCADA simulation and their functionality is described in Section 3.6.

Scripts for both topologies are located in the *topologies* folder. Listing 1 shows redactions parts from the extended topology script (*SmartGridTopology.py*). It demonstrates how a networking device is created, how hosts are created (RTUs), how links between devices are established, how devices are started and lastly, how custom scripts are defined.

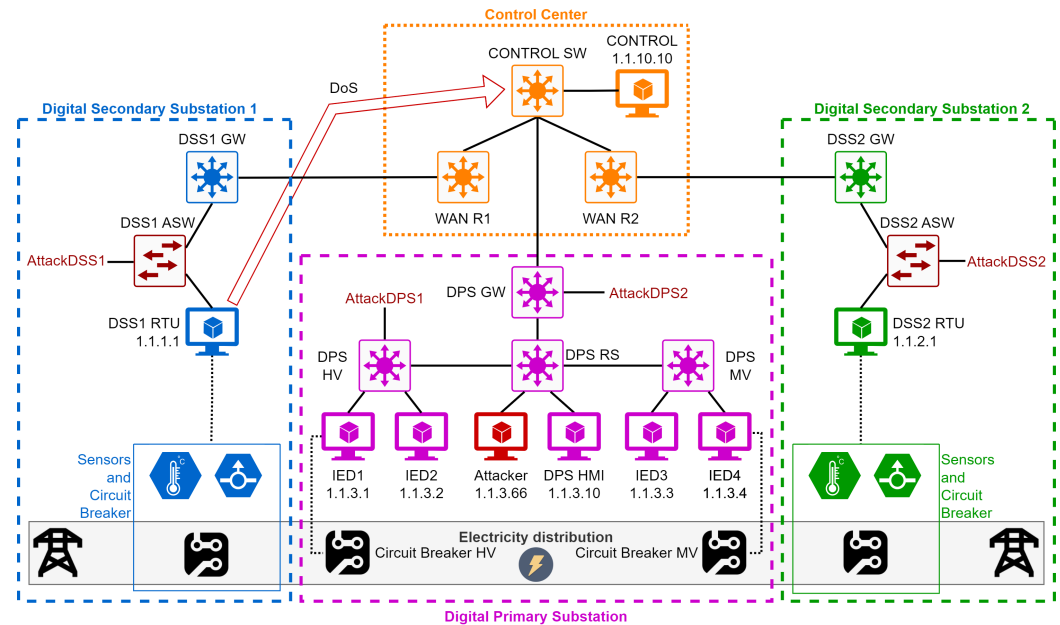


Figure 2. Smart Grid (extended) topology composed from one primary and two secondary substations.

Listing 1. Redacted parts from the topology script.

```

1  info( '*** Adding networking devices\n')
2  DSS1GW = net.addSwitch('DSS1GW', cls=switchType, dpid='1',
3      failMode='standalone')
4
5  info( '*** Adding hosts \n')
6  DSS1RTU = net.addHost('DSS1RTU', cls=Host, ip='1.1.1.1',
7      defaultRoute='1.1.10.10', mac='00:00:00:00:00:06')
8  DSS2RTU = net.addHost('DSS2RTU', cls=Host, ip='1.1.2.1',
9      defaultRoute='1.1.10.10', mac='00:00:00:00:00:07')
10
11 info( '*** Adding links\n')
12 net.addLink(WANR2, CONTROL SW)
13 net.addLink(WANR1, CONTROL SW)
14
15 info( '*** Starting networking devices\n')
16 net.get('DSS1GW').start([])
17 net.get('DSS2GW').start([])
18
19 info( '*** Preparing custom scripts \n')
20 CLI.do_sgsim_startcom_104 = sgsim_startcom_104
21 CLI.do_sgsim_startcom_goose = sgsim_startcom_goose

```

3.3. Emulated Communication

Communication between devices is realized using libIEC scripts for emulating IEC104, GOOSE and SV protocols. These protocols are responsible for:

- IEC104—used for communication between secondary substations and control center. The communication uses TCP for reliable delivery.
- GOOSE—used for local layer 2 (Ethernet) communication between IED1, IED4 and DPS HMI (Human Machine Interface).
- SV (Sampled Values)—used for local layer 2 (Ethernet) communication between IED2, IED3 and DPS HMI.

All the communication scripts were modified to correspond to the real traffic as provided by the National Smart Grid Laboratory (NSGL) [20] in order to fulfill the realistic needs of the project partners.

3.3.1. IEC 60870-5-104

IEC 60870-5-104 (IEC104) communication uses TCP (Transmission Control Protocol) for reliable delivery over WANs (Wide Area Network). It is primarily used for communication between the control center and substations. In the model it is implemented for communication between RTUs in secondary substations and the control center.

The following two modes of communication are implemented:

- Periodical communication—RTUs are sending messages of the *I* type with data from 5 sensors every 1 s. These sensors are identified by IOA (Information Object Address). Every successful delivery of the *I* message type by the control device is acknowledged by the *S* message type. Data values from the sensors are generated randomly from the following ranges (as to correspond to the real data traffic):
 - IOA 15: 52.2–52.9
 - IOA 16: 45.5–47.9
 - IOA 19: 0.25–0.27
 - IOA 23: 45.5–47.9
 - IOA 24: 0.003–0.105

IOA numbers are also corresponding to real sensor values.

- Read—request communication—the control device is sending *I* message type request in randomly generated time periods from the interval 1–60 s. The request is randomly generated from one sensor (IOA) in the range 1–7. The RTU replies with the *I* message type, which is then acknowledged by the *S* message type sent from the control device. The data message response is unique for every sensor and set either statically, or randomly generated from a defined interval.

Both types of the communication are included in scripts for RTU devices (located in *comlib_dss/sgdevices/RTU/rtu.c*) and the CONTROL device (located in *comlib_dss/sgdevices/CONTROL/control.c*) and automatically launched for corresponding devices upon executing the *sgsim_startcom_104* command.

3.3.2. GOOSE

GOOSE is a protocol used mostly between IEDs for low-latency communication during specific events such as line faults. GOOSE is a layer 2 protocol and it uses Ethernet frames for local delivery with layer 2 multicast for publisher-subscriber model communication. This is how this communication model is implemented in the SGSim:

- Layer 2 multicast—most switches treat layer 2 multicast as broadcast and the messages are flooded to all the ports (except the receiving port). This behavior is also utilized in the model as it uses Open vSwitches [21] which broadcast the GOOSE traffic (MAC destination address starting with 01:0c:cd:...).
- Local delivery—GOOSE messages are stopped at the boundary of the substation by an OpenFlow rule blocking all the traffic on the DPS GW device.

The GOOSE protocol communication in the SGSim is implemented in scripts for IED devices (located in *comlib_dps/sgdevices/IED_GOOSE/ied_goose.c*) and the HMI (located in *comlib_dps/sgdevices/DPSHMI_GOOSE*).

Listing 2 shows an example of how the main header fields and the payload of a GOOSE message are created (stNum, sqNum and the boolean value are passed as function parameters).

Listing 2. GOOSE payload configuration.

```

1 // Add main header fields
2 GoosePublisher_setGoID(publisher,
3     'SIP/VI3p1_27Undervoltage1/LLN0/Control_DataSet');
4 GoosePublisher_setStNumSqNum(publisher, stNum, sqNum);
5
6 // Add main payload
7 LinkedList_add(dataSetValues,
8     MmsValue_newFloat(generateRandomNumber(5.001, 5.01)));
9 LinkedList_add(dataSetValues, MmsValue_newBoolean(boolRand));
10 LinkedList_add(dataSetValues, MmsValue_newBitString(0));

```

3.3.3. SV

Sampled Values (SV) protocol uses periodical messages mostly for reporting sensor values. Like GOOSE, SV messages use publisher-subscriber model and are defined in scripts for IED devices (located in *comlib_dps/sgdevices/IED_SV/ied_sv.c*) and the HMI (located in *comlib_dps/sgdevices/DPSHMI_SV/dpshmi_sv.c*). Messages are typically send with a very high frequency (often under 1 ms). The SV communication in the model allows the frequency to be easily set (the default frequency is 50 ms), but values under 1 ms are not supported due to the hardware and OS limitations. Payload in this stage contains only artificial float data with timestamps. VLAN tagging is not used in the model.

3.4. Simulation of End Devices

End devices from the networking perspective (IEDs, RTUs, HMI and CONTROL) are emulated as Linux hosts. This means that they do not run any specialized power grid software, but use specific scripts to generate (or accept) appropriate network traffic (such as GOOSE, SV or IEC104).

These devices are created in the Mininet topology script (*/topologies/SmartGridTopology.py*), where their IP and MAC addresses can be configured. Listing 3 shows configuration of the host IED1 representing an IED from SIEMENS (MAC address starting with the B4:B1:5A).

Listing 3. Configuration of the host IED1.

```

1 IED1 = net.addHost('IED1', cls=Host, ip='1.1.1.3.1',
2     defaultRoute='1.1.10.10', mac='b4:b1:5a:00:00:01')

```

3.5. Impact Simulations

SGSim shows the impact of attacks on grid communication in real time. Based on the attack type, this can be observed via Wireshark (real time traffic monitoring), by using included performance monitoring scripts, or it can be visualized in the SCADA simulation (described in Section 3.6).

SGSim can be connected to external devices (real or virtualized) which can then start an attack. To simplify the process, SGSim also integrates DoS and FDI attacks for easy and quick verification of their impacts. These attacks work as follows:

- DoS attack assumes a successful infiltration of the DSS1 and spoofing of the RTU. The attack executes a script on the RTU which starts the *hping3* tool. The tool is launched with the command *sudo hping3-S-flood 1.1.10.10*. This will start generating as much TCP SYN traffic as possible with the CONTROL device as the target. The impacts can be observed by starting the IEC104 traffic in the “performance monitor” mode, which will display network performance parameters and compare them with predefined thresholds—classifying grid traffic into either functional or disrupted.
- False data injection attack assumes a successful infiltration of the DPS and physical connectivity to the DPS RS. The attack executes a script on the Attacker device which will start sending the same GOOSE messages which are already present in the topology,

but with higher sqNum values, with goID set to under voltage and with a purposefully high oil temperature reading, which will trigger the circuit breaker opening. This impact can be observed in the SCADA simulation.

The SGSim topologies contain pre-defined external connections (AttackDSS1/2 and AttackDPS1/2) for connecting external devices. These can be used for connecting another VM (as shown in Figure 1 with the Attacker (optional) Kali Linux VM), or other physical devices to run more complex attacks. These connections are created in the topology script via the *Intf* class using the name of a real interface located on the device as shown in Listing 4. These names might have to be modified to correspond to the interfaces on the target device.

Listing 4. Definition of external connections.

```

1 info( '*** Starting external connection\n')
2 DSS1ASW = net.addSwitch('DSS1ASW', cls=switchType, dpid='10',
3   failMode='standalone')
4 DSS2ASW = net.addSwitch('DSS2ASW', cls=switchType, dpid='11',
5   failMode='standalone')
6 Intf( 'enp0s3', node=DSS1ASW )
7 Intf( 'enp0s8', node=DSS2ASW )
8 Intf( 'enp0s9', node=DPSHV )
9 Intf( 'enp0s10', node=DPSGW )

```

3.6. SCADA Simulation

SGSim includes a simple SCADA simulation implemented as a web application. The application uses PHP server, SQL database, HTML with JavaScript and the AnyChart framework [22] for graphical visualization. The web part of the application has two web pages:

- Quick overview—shows a table with communication status from Digital Secondary Substations with information such as status (connected, disconnected), a timestamp of the last received IEC104 message and its ASDU type.
- Topology—provides an overview of the Smart Grid topology. It shows real-time visualization of the state of the grid. Rectangles around sensor and devices distinguish operation states (green color = normal state, red color = malfunction) and hovering above a sensor shows last received data (updated every 1 s). The functionality is further demonstrated in the set of videos available at [23].

Figure 3 shows the SCADA simulation on a running model with all communication connected except IED4 making the circuit breaker MV to open. Circuit breakers statuses are at this stage based only on values received from GOOSE messages from IED1 and IED4 respectively at the HMI device. If the received value is higher than 5, or there is no received message, the circuit breaker will open, which is visualized by the red color and text explaining the reason (either oil temperature too high, or no data received from the IED).

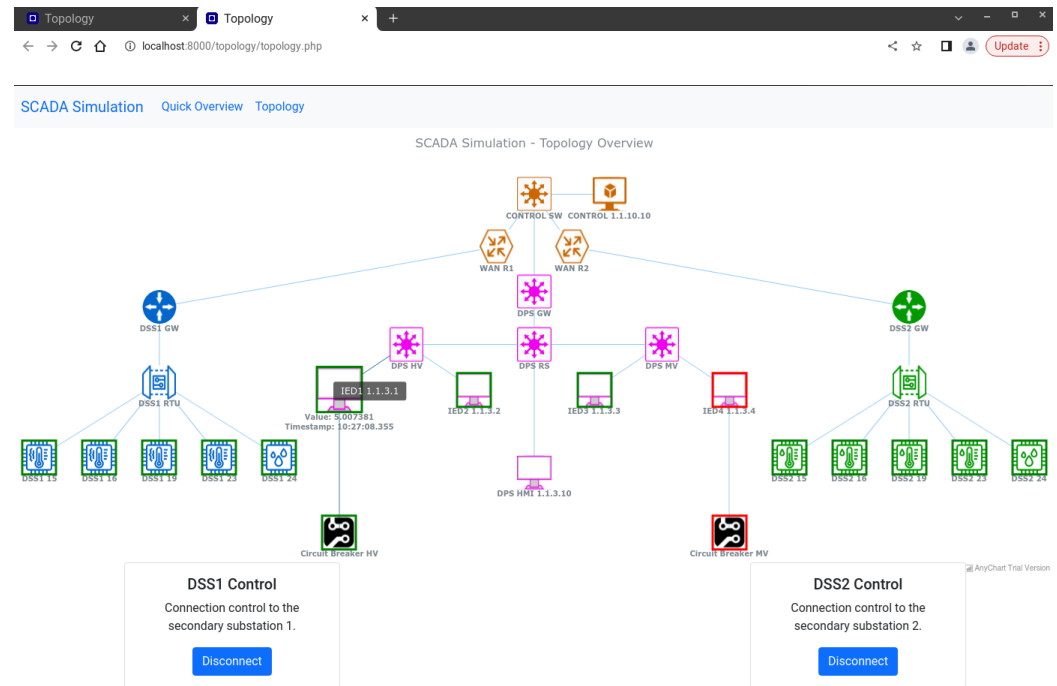


Figure 3. SCADA simulation.

3.7. SGSim CLI

The entire smart grid topology is emulated within the Mininet tool and is started either by user clicking on the desktop shortcut (launching the underlying Mininet scripts), or executing the scripts directly. The Mininet CLI will display current progress and provide the option to control the topology. This is achieved via custom functions defined in the Mininet's topology script. The following functions were implemented:

- `sgsim_startcom_104`—starts the IEC104 communication between the RTUs from DSS1 and DSS2 and the CONTROL device.
- `sgsim_startcom_goose`—starts the GOOSE communication between the IED1, IED4 and the DPS HMI.
- `sgsim_startcom_sglab_goose`—starts a modified GOOSE communication used for machine learning on generated custom traffic data.
- `sgsim_startcom_sv`—starts the SV communication between IED2, IED3 and the DPS HMI.
- `sgsim_startperfmon`—starts the IEC104 communication between the same devices, but with modified scripts with added timestamps for measurement and display of network performance parameters (delay, IPDV, jitter, packet loss and threshold exceeding percentage based on 60 s measurement).
- `sgsim_attackmirror`—creates SPAN (Switched Port Analyzer) ports on the DSS1 and DSS2 ASW on the ports leading outside the topology (where an attacker device can be connected physically or virtually through another VM). All the traffic from the other two ports will be mirrored to the SPAN port. This is achieved by inserting OpenFlow rules into the switches.
- `sgsim_attack_goose_fdi`—starts the FDI attack from the attacker device inside the DPS.
- `sgsim_attack_dos`—starts the DoS attack from the DSS1 RTU to the CONTROL device (assumes compromised RTU).

Listing 5 shows output of the help command used in the running model's Mininet console. Additional documentation for every command can be displayed with keyword help and the commands as a parameter as shown on the line 17 of the listing.

Listing 5. Overview of available commands.

```

1
2 mininet> help
3
4 Documented commands (type help <topic>):
5 =====
6 EOF      iperfudp      pingpair      sgsim_attackmirror
7 dpctl    link          pingpairfull  sgsim_startcom_104
8 dump     links         ports         sgsim_startcom_goose
9 exit     net          px            sgsim_startcom_sglab_goose
10 gterm    nodes        py            sgsim_startcom_sv
11 help     noecho       quit          sgsim_startperfmom
12 intfs    pingall      sgsim_attack_dos  sh
13 iperf    pingallfull  sgsim_attack_goose_fdi  source
14 switch   time          wait          xterm
15 x
16
17 mininet> help sgsim_startcom_104
18 Starts the IEC104 communication (periodical and read requests)
19 for both secondary substations.

```

3.8. Running the SGSim

Figure 4 shows running SGSim with IEC104 communication, GOOSE and SV communication scripts on the left side, the main console CLI on the bottom, the SCADA simulation in the right bottom corner and the Wireshark packet capture in the right upper corner.

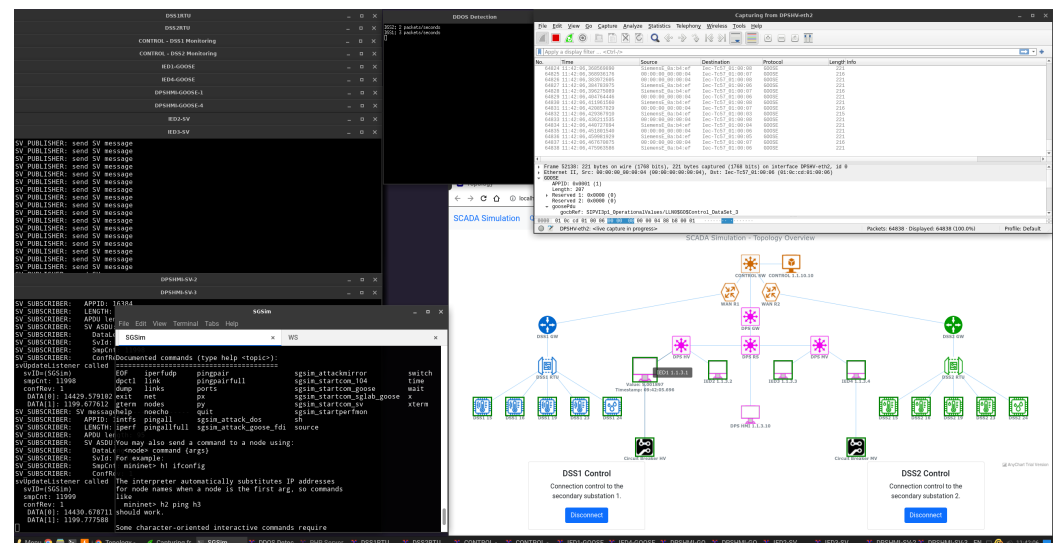


Figure 4. Running SGSim.

4. SGSim Use Case Scenarios and Validation

This section describes practical validation of the SGSim as well as experience from real world usage of the tool.

4.1. Use Case Scenarios of SGSim

SGSim is suitable for a wide range of use cases and was verified in the following use case scenarios conducted together with corresponding specialists:

- Impact simulations—the tool was used for verification of critical threats identified by the cyber security threats and attacks analysis approach within the mentioned projects. The results were used for enhancing the cyber security of digital substations.

- Cyber security awareness and training—the tool was used for visualization of impacts of cyber attacks and is currently being extended to be used as a training platform for substation operators and students of relevant fields.
- Configuration testing—SGSim was provided to substations operators for interconnection to the real hardware equipment and testing of various configuration of networking devices such as firewalls.
- Data generation—the tool was used as a platform to generate a large amount of real world traffic data for machine learning and AI applications. This was used in connection with the NSGL which provided limited data sets (as the time of the lab is strictly divided between various projects and is quite expensive) to which SGSim traffic scripts were modified and then run for the desired length to capture requested traffic volume.

4.2. Data Generation

The libIEC61850 [18], used in SGSim, provides example scripts for GOOSE and SV values communication, which can be easily modified to correspond to the real substation traffic. We have modified the library to correspond to the real traffic data provided by the NSGL and place the script into *comlib_dps/sgdevices/IED_GOOSE/ied_goose_sglab.c*. The script can be launched with the command *sgsim_startcom_sglab_goose* and the source code shows how to prepare similar data generation with custom values.

The script will set the custom undervoltage GoID, stNum, sqNum and values—float (randomly generated value from the interval 5.001–5.01), boolean and a bit string. The generated data was used for improving accuracy of a machine learning based anomaly detection algorithm which validated that the generated data corresponds to the captured real data.

4.3. SGSim Validation

SGSim was validated in several different scenarios described below.

Figure 5 shows Wireshark comparison of the captured traffic. The left window shows legitimate GOOSE traffic and the right one shows FDI traffic. The attack generates an identical message except in two fields: goID is set to *Undervoltage* and the float value is set higher than 10 (simulating too high oil temperature resulting in opening of the circuit breaker). Also note the automatically increased stNum (2) resulting in a smooth reception by the subscribers.

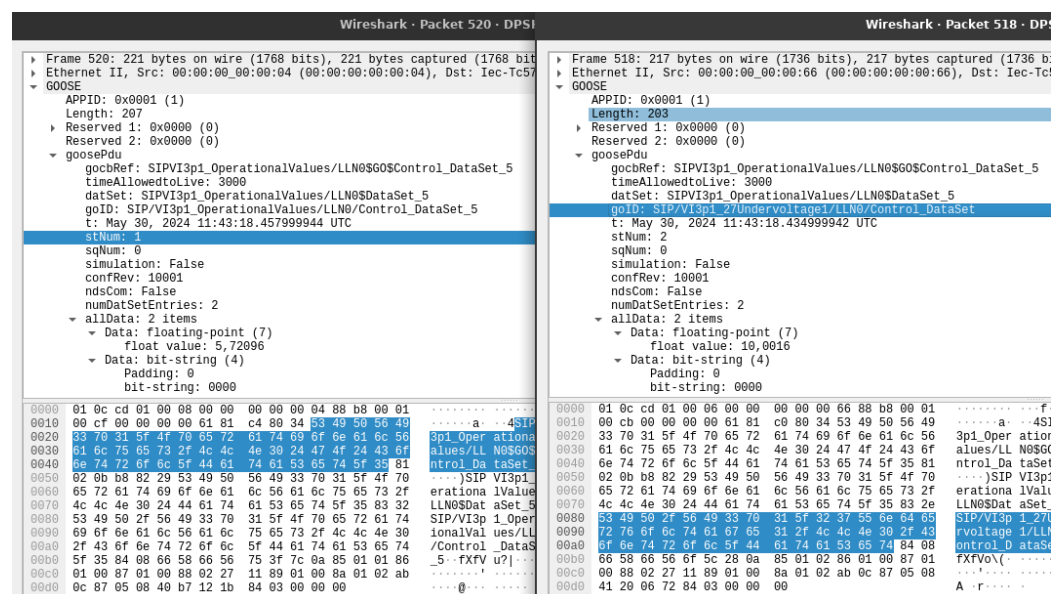


Figure 5. Wireshark analysis of legitimate GOOSE traffic (left) and FDI (right).

Figure 6 shows the impacts of the DoS attack on the GOOSE traffic. This was achieved by launching the iPerf measurement between IED2 and IED3. The lower part of the figure shows the overall amount of traffic in the substation (red increase is the DoS attack) and the upper part shows the transmission of GOOSE traffic. Spaces during the attack clearly show GOOSE traffic disruption, which results in the circuit breaker opening followed by a power outage.

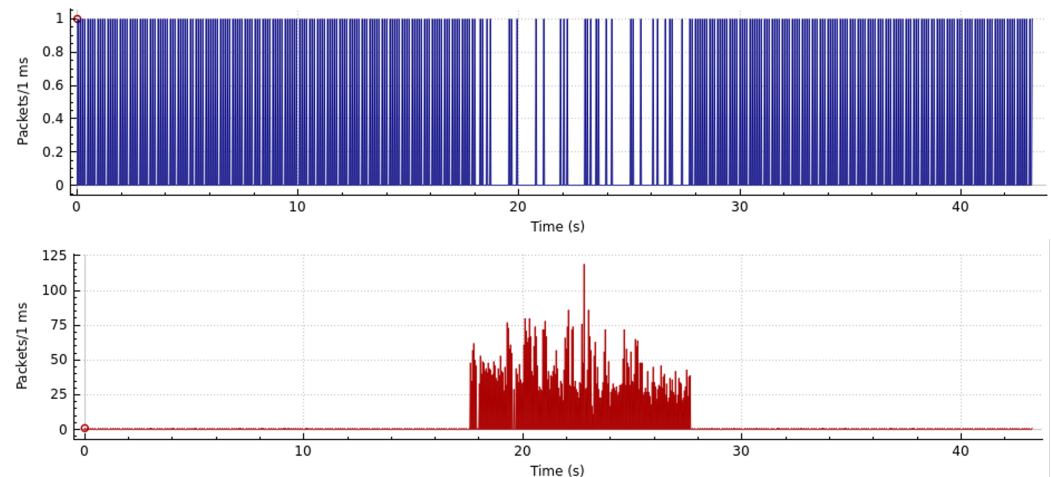


Figure 6. Impacts of DoS attack on the GOOSE traffic.

Finally, Figure 7 shows the tool performance in SV messages generation on an average laptop running in a VM with 2xCPU and 4 GB of RAM. The figure shows that with this configuration, the tool performs well when the messages are generated with spaces of at least 10 ms (error under 10%). If the frequency of messages is higher, the VM does not have enough performance and the timing between messages rises above the requested interval. For 5 ms intervals, it is around 30% so on average 6.5 ms, while for 1 ms the error reaches 90%, almost doubling the interval.

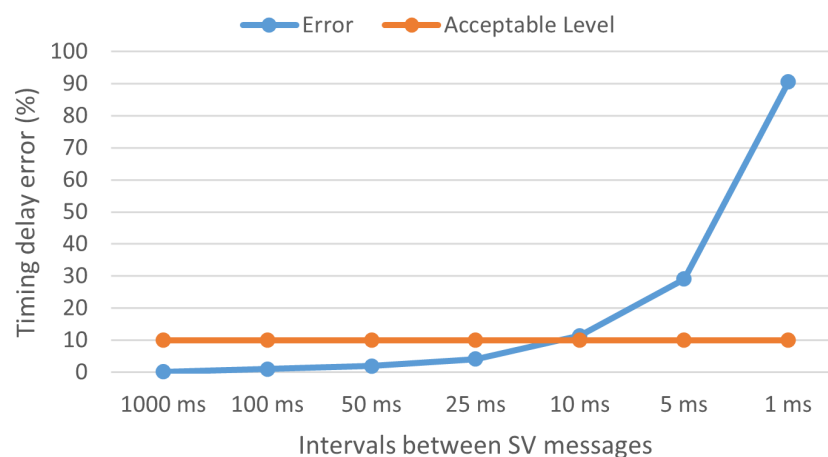


Figure 7. Timing delay error in SV messages generation.

4.4. Comparison with Other Tools

Table 1 shows SGSim comparison with the tools described in Section 2. Three categories are compared: portability (if the tool require specific devices), real traffic generation (if the tool can be used for generating realistic communication in large quantities) and network configuration testing (if the tool can be connected with substation network equipment such as firewalls and access control lists).

Table 1. Comparison of SGSim with similar tools.

Tool	Method	Portability	Real Traffic Generation	Network Configuration Testing
Multi Scale Model [11]	Mathematical model	Yes	No	No
Learning Platform [12]	RTDS HiL	No	Yes	Yes
HiL for Resiliency [13]	Real HW + Omicron	No	Yes	Yes
Cyber events impacts [14]	RTDS, PMUs + NS3	No	No	No
SGSim	VM + Mininet	Yes	Yes	Yes

Results show that SGSim is the only tool which complies with all three categories because it has the form of a virtual machine, and consequently can be used for realistic traffic generation and can be connected to real substation network equipment for configuration testing. SGSim also excels in cost-effectiveness when compared with the other tools. However, SGSim cannot compete with HW-based tools in terms of performance and emulation of grid equipment.

5. SGSim Survey Evaluation

SGSim was evaluated during a survey which took place between September–December 2023. In this survey, a scenario was prepared for participants in which they imported the SGSim VM on their machines and verified the tool in a 7-step interactive process. This covered features from basic functionality of the SGSim to attack simulations. As an alternative for participants who were not able to import the VM (due to restrictions to install any software in their company), a video demonstrating all the steps of the survey was prepared and used instead. Both types of participants were then asked to fill out the survey.

5.1. Survey's Methodology

The survey can be classified as descriptive research according to [24] with the goal of gathering information about SGSim at a single point in time (a bounded time frame). It was conducted according to the methodology presented in [25]. First, the population of interest was selected for sampling and divided by workplace areas and expertise as shown in Tables 2 and 3. The goal was to have at least 1 participant for every workplace area and with at least medium expertise in every area.

Table 2. Participants' workplace background.

Workplace Area	Percentage
Electricity grid operator	14.3%
University	42.9%
Research organization	28.5%
Energy regulation authority	14.3%

Table 3. Expertise of the participants.

Expertise area	Low	Medium	High
Linux	29%	42%	29%
Communication networks (IP)	14%	57%	29%
Cyber security of communication networks	29%	14%	57%
Electricity grid functionality	29%	14%	57%
Smart grid networks (IEC61850)	14%	72%	14%

Next, a combination of quantitative and qualitative research strategies was chosen. The quantitative questions were used for getting concrete answers for background and

technical questions. The qualitative questions were inserted to give the participants the opportunity for adding custom comments and suggestions.

Finally, a self-administered questionnaire method was selected for data collection and implemented with the Nettskjema platform [26] with all responses made anonymous.

5.2. Participants' Workplace Background

The first question collected participants' workplace background with the results shown in Table 2. Every sector was represented by at least one person with majority of people being from universities (researchers) and research organizations. This was an important goal of the survey as it allowed us to verify the SGSim's usability as a platform for professionals as well as for education and training.

The next two questions were focused on the expertise of the participants. More than 70% of the participants classified themselves as "senior position" as opposed to 14% for junior or prefer not to/cannot answer. More importantly, the distribution of the positions ensured that there was still at least 1 senior person for every workplace category.

The more detailed distribution of participants' expertise is shown in Table 3. This again demonstrated the diversity across all options among participants with slightly more frequent participation of medium to highly skilled experts.

5.3. Survey Results

The main questions of the survey were focused on evaluating various parameters of the SGSim as shown in Table 4. Every question allowed the participants to rate in a range of 1–5 (whole numbers only) and there was no option not to answer.

Table 4. SGSim evaluation results.

Area	Rating 1–5 (Higher Is Better)				
	1	2	3	4	5
Convenience of the VM form distribution	0%	0%	29%	14%	57%
Interface usability	0%	0%	14%	57%	29%
Documentation clarity	0%	0%	0%	43%	57%
Study instruction clarity	0%	0%	0%	33%	67%
Usefulness in terms of the project goals	0%	0%	43%	14%	43%
Security awareness potential	0%	0%	29%	29%	43%

Results show that the overall satisfaction with the SGSim and the provided documentation and study instructions (scenarios used for the survey) is very high with no answers being rated worse than 3 in technical aspects and not worse than 4 in documentation aspects.

The lower rating of the convenience of the VM distribution is connected with professionals with limited ability to install an additional software on their devices, while for researchers, the rating is very convenient. For this reason, a cloud version of the SGSim was being considered, but was not pursued further as it would require major changes, hardware GUI support and would limit the possibility of interconnection with hardware (or virtualized) local devices.

Interface usability was also well rated despite the fact that most functionality is being provided via a CLI. This demonstrates that end users of the tool are accustomed to similar interfaces and the lack of GUI is not seen as a problem.

Documentation and study instruction clarity were the best rated areas. We understand the importance of delivering concise documentation and instructions as the clarity helps not only the tool developers but all parties using the tool.

The usefulness of the SGSim in terms of the project goals was the lowest rated. This might be due to the two reasons. First, the ECODIS project [9] is mostly focused on physical development of digital substation pilots and the cyber security work supported with SGSim is only part of a single work package. Second, two of the survey participants were not

affiliated with the project and therefore might have decided to answer neutrally with the rating 3. Despite these reasons, the overall usefulness of the SGSim is rated very highly.

Finally, the security awareness potential of the SGSim is also rated very high with additional tips and comments provided in the final open questions.

5.4. Additional Comments/Open Questions

Participants had an option to fill two “open text” questions at the end of the survey. Condensed inputs are listed below:

- (1) Are there any features you are missing from the SGSim? Please list them.
 - ML (Machine Learning) in the detection of attacks and risk identification for various scenarios.
 - More detailed description of attacks (how do they originate in the network) and simulation of the first step—infection by malware or email.
- (2) Do you have any suggestions for further improvements of the SGSim or any other comments? Please list them.
 - Add more videos including a detailed walkthrough over all the SGSim features.
 - Add short CTF (capture the flag) exercises.
 - Add more attacks including a replay attack.
 - Add merging units which would allow to simulate grid operation and could be used for better staff training and students education.
 - Add more visualization features (for example attack detection and anomaly analysis).

These answers confirmed the high level of satisfaction with the tool and formed the direction for the future SGSim extensions and modifications.

5.5. Discussion

The overall rating of the SGSim was very positive. There were several additional positive comments provided in the last questions we have not included in Section 5.4. These comments mostly confirmed the validity of implemented topology and used communication patterns, relevance of the presented attacks and overall satisfaction with the tool.

We have also been asked to provide the tool to become a platform for the education of students of critical infrastructure protection courses. Finally, the tool is being used as an experimental platform within the COCOON project [27] which explores the data plane programmability concept for the protection of future smart grid networks.

6. Conclusions

In this work, we have summarized our approach for analyzing cyber security threats and attacks in digital substations based on a previous case study from Norway by showing the last step—impact simulations—in detail. We have shown how we approached the problem and developed a tool called SGSim which can be used not only for performing the impact simulations, but also in areas of general cyber security awareness, configuration testing and generation of large data sets. The tool is distributed as a virtual machine image without any requirements on specific hardware making it very portable and cost-efficient when compared to more complex tools based on hardware in the loop technology as well as real devices.

We validated the tool by comparing it with real data and by deploying it in several use case scenarios with corresponding experts in the area. We presented the tool results in term of SV messages generation performance and demonstrated the impact of a DoS attack on the GOOSE traffic.

Finally, we evaluated SGSim in a survey carried out among the experts. The survey demonstrated a high level of general satisfaction with the platform and provided several useful tips for further improvement of the tool, which we are planning to implement in our future work. This includes extending SGSim with additional attack scenarios, capture the flag exercises and working on providing better interactivity and visualization. We

will also continue expanding on tool dissemination through informative videos, as well as increasing tool usage among companies and educational institutions.

Author Contributions: Conceptualization, F.H., G.B.O. and S.Y.Y.; methodology, F.H., S.Y.Y. and G.B.O.; software, F.H.; validation, F.H.; formal analysis, F.H.; investigation, F.H.; resources, F.H.; data curation, F.H.; writing—original draft preparation, F.H., S.Y.Y. and G.B.O.; writing—review and editing, F.H., S.Y.Y. and G.B.O.; visualization, F.H.; supervision, S.Y.Y.; project administration, S.Y.Y.; funding acquisition, S.Y.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the ECODIS project funded by the Research Council of Norway under the grant number RCN 296550.

Data Availability Statement: No new data were created in this study. Source codes of the SGSim are available at: <https://github.com/filipholik/SmartGridSim>.

Acknowledgments: We would like to thank all the survey participants for their feedback about the SGSim and to Oscar Mathey for his contribution in development of the SCADA simulation GUI.

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

Abbreviations

The following abbreviations are used in this manuscript:

ASDU	Application Service Data Unit
CLI	Command Line Interface
CPU	Central Processing Unit
CTF	Capture The Flag
(D)DoS	(Distributed) Denial of Service
DPS	Digital Primary Substation
DSS	Digital Secondary Substation
FDI	False Data Injection
GOOSE	Generic Object Oriented Substation Event
GUI	Graphical User Interface
GW	Gateway
HiL	Hardware-in-the-Loop
HMI	Human Machine Interface
HV	High Voltage
HW	Hardware
ICT	Information and Communications Technology
IEC	International Electrotechnical Commission
IED	Intelligent Electronic Device
IOA	Information Object Address
IP	Internet Protocol
IPDV	Instantaneous Packet Delay Variation
LV	Low Voltage
MAC	Media Access Control
MitM	Man-in-the-Middle
MU	Merging Unit
MV	Medium Voltage
NSGL	National Smart Grid Laboratory
NS-3	Network Simulator 3
OT	Operational Technology
PCAP	Packet Capture
PMUs	Phasor Measurement Unit
RAM	Random-Access Memory
RTDS	Real Time Digital Simulator
RTU	Remote Terminal Unit
SCADA	Supervisory Control and Data Acquisition
SGSim	Smart Grid Simulator

SPAN	Switched Port Analyzer
SV	Sampled Values
SQL	Structured Query Language
SYN	Synchronize
TCP	Transmission Control Protocol
VM	Virtual Machine
VPN	Virtual Private Network
WAN	Wide Area Network

References

- Smith, J.C. Effective Security by Obscurity. *arXiv* **2022**, arXiv:2205.01547.
- Jack, V. Europe's Grid Is under a Cyberattack Deluge, Industry Warns. 2023. Available online: <https://www.politico.eu/article/energy-power-europe-grid-is-under-a-cyberattack-deluge-industry-warns/> (accessed on 23 April 2024).
- Casanovas, M.; Nghiem, A. Cybersecurity—Is the Power System Lagging Behind? 2023. Available online: <https://www.iea.org/commentaries/cybersecurity-is-the-power-system-lagging-behind> (accessed on 23 April 2024).
- Intelligence Community. Annual Threat Assessment of the U.S. Intelligence Community. Annual Threat Assessment, National Intelligence Council. 2023. Available online: <https://www.dni.gov/files/ODNI/documents/assessments/ATA-2023-Unclassified-Report.pdf> (accessed on 2 June 2024).
- ENISA. NIS INVESTMENTS Cybersecurity Policy Assessment. Technical Report TP-02-23-138-EN-N, ENISA. 2023. ISBN 978-92-9204-648-4. Available online: <https://www.enisa.europa.eu/publications/nis-investments-2023> (accessed on 11 June 2024).
- Boje, C.; Guerriero, A.; Kubicki, S.; Rezgui, Y. Towards a semantic Construction Digital Twin: Directions for future research. *Autom. Constr.* **2020**, *114*, 103179. [CrossRef]
- Hussain, S.; Iqbal, A.; Hussain, S.; Zanero, S.; Shikfa, A.; Ragaini, E.; Khan, I.; Alammari, R. A novel hybrid methodology to secure GOOSE messages against cyberattacks in smart grids. *Sci. Rep.* **2023**, *13*, 1857. [CrossRef] [PubMed]
- Yildirim Yayilgan, S.; Holik, F.; Abomhara, M.; Abraham, D.; Gebremedhin, A. An Approach for Analyzing Cyber Security Threats and Attacks: A Case Study of Digital Substations in Norway. *Electronics* **2022**, *11*, 4006. [CrossRef]
- SINTEF. ECODIS—Engineering and Condition Monitoring in Digital Substations. 2024. Available online: <https://www.sintef.no/en/projects/2019/ecodis/> (accessed on 22 January 2024).
- Lnett. InterSecure. 2024. Available online: <https://www.l-nett.no/fou-og-innovasjon/fou-prosjekter/intersecure> (accessed on 22 January 2024).
- Hao, W.; Yang, Q.; Li, Z.; Hu, S.; Liu, B.; Ruan, W. Multi-Scale Traffic Aware Cybersecurity Situational Awareness Online Model for Intelligent Power Substation Communication Network. *IEEE Internet Things J.* **2023**, *10*, 1666–1681. [CrossRef]
- Manbachi, M.; Nayak, J.; Hammami, M.; Bucio, A.G. Virtualized Experiential Learning Platform for Substation Automation and Industrial Control Cybersecurity. In Proceedings of the 2022 IEEE Electrical Power and Energy Conference (EPEC), Virtual Event, 5–7 December 2022; pp. 61–66. [CrossRef]
- Jorgensen, P.A.; Waltoft-Olsen, A.; Houmb, S.H.; Toppe, A.L.; Soltvedt, T.G.; Muggerud, H.K. Building a Hardware-in-the-Loop (HiL) Digital Energy Station Infrastructure for Cyber Operation Resiliency Testing. In Proceedings of the 2022 IEEE/ACM 3rd International Workshop on Engineering and Cybersecurity of Critical Systems (EnCyCriS), Pittsburgh, PA, USA, 16 May 2022; pp. 9–16. [CrossRef]
- Liu, R.; Vellaithurai, C.; Biswas, S.S.; Gamage, T.T.; Srivastava, A.K. Analyzing the Cyber-Physical Impact of Cyber Events on the Power Grid. *IEEE Trans. Smart Grid* **2015**, *6*, 2444–2453. [CrossRef]
- RTDS Technologies Inc AMETEK. Real-Time Simulation and Hardware-in-the-Loop Testing. 2024. Available online: <https://www.rtds.com/> (accessed on 13 May 2024).
- nsnam. ns-3 Network Simulator. 2024. Available online: <https://www.nsnam.org/> (accessed on 13 May 2024).
- OpalRT. OpalRT Technologies. 2024. Available online: <https://www.opal-rt.com/> (accessed on 7 May 2024).
- Zillgith, M. libIEC61850/lib60870. 2022. Available online: <https://libiec61850.com/> (accessed on 29 November 2023).
- Holik, F. SmartGridSim—GitHub. 2024. Available online: <https://github.com/filipholik/SmartGridSim> (accessed on 11 June 2024).
- SINTEF. Smart Grid Laboratory. 2020. Available online: <https://www.sintef.no/en/all-laboratories/smartgridlaboratory/> (accessed on 20 February 2024).
- Linux Foundation. Production Quality, Multilayer Open Virtual Switch. 2023. Available online: <https://www.openvswitch.org/> (accessed on 17 April 2024).
- AnyChart. AnyChart. 2024. Available online: <https://www.anychart.com/> (accessed on 16 May 2024).
- Holik, F. Youtube: SGSim—Smart Grid Simulation Model. 2023. Available online: <https://www.youtube.com/playlist?list=PLL3pXQi4Q7Ur03kmHE2SqvD4RFFRJde7> (accessed on 22 January 2024).
- Kelley, K. Good practice in the conduct and reporting of survey research. *Int. J. Qual. Health Care* **2003**, *15*, 261–266. [CrossRef] [PubMed]
- Ponto, J. Understanding and Evaluating Survey Research. *J. Adv. Pract. Oncol.* **2015**, *6*, 168–171. [PubMed]

-
26. University of Oslo. Nettskjema. 2024. Available online: <https://nettskjema.no/> (accessed on 11 June 2024).
 27. COCOON. Cooperative Cyber Protection for Modern Power Grids. 2024. EU Horizon Europe Project Number: HORIZON-CL3-2022-CS-01. Available online: <https://www.cyber-cocoon.eu> (accessed on 17 April 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.