Personal Health Metrics Data Management Using Symmetric 5G Data Channels

Razvan Bocu 1,2,*,†, Anca Vasilescu 1,‡ and Delia Monica Duca Iliescu 1,‡

1 Department of Mathematics and Computer Science, Transilvania University of Brașov, 500036 Brașov, Romania; vasilex@unitbv.ro (A.V.); delia.duca@unitbv.ro (D.M.D.I.)
2 Department of Research and Technology, Siemens Industry Software, 500203 Brașov, Romania
* Correspondence: razvan.bocu@unitbv.ro; Tel.: +40-732011010
† Current address: Blvd. Iuliu Maniu Nr. 50, 500091 Brașov, Romania.
‡ These authors contributed equally to this work.

Abstract: The integrated collection of personal health data represents a relevant research topic, which is enhanced further by the development of next-generation mobile networks that can be used in order to transport the acquired medical data. The gathering of personal health data has become recently feasible using relevant wearable personal devices. Nevertheless, these devices do not possess sufficient computational power, and do not offer proper local data storage capabilities. This paper presents an integrated personal health metrics data management system, which considers a virtualized symmetric 5G data transportation system. The personal health data are acquired using a client application component, which is normally deployed on the user’s mobile device, regardless it is a smartphone, smartwatch, or another kind of personal mobile device. The collected data are securely transported to the cloud data processing components, using a virtualized 5G infrastructure and homomorphically encrypted data packages. The system has been comprehensively assessed through the consideration of a real-world use case, which is presented.

Keywords: data privacy; homomorphic encryption; personal health data; 5G data links; data privacy; distributed system

1. Introduction

The ubiquity of personal wearable devices essentially changes the collection of personal health metrics data. These devices are normally featured by sensors, which acquire the necessary medical data. This assertion suggests that the patients’ medical parameters can be relatively easily collected, and used in the context of relevant medical scenarios. This data collection process generates significant amounts of personal health information data. The restricted storage and computational features of the personal wearable devices imply that the local processing of the acquired data is impossible. Consequently, the collected data must be handled by external software components. Considering the personal and medical nature of the acquired data, any relevant system should implement the necessary data privacy mechanisms. In this context, the data privacy denotes two aspects of the relevant problematic. First, the data transportation channel that links the wearable mobile devices to the data processing components should deliver data in a secure manner. This requirement is implemented through the proper configuration of virtualized 5G data channels, and also considering homomorphically encrypted data packets. Second, the data processing components should address the acquired data without any knowledge regarding the patient’s identity, or the values of the acquired personal health information data. The integrated system that is presented in this paper considers a comprehensive homomorphic encryption model that implements a complete data privacy mechanism. It also relies on the properly implemented secure 5G data channels. The homomorphic encryption routines implement the proper computations on encrypted data. The results of the data processing routines are
encrypted and fully consider the bit (binary) values of the related plain text data [1]. Consequently, the decrypted results correspond to the outcomes that are obtained considering the same operations that are conducted on plain text data.

1.1. Remarks Concerning the General Principle of Data Privacy in e-Health Systems

This type of approach to preserve the data privacy is efficient regarding the system that is presented in this article. Furthermore, this is a necessary approach considering the use cases that involve personal health data, which are collected by mobile wearable devices in the realm of e-Health systems. The following paragraphs discuss on the relevant contributions that have been accomplished in the scope of privacy-preserving information systems.

The contributions that are reported in [2] discuss on certain encryption schemes [3] that ensure data privacy [4]. Nevertheless, the collected data are aggregated at the level of the client [5]. Thus, the Sum aggregate and Min aggregate routines are specified using an additive homomorphic encryption model in [2]. It is relevant to note that these operations consider a set of values as input, and generate one value as output. Moreover, they represent flexible and computationally efficient routines that are used in other contexts that presume the aggregated processing of data. As an example, let us consider the queries that are conducted on relational databases, which involve multiple table columns. The relevant computations are designed to be conducted on the client devices in the case of most of the currently available similar approaches. The operations that implement the homomorphic encryption routines are generally computationally expensive. The computational models that are reported by the existing literature are generally unsuitable for the use case that is presented in this paper. The contributions that are reported in [3, 4] consider the transfer of required data processing operations to the cloud. Moreover, the authors of [5] described a privacy preserving sum aggregation, which performs a significant part of the data processing operations at the level of the client devices. Furthermore, the authors of [6] proposed a homomorphic encryption scheme that is solely proper in order to perform the sum operation, which is essentially insufficient for the computational use case that this paper describes. It is immediate to understand that it is necessary to describe a data computation mechanism, which allows for the basic arithmetic operations to be conducted over the encrypted data. This is referred to as verifiable computation.

The concept of verifiable computation has been initially described by Gennaro et al. [7]. This mechanism generally allows computationally constrained mobile client devices to offload their data processing tasks to one or several third-party workers. Furthermore, the client devices have the opportunity to check the correctness of the results that are received. In [8], the authors presented a verifiable computational approach, which considers the input in a plain text format. Moreover, [9] described a data processing scheme, which conducts the homomorphic data aggregation relative to e-Health information systems. Nevertheless, this model cannot reliably check the correctness of the obtained results, which is essential for the integrated system that is described in this paper. Moreover, this paper also proposes a publicly verifiable data processing scheme, which pertains to large polynomials and matrices. Additionally, the authors of [10] proposed a verifiable delegated data processing model, which uses set structures and operations, such as set union, set intersection, and set difference. It is relevant to note that these algorithmic models are applicable only to input data that are supplied in plain text format.

1.2. Relevant Existing Contributions and Essential Research Gaps

The contribution that is reported in [11] concerns a verifiable data processing model, which is applied to encrypted input data relative to m-Health (mobile health) information systems. The algorithmic model of the accumulation tree is described in [12], with the goal to check the results of geographical proximity tests. Moreover, the authors of [13] presented the results that were obtained regarding the verifiable computation, which considers encrypted input data. Thus, the majority of the existing approaches conduct the
data processing on the client devices. This approach is essentially unfeasible relative to the integrated data management system that is described in this paper.

The positive aspects regarding the data storage and processing in the cloud are easily discernible. Nevertheless, some drawbacks may be easily identified [14]. The determination of the proper security models, which provide the security of private data, represents a relevant problematic that poses significant issues to the cloud service providers [15].

These service providers generally design and deploy multiple layers of complex security models. Nevertheless, the unencrypted data may still be used through the consideration of appropriate intrusion techniques. Therefore, it is necessary to encrypt the data prior to sending them to the external data processing components, and retrieve them using various search techniques relative to the encrypted data. The surveyed relevant contributions determine a significant computation overhead relative to the mobile client devices [16]. This remark is particularly applicable to the personal mobile devices, which acquire the medical data that are processed by the integrated system that is described in this paper. Generally, it can be stated that there are approaches, such as those that are reported in [17], which do not implement proper security mechanisms that should protect the data privacy [18], during their transmission through the proper data channels. The adequate manipulation of personal health information (PHI) data is related to ethical principles and formal regulations [19]. The integrated data processing system is based on an architecture that considers all the necessary constraints. The authors of [20] presented the general features and the life cycle of the services that are deployed into a cloud-based data processing environment.

Since C. Gentry proposed the concept of homomorphic encryption in 2009 [1], a substantial research effort has been put into the improvement [21] of this computationally expensive approach. Therefore, it should have been normally compatible with particular real-world scenarios, which required certain powerful hardware resources [22]. Additionally, the early homomorphic encryption models were excessively computationally expensive relative to the intended use cases [23]. Consequently, the related algorithmic apparatus has been enhanced considering multiple development stages [24]. The authors of papers presented relevant mathematical and algorithmic models, which make the homomorphic encryption routines more efficient. It is equally relevant to mention the papers [25–27], as they extended the initial set of algorithms with useful computational features. The algorithmic apparatus that is described in [28] has been considered during the specification of the integrated data management system’s data processing components [29,30]. Nevertheless, the comprehensive validity assessment that we conducted [31] showed that even the optimized homomorphic encryption models are not feasible [32] for the timely processing of the acquired medical data on the mobile client devices [33]. Moreover, it is significant to mention the relevant contributions that pertain to the realm of ubiquitous systems. Thus, the scientific contribution that is reported in [34] presents a software application that is determined by two interesting functional requirements. First, the system is capable to perform the semantic analysis of data that are generated by user interactions relative to various contextual parameters during the usual activities of daily living (ADL). This is accomplished with the purpose to determine the set of relevant behavioral patterns that support the involved complex activities. Furthermore, the software system includes an algorithmic routine that is used in order to support the appropriate decision-making processes. The architectural model that is reported in this paper may influence the specification and implementation of a near real-time cardiac abnormality detection module, which may be included in a future version of the integrated medical data management system. It is relevant to mention the related contribution reported in [35]. Furthermore, the authors of [36] presented a general architecture of an ubiquitous system that is intended for general medical use case scenarios. Additionally, it is also useful to mention the survey work that is reported in [37].

The existing relevant computational models are usually improper for the design and implementation of an efficient privacy-preserving personal data processing system relative
to the four fundamental functional aspects: the collection of the medical data using the mobile client devices, their transfer to the central data processing components, the proper and secure storage of these data, and the privacy-preserving data processing [38]. The integrated data processing system is one of the few relevant information systems that considers both the distinction between the long-term data storage and data processing paths, and the functional requirement to efficiently enroll any compatible client medical data collection devices [39]. Moreover, the data processing components are able to use the storage and processing services that are used by the related cloud-based platform. This ensures the long-term scalability of the system [40]. The following sections describe the integrated data management system relative to the features that distinguish it from most existing similar approaches. Furthermore, it is the only relevant similar system, which considers virtualized 5G data transmission channels. This contributes to the versatile and economical deployment of its components on virtually any physical computational infrastructure.

The rest of the paper is structured considering the following sections. The next section describes the system considering its architectural components. The following section describes the technical architecture of the broadband 5G data transmission subsystem. Furthermore, relevant considerations concerning the optimization model are discussed. Consequently, the necessary implementation details relative to the specific use case are provided, and the practical system performance is evaluated through the consideration of a real-world use case. The last section concludes the paper.

1.3. Conclusive Remarks Concerning the Literature Review

The survey that was conducted suggests that although certain progresses have been made regarding the verifiable computation over encrypted data, no suitable system architecture was described, which would accommodate the requirements that are considered by the integrated data management system that is described in this paper. These are the following.

- The collection of the medical data using the mobile client devices.
- The data transfer to the central data processing components.
- The proper and secure storage of these data, and the privacy-preserving data processing.
- The specification of a flexible and decoupled system architecture, which would allow for an efficient extension and re-structuring of the system to occur in the future.
- The consideration of all the legal and formal requirements that are enforced by American and European regulations.
- The efficient integration of the system in the software frameworks of hospitals, clinics, and other medical facilities, while considering the logical differences that exist among the relevant types of actors, such as patients, doctors, and nurses.

The integrated medical data management system, which is described in this paper, considers all of these requirements.

2. Essentials Concerning the System Architecture

The generally considered encryption schemes, such as AES (Advanced Encryption Standard), are not compatible with the implementation of arithmetic operations relative to the encrypted data, because their arithmetic architecture does not support the specification of the basic mandatory operations, addition and multiplication, directly over the encrypted data. Thus, using the secret decryption key in order to obtain the plain text data represents the only possible operation. Consequently, it can be stated that the standard encryption models propose a mechanism that allows for the secure storage of the data, without the possibility to compute them.

The Fully Homomorphic Encryption (FHE) models provide the opportunity to implement data processing routines relative to the encrypted data, while completely disregarding the significance of the original unencrypted data. The integrated data processing system relates to the usage of the fully homomorphic encryption models, which implies that the personal health information (PHI) is safely acquired and processed. The personal health
data are computed by the data processing components considering their encrypted form. Thus, the implemented data privacy model is optimal for the considered use case, and the resulting system performance does not negatively impact the end user’s experience.

The architecture of the integrated data processing system is described in Figure 1. The blue arrows represent an indication concerning the directional flow of the data between the human actors and the relevant system components, and also through the internal system structures. The privacy of the processed data is relevant considering the four main stages that support the operation of the data transmission pipeline. Thus, the first stage pertains to the data acquisition using each enrolled individual’s wearable or mobile device. Consequently, the second stage relates to the safe data transmission to the backend components using the secure 5G data channels. The third stage pertains to the actual storage of the collected personal health information data, while the last stage implements the privacy-preserving data processing. The relevant components are installed on the IBM Cloud [41] platform. Furthermore, the acquired data are effectively stored using IBM Cloudant [42]. The relevant entities are modeled through the appropriate JSON documents, which essentially retain the informational structure of the main classes that are part of the integrated data management system’s solution. The relevant non-disclosure agreement, which legally formalizes this research process, prevents the full structure of the database to be revealed. Nevertheless, the interested reader may consult the main database entities in the component diagram that is presented in Figure 3. Moreover, the required data computations are conducted through the consideration of the Apache Spark platform. The data processing events are handled using the Apache OpenWhisk programming service [43]. The following sections provide further details concerning this integrated data storage and processing system.

![Figure 1. General System Architecture.](image)

The data transmission channel is essentially connected to the data processing components relative to the last two stages: the data storage, and the privacy-preserving data processing. The outcomes of the operations are returned to the mobile client devices in an encrypted form using the available secure 5G data transmission channel. This integrated system architecture fundamentally differs from most of the existing similar approaches, considering that it implements an end-to-end privacy-preserving data protection mechanism, which discloses the data only when they arrive on the requesting mobile client data device, without significantly affecting the execution times and the end user’s experience. These features are not characteristic to most of the existing similar approaches.

Technical Architecture of the 5G Data Transmission Pipeline

The personal health data collection devices mostly use 5G data communication channels in order to send the acquired data to the central data processing components. Therefore, it is essential that the architecture of the 5G data collection channels ensure an additional layer of data security, which is added to the intrinsic private nature of the homomorphically encrypted data. This essentially creates a system that ensures absolute data privacy...
The design of efficient and secure architectural models for the design and implementation of 5G data networks represents the object of intense research efforts. Thus, any proper approach envisions two perspectives [44]. The data perspective determines the real-time analysis of the transported data using software-defined data channels. The control perspective is connected to the proper implementation of the administrative tasks.

3. Essential Concepts Regarding the Optimization Model

3.1. Fully Homomorphic Encryption Model

The fully homomorphic encryption model that the integrated data management system considers is thoroughly presented in [26]. It is generally known as the Brakerski–Gentry–Vaikuntanathan (BGV) fully homomorphic encryption (FHE) model. We have extensively assessed and tested the existing FHE schemes considering simulated test infrastructures. It was determined that most of the FHE schemes are excessively resource intensive, even relative to sufficiently powerful hardware, particularly as a consequence of the expensive noise elimination (recrypt) operations, which are performed after each multiplication operation [1,27]. We have found the BGV model to be the only computationally proper solution relative to the integrated data management system. This is justified by the BGV scheme, which specifies a leveled FHE scheme that disregards the noise elimination operations. This approach envisions a more efficient noise management algorithm, which is referred to as modulus-switching. This optimization model is completely explained in [21]. It implies that cascaded homomorphic multiplications ($X_h$) can be performed, while circumventing the possibility to face decryption errors. This potential problem would render the precise privacy-preserving function of the system impossible. The following paragraphs present the four FHE operations that are implemented. Essentially, the system considers a parameter $L$ (the Level), which must be precisely determined before starting any effective data processing instruction. The level $L$ depends on the number of the multiplication operations that are necessary considering the particular computational context.

The first kind of FHE operation that the integrated data management system implements is the homomorphic addition ($+_{h}$). This operation takes as input two ciphertexts that relate to slot-wise XOR operations of the respective unencrypted elements. The second type of FHE operation that the integrated data management system implements is the homomorphic multiplication ($X_h$). This operation takes as input two ciphertexts that relate to a slot-wise AND function that is applied on the respective unencrypted elements. Each multiplication increments by 1 the related level $L$. Thus, the depth of the multiplication operations determines the calibrated value of the level $L$. Following, the rotate ($<<<h,>>>h$) represents an operation that allows for the defining storage bits to be rotated. Additionally, the select($select_{mask}$) is an operation that, in essence, recovers the potentially altered slots (bits) of the data elements that are generated by the rotate operation. Consequently, the select operation preserves the consistency of the processed data.

3.2. The Improved Fully Homomorphic Encryption Model

The integrated data management system considers the optimal implementation of the data processing components, which must safely compute, as required, the personal health information data. The communication data path is described in Figure 2, which suggests that each bit of the unencrypted data is adequately concatenated to the related plain text message. The generation of the ciphertext occurs according to the steps determined by the top data path. The direct processing of the encrypted data is the fundamental advantage of this privacy-preserving data processing scheme. The transformation of the processed data into a binary format is related to the bottom data processing path illustrated in Figure 2. It is implemented using the functions that are called computation ($f_{,}(.)$) and aggregation ($f_{a}(.)$).
Furthermore, let us observe the architectural structural of the main data management module, as it is described by the component diagram, which is represented in Figure 3.

In Figure 3, the component that is labeled Data Collection Interface represents the software routines that are necessary in order to properly acquire the data that are collected by the user-side devices. Moreover, the components that are depicted in the central part of the diagram describe the main types of actors that can interact with the system: the medical actors, and the administrative users. Furthermore, the rightmost section of the diagram relates to the secure data management components. This includes the software module that realizes the connection to the Cloudant non-relational database engine, which is distinctly highlighted.

3.3. Scope of the System

The system considers a functional architecture, which offers sufficient flexibility that is enough to accommodate any use case scenario that requires the collection of sensitive client data using a particular mobile or wearable device, and their safe transportation, storage and processing by the data processing components. The system can be customized in order to accommodate various current and future data collection sensors. The validity and appropriateness for the intended scope of the integrated data management system is assessed through the collection of cardiac rhythm data. The data storage and processing components possess the functional capability to persist the processed medical data and
supply them as necessary to the requesting client using the available secure 5G data channels. Moreover, the system also evaluates the medical condition, which is called the delayed repolarization of the heart syndrome (DRHS) [29].

4. Relevant Implementation Details

4.1. Remarks Regarding Virtualized Wireless Network Function

The virtualized wireless network function (VWNF) is fundamental for the efficient design and implementation of 5G networks. This has been considered in order to deploy the core of the virtual 5G network that supports the function of the integrated data management system. This approach allows for the self-sufficient specification of the logical 5G network, which supports the overall operation of the integrated data management system. Moreover, it also creates the possibility for the specialized logical 5G network that supports the function of the integrated data management system to be deployed on certain hardware and software infrastructures, such as those that are offered by cloud service providers or telecommunications service providers [45,46]. This mechanism was used in order to specify and deploy the necessary specialized networked services. We have observed that the virtualized networked environment supplies the necessary logical flexibility and scalability [47].

We have effectively observed that this logical mechanism is compatible with the proper processing of the personal health information data traffic that is sent through the logically defined 5G network. The contribution that is presented in [48] describes a possible use case of this logical networking mechanism [49].

We have also analyzed the logical 5G networked structure considering the utilization of the radio resources, and it was determined that this approach optimizes the allocation and usage of the necessary radio resources. Thus, we were able to define logical sub-networks that are conducting distinct analyses of the 5G data traffic using individual instances of the integrated data management system. It is immediate to note that the findings that are reported in this paper extend and refine the research effort that is reported in [50]. It can further be stated that the experimental work that we conducted acknowledges that logical 5G networks that are adequately defined and sized are capable to support even information systems that work with large amounts of real-time data, such as the integrated data management system.

4.2. General Algorithmic Remarks

The real-world operation of the integrated data management system relies on two variables. Thus, the properly calibrated level $L$ is essential for the time efficiency of the system. Furthermore, the efficient functional behavior of the system depends on the number of multiplication and rotation operations, which are computationally expensive. The multiplication operation is also important because it determines the calibration of the level $L$. The integrated data management system implements a series of enhancements that relates to the decrease of the level $L$, and the general number of the considered FHE iterations. They also ensure that the number of FHE operations is kept at the minimum possible level. It can easily be understood that the system is designed in order to efficiently calibrate the level $L$. This essential operation uses $N_{CT}$ ciphertexts as input data. This encrypts an array of $n$ bits that privately stores cardiac rhythm data. The complexity of the main processing algorithmic routine depends on the value of the level $L$. Thus, ignoring the logarithmic factors, the model is characterized by a complexity of $\tilde{O}(L^2)$, which depends on the value of the level $L$. It is immediate to note that it is required to maintain the value of the level $L$ at an appropriate low level during the fully homomorphic encryption computational processes.

4.3. The Detection of the Average Heart Rate

The determination of this physiological parameter is based on $N_{CT}$ ciphertexts. The algorithmic optimization that is implemented designates two main types of enhancements.
The first one relates to the decrease of the computationally expensive multiplication operations. Moreover, the decrease of the computation operations depth is also applied, while the level $L$ is also optimally calibrated.

The addition operation is enhanced considering two functional aspects. They are referred to as the additive compression and the prefixed parallel addition. The additive compression translates three data inputs $(H, M, F)$, each of them structured considering $n$ bits, into two outputs. The outputs are designated by the $A_R$ (addition result), and $L_{OVER}$ (leftover). The $A_R = H \Delta MAF$, and $L_{OVER} = [(H \times M) \nabla (H \times F) \nabla (M \times F)]$. Here, $\Delta$ designates an additive single instruction multiple data (SIMD) operation. Furthermore, the nabla operand ($\nabla$) refers to a SIMD operation that relates to the entire set of $n$ bits of the input data considering a parallel data processing model. The prefixed parallel addition has been designed and implemented using the algorithmic model, which is described in [27].

The calculation of the average heart rate uses $N_{CT}$ ciphertexts. These encrypt the input messages that are described by $n$ bits. The first stage of this data processing mechanism considers the additive compression in order to translate $N_{CT}$ ciphertexts into two ciphertexts. Moreover, the two ciphertexts that are obtained are summarized using the operation of prefixed parallel addition. The comprehensive real-world performance assessment suggests that the system efficiently processes the personal health information data, while the 5G data channels do not place a perceptible overhead on the enrolled users' experience.

### 4.4. Determination of the Delayed Repolarization of the Heart

The determination of this abnormal cardiac condition relates to the computational model that is presented in [33]. The fundamental equation that is described in [33] is improved. Consequently, let us study the following two mathematical expressions.

\[
\frac{T_{QT}}{\sqrt{T_{RR}}} > 475 \text{ms} \Rightarrow T_{QT}^2 > T_{RR} \times 225,625 \tag{1}
\]

\[
\Rightarrow T_{QTH} > T_{RRH} \tag{2}
\]

The expressions $T_{QT}^2 = T_{QTH}$ and $T_{RR} \times 225,625 = T_{RRH}$ are processed by the client mobile devices, which are illustrated in Figure 1. The $T_{QT}$ and $T_{RR}$ denote the time intervals that are measured and recorded during any electrocardiogram test. Thus, $T_{QT}$ describes the time that is required for the ventricular depolarization and repolarization, and $T_{RR}$ quantifies the variability concerning the timing of the heartbeats. The subscript $H$ suggests the homomorphic nature of the comparison, which is required in order to detect the presence of the DRHS condition. It is relevant to note that a comprehensive set of calibration tests has been implemented, which was used in order to fine-tune Equation (1). Thus, the equation is improved concerning the accuracy of the detection results, and also regarding the efficient usage of the computational resources. The equation ensures that the integrated data management system precisely detects this problematic medical condition with no false positives. This is accomplished using just the absolutely required FHE operations. The data processing components aggregate the results of the particular comparison operations. The data processing flow implies that the mobile client devices send a request to the data processing components, which essentially asks for a medical report relative to a certain period of time. The mobile client device decrypts the results that are received and checks whether at least one bit is equal to 1. If at least one such bit is found, then it is immediate to infer that the comparison $T_{QTH} > T_{RRH}$ was true at least once. As a consequence, it can be stated that the DRHS cardiac malfunction occurred, with a high probability, at least once.

### 4.5. Determination of the Minimum and Maximum Heart Rates

The determination of minimum and maximum heartbeat rates represents a functional feature of the integrated data management system. This is specified through the consideration of the $f_c(.)$ function, which is described in Figure 2. This function converts the
input data in a binary manner. This is efficiently computed by the relevant components of the integrated data management system. The relevant considerations that have been made in Sections 5.1 and 5.2 suggest that two \( n \)-bit numbers are effectively compared, and they produce an output that is also coded through a proper sequence of \( n \) bits. Let us suppose that the first number is greater than the second number. Consequently, it is valid to assert that the output of the operation is determined by one bit of 1, and \( n-1 \) bits that are 0. Additionally, it is relevant to note that the result is composed only of bits that are 0 if the first number is less than the second number. The integrated data management system preserves the validity of the processed data through the usage of the basic rotate and select operations. In essence, the result of this algorithmic model is specified through \( n \) bits that represent only values of 1.

The quantitative assessment of the minimum and the maximum cardiac rate values essentially calculates the minimum and maximum values of \( N_{CT} \) ciphertexts. This algorithmic and computational model encrypts an array of messages, which are composed of \( n \) bits. Therefore, the computational process considers the following functions: \( \min(f_c(.)) \) and \( \max(f_c(.)) \). Here, the value of the level \( L \), which supports the data processing flow of the fully homomorphic encryption routines, is computed through the following formula that ensures the continued calibration of this fundamental parameter.

\[
L > (\log_2 n + 2) \times \log_2 N_{CT}
\]

5. Real World System Performance

5.1. Considerations Regarding the System Architecture

The architecture of the integrated data management system is described in Figure 1. The system is compatible with any type of mobile data acquisition device, with the condition that it is technically suitable and proper for the data collection process. The structural stability of the system, which is demonstrated in Figure 1, is determined by the invariability of the data processing components.

The software component that is installed on the mobile client devices sends the acquired data to the data processing components considering a real time pattern. If the 5G data connection is not available, then the acquired data are cached locally. The locally stored data are transferred to the data processing components through a secure 5G data channel in a homomorphically encrypted format, as soon as the 5G data transfer channel becomes available.

It can be stated that several cardiac sensors have been assessed. Thus, it was determined that the Polar H10 Heart Rate Sensor produced the most precise results [51]. Thus, it has been selected as the user-side data collection device in order to assess the field trial deployment of the integrated data management system. The personal health information data, which are necessary to evaluate the system’s capacity to detect the delayed repolarization of the heart syndrome (DRHS), are supplied by a medical dataset that stores 750 patients. The Polar H10 sensor has been used by all the individuals that are enrolled into the system during the field trial.

The system architecture relates to the usage of certain software and hardware components. The cardiac data are acquired by the Polar H10 personal sensor. The data that are gathered are sent to each patient’s Android smartphone, which uses the allocated secure 5G data channel in order to communicate with the data processing components. The integrated data processing system’s client component is installed on the patient’s smartphone. It collects the data, which are effectively acquired by the Polar H10 sensor. Furthermore, the data are encrypted, and they are transmitted to the data processing components, which are deployed on the IBM Cloud infrastructure.

The central data processing components’ algorithmic core is designed and implemented using an improved version of the algorithmic model that is presented in [31]. This version includes the enhancements that have been presented in the previous section. The central data processing components are installed on the IBM Cloud platform through
an adequate buildpack. The Apache Spark engine is considered in order to enhance the data access layer. The data that are acquired by the client software components are securely stored, in a fully homomorphically encrypted format, through the utilization of the IBM Cloudant platform. This is a non-relational database engine, which has been found as suitable for the storage of the collected and processed medical data. The newly collected personal health information data are detected by the Apache OpenWhisk programming service. As a consequence, the proper event handlers are called. This ensures that the newly imported data are automatically processed and safely stored. The data processing is conducted by the data processing components using the algorithmic structure and data processing pathways, which have been described.

5.2. Performance Metrics

The real-world evaluation of the system is made through the usage of four relevant metrics. The first one is the network capacity that is necessary in order to send the data between the client mobile devices and the data processing components, considering both directions. This metric is relevant considering the large amount of data that are generated by the fully homomorphic encryption modules. Thus, the $XFER_{IN}$ determines the data that are sent from the mobile client data collection devices to the data processing components, while the $XFER_{OUT}$ represents the data that are transferred from the data processing components to the user-side mobile devices.

The second performance metric is related to the load that is placed on the 5G data channel, relative to the entire capacity of this data link. Let us designate this metric with $L_{SG}$.

The third metric is determined by the storage ratio ($S_R$). This quantifies the storage capacity that is required in order to persist one byte of plain text data using the fully homomorphic encryption. Thus, if $S_R = 1000$, then it is immediate to state that considering one byte of plain text data, 1000 bytes are necessary in order to store the respective byte in the fully homomorphically encrypted format.

The fourth metric is defined by the processing speed ($P_S$). This metric is determined by the following expression.

$$P_S = \frac{P_{TO}}{P_{IN}}$$

Relative to this mathematical expression, the numerator designates the time that is necessary to transmit the data from the client devices to the data processing components. The denominator represents the time that is necessary for the backend components to process the data that are received.

5.3. Outcomes of the Performance Evaluation

The evaluation considers the data that are collected from the enrolled 750 patients using the Polar H10 cardiac sensors. The field trial has lasted for a period of two months.

The actual states of the presented metrics are described in Table 1. The table columns present, in this order, the load of the 5G data channel, the number of ciphertexts, the level $L$, the data that are received and sent by the data processing components in Gigabytes (GB), the storage ratio, and the processing speed. The values of the performance parameters scale efficiently with the size of the input data, and it is more efficient than similar reported contributions, such as the one that is presented in [31]. The integrated data management system that is presented in this paper essentially differs from existing similar approaches, considering that it offers a unified platform for the collection, transport, processing, and storage of the medical data in a fully private manner. Moreover, it can be stated that the system is scalable, as it can be observed in Table 1. Furthermore, the values of the main performance parameters, $L_{SG}$, $N_{CT}$, the level $L$, $XFER_{IN}$, and $XFER_{OUT}$, are kept at lower values, which further demonstrates the efficient behavior of the system.
Table 1. Values of performance metrics (lower is better).

<table>
<thead>
<tr>
<th>Data Reading Interval</th>
<th>( L_{SG} )</th>
<th>( N_{CT} )</th>
<th>Level ( L )</th>
<th>( XFER_{IN} )</th>
<th>( XFER_{OUT} )</th>
<th>( S_R )</th>
<th>( P_S )</th>
</tr>
</thead>
<tbody>
<tr>
<td>One minute</td>
<td>0.01</td>
<td>2</td>
<td>10</td>
<td>5.3</td>
<td>3201.3</td>
<td>32.1</td>
<td>0.54</td>
</tr>
<tr>
<td>Five minutes</td>
<td>0.07</td>
<td>12</td>
<td>12</td>
<td>6.4</td>
<td>1298.8</td>
<td>39.4</td>
<td>0.24</td>
</tr>
<tr>
<td>Fifteen minutes</td>
<td>0.19</td>
<td>40</td>
<td>15</td>
<td>6.9</td>
<td>669.2</td>
<td>47.5</td>
<td>0.23</td>
</tr>
<tr>
<td>Thirty minutes</td>
<td>0.33</td>
<td>44</td>
<td>16</td>
<td>10.6</td>
<td>1102.6</td>
<td>88.3</td>
<td>0.36</td>
</tr>
<tr>
<td>One hour</td>
<td>0.39</td>
<td>86</td>
<td>18</td>
<td>8.1</td>
<td>643.7</td>
<td>91.4</td>
<td>0.35</td>
</tr>
<tr>
<td>Three hours</td>
<td>0.52</td>
<td>258</td>
<td>20</td>
<td>9.6</td>
<td>221.9</td>
<td>101.2</td>
<td>0.37</td>
</tr>
<tr>
<td>Six hours</td>
<td>0.63</td>
<td>519</td>
<td>21</td>
<td>11.6</td>
<td>108.8</td>
<td>108.5</td>
<td>0.36</td>
</tr>
<tr>
<td>Twelve hours</td>
<td>0.72</td>
<td>1021</td>
<td>23</td>
<td>12.1</td>
<td>45.9</td>
<td>117.4</td>
<td>0.39</td>
</tr>
<tr>
<td>One day</td>
<td>0.81</td>
<td>2099</td>
<td>25</td>
<td>15.2</td>
<td>26.4</td>
<td>128.1</td>
<td>0.42</td>
</tr>
</tbody>
</table>

5.4. Comparative Performance Evaluation

We have conducted a comparative study regarding the real-world performance of our algorithmic model relative to the reference Brakerski–Gentry–Vaikuntanathan (BGV) fully homomorphic encryption model. Thus, two instances of the system have been implemented and deployed under identical software and hardware conditions through the consideration of the general system architecture that is presented in Figures 1 and 2.

The value of the performance metrics that are presented in Table 2 prove that our algorithmic variant performs better than the reference BGV model. This is particularly important, as it ensures a superior level of scalability considering the inherently large amounts of transferred and processed data.

Table 2. Values of performance metrics for the BGV variant (lower is better).

<table>
<thead>
<tr>
<th>Data Reading Interval</th>
<th>( L_{SG} )</th>
<th>( N_{CT} )</th>
<th>Level ( L )</th>
<th>( XFER_{IN} )</th>
<th>( XFER_{OUT} )</th>
<th>( S_R )</th>
<th>( P_S )</th>
</tr>
</thead>
<tbody>
<tr>
<td>One minute</td>
<td>0.02</td>
<td>2</td>
<td>11</td>
<td>6.42</td>
<td>4104.3</td>
<td>32.1</td>
<td>0.79</td>
</tr>
<tr>
<td>Five minutes</td>
<td>0.09</td>
<td>12</td>
<td>14</td>
<td>9.47</td>
<td>1681.8</td>
<td>39.4</td>
<td>0.41</td>
</tr>
<tr>
<td>Fifteen minutes</td>
<td>0.22</td>
<td>40</td>
<td>18</td>
<td>8.1</td>
<td>865.2</td>
<td>47.5</td>
<td>0.32</td>
</tr>
<tr>
<td>Thirty minutes</td>
<td>0.36</td>
<td>44</td>
<td>21</td>
<td>11.79</td>
<td>1602.9</td>
<td>88.3</td>
<td>0.43</td>
</tr>
<tr>
<td>One hour</td>
<td>0.42</td>
<td>86</td>
<td>24</td>
<td>9.85</td>
<td>814.8</td>
<td>91.4</td>
<td>0.41</td>
</tr>
<tr>
<td>Three hours</td>
<td>0.55</td>
<td>258</td>
<td>27</td>
<td>10.87</td>
<td>314.8</td>
<td>101.2</td>
<td>0.44</td>
</tr>
<tr>
<td>Six hours</td>
<td>0.67</td>
<td>519</td>
<td>31</td>
<td>12.9</td>
<td>198.9</td>
<td>108.5</td>
<td>0.39</td>
</tr>
<tr>
<td>Twelve hours</td>
<td>0.76</td>
<td>1021</td>
<td>35</td>
<td>13.84</td>
<td>87.9</td>
<td>117.4</td>
<td>0.45</td>
</tr>
<tr>
<td>One day</td>
<td>0.86</td>
<td>2099</td>
<td>39</td>
<td>16.86</td>
<td>26.4</td>
<td>208.1</td>
<td>0.46</td>
</tr>
</tbody>
</table>
5.5. Analytical Discussion Regarding Similar Contributions

The theoretical and practical value of the contribution that is reported in this paper may be better understood by the interested reader through an analytical analysis relative to relevant existing contributions. Thus, the following paragraphs discuss on this problematic.

The contribution that is reported in [52] focuses on an extensive review of current and existing approaches and mechanisms that are used in order to handle security and privacy related matters relative to e-Health software systems. Thus, strengths and weaknesses of some of these approaches are enumerated. Reviewed articles were narrowed down to the respective reported scientific contributions because of similarity observed in the models adopted by some researchers. Additionally, the authors provided an acceptable and standard definition regarding the general concept of e-Health system. Furthermore, a classification of cloud-based models was accomplished, and the relevant security and privacy requirements, as recommended by the Health Insurance Portability and Accountability Act (HIPAA) [53], were also discussed and analyzed. The authors proposed a secured and dependable architecture, which is suitable for electronic health scenarios that could guarantee efficiency, reliability, and a properly regulated access framework to health information. Nevertheless, the proposed architecture does not ensure the distributed nature of the system and its required scalability, while the encryption mechanisms are based on standard asymmetric encryption models that do not ensure the required level of health data privacy.

The paradigm of cloud-based healthcare computing has changed the face of healthcare in many ways. The main advantages of cloud computing in healthcare are represented by the scalability of the required services, and the possibility to upscale or downsize the data storage, or the required computational resources. There are papers that examine various research studies, which assess the relevant aspects that relate to the mandatory specification and implementation of the relevant security and data privacy preserving mechanisms. In this respect, there are various significant legal and technological aspects that should be analyzed. Thus, the authors of [54] analyzed a series of scientific contributions that lack, at least, some of the technical features that clearly distinguish them from the integrated data management system, such as the end-to-end data privacy mechanisms, the consistent scalability, and the possibility to accommodate various technical platforms and frameworks regarding the client and backend components.

Considering that information and communication technology has advanced towards an improved economical environment, which provides enhanced services to consumers and business actors, it is relevant to note that the health sector also benefits from these theoretical and practical advancements. Despite the visible and significant benefits that a cloud-based system deployment provides, there are still security and privacy challenges that are preventing the full array of benefits from being considered. Thus, the authors of [55] described a distributed system, which provides different levels and models of encryption relative to the various distributed software modules. This heterogeneity determines multiple administrative, functional, and security issues, which make the reported model unsuitable for the real-time implementation of large-scale medical data processing systems.

The emergence of Internet of Things (IoT) as a theoretically and practically relevant paradigm, and also the sustained development of cloud computing technologies, the design and development of electronic health systems are perceived as significant and active domains of scientific research, which enable the development of medical practices in a convenient and economical way. In this context, it can be stated that the authors of [56] discussed about an important problematic that pertains to the fully secure preservation of the personal data privacy. The paper presented an access control model for cloud-based data, which uses a certificate-based mechanism. The fundamental features of the described model are represented by the integration of the relevant trust-related mechanisms with the proper data monitoring models, which may provide a superior level of security relative to the access control mechanisms. The authors explained the methodology of the proposed approach through experimental evaluation results, which apparently demonstrate an
improvement of the system’s security and performance, through the minimization of the time that is spent in order to define and implement the permissions that are necessary to access the relevant services, and also through the optimization of the overall system resources utilization. Although the algorithmic and data processing structures are designed in a more uniform manner, as compared to similar reviewed approaches, the reported model is still unable to provide the required scalability, and also the mandatory end-to-end medical data privacy between the client devices, and the backend data processing components. This also favorably differentiates the integrated medical data processing system, which we describe in this paper.

Cloud computing in healthcare has witnessed a major development in recent years due to its remote access capabilities, among other factors. The reviewed scientific studies have shown that it has attracted significant attention in the field of healthcare. Nevertheless, the surveyed research papers demonstrate that a relatively high number of healthcare consumers are yet to accept the technology, especially in developing countries due to reasons, such as the data security and the improper and unfriendly utilization of this approach relative to the end users, with limited or no technical skills [57]. This is another significant aspect, which relates to the realm of user experience, that is consistently approached by the integrated data management system, which we report in this paper. Thus, the system is compatible with the most affordable mobile devices that feature a data connection. This also allows for the medical data collection to occur in a seamless manner for the enrolled patients, without any costly changes that would be required to their personal mobile devices assets.

Considering the Attribute-Based Encryption (ABE) schemes, patients encrypt their electronic health record (EHR), attach the proper attributes, and transmit them over to the cloud. Doctors and entitled medical practitioners receive the encrypted EHR, which corresponds to their area of interest, from the cloud-based systems. The decryption of the received encrypted EHR involves that the medical practitioners receive the secret keys from the key generation center (KGC). Considering that the KGC knows the secret keys of all the encrypted EHR records, it may consequently decrypt the patients’ records. A decentralized ABE scheme overcomes this issue, but it requires high computation and communication costs. Moreover, considering a scheme with such an architecture, any unauthorized doctor may be able to access the patients’ private EHR data. Moreover, the KGC’s secret keys privacy and the doctor’s attribute privacy also represent serious concerns. The authors of [58] described a privacy-preserving e-Health (CP2EH) scheme over the cloud that overcomes the problems of both unauthorized access of patient records by a doctor, and a doctor’s attribute privacy in an ABE scheme. In the context of this CP2EH scheme, it is relevant to mention the incorporation of oblivious transfer (OT) and zero-knowledge proof (ZKP) protocols into the centralized ABE scheme. The OT protocol maintains the secret keys’ privacy and the doctor’s attribute privacy. Nevertheless, the described system proves to be rigid concerning the accepted data collection devices. Additionally, it is compatible with only certain software frameworks, it does not scale well relative to an increased number of enrolled patients, and it does not implement end-to-end data privacy mechanisms. In contrast, the integrated data management system addresses all of these shortcomings.

The authors of [59] described a novel attribute-based encryption (ABE) based on an access control scheme, which may impose multi-level and controlled access delegation. Furthermore, it is assessed how such a system may be deployed in an e-Health environment, in order to securely share the outsourced EHR data of the enrolled patients. Furthermore, the authors inferred that the proposed scheme is secure against chosen plaintext attacks, as well as attacks mounted via attribute collusion [60]. Nevertheless, even if this seems to be one of the most promising approaches, which have been reviewed, it still does not design an end-to-end privacy-preserving medical data processing pipeline. Additionally, it still suffers from the essential architectural and functional shortcomings that have already been mentioned.
The comprehensive literature review that was conducted demonstrates that although interesting contributions are reported in the relevant scientific literature, virtually all of the existing algorithmic and functional models lack some of the fundamental technical features, which are necessary in order to implement a real-time, scalable, and completely private integrated medical data management system. Consequently, it can be asserted that the integrated medical data management system, which is reported in this paper, is one of the few that fulfills all of the mandatory algorithmic and technical constraints. This clearly differentiates it from most of the existing systems, which fully warrants the value of the contribution that is reported in this paper.

6. Conclusions and Planned Developments

The efficient acquisition of personal health information data has become progressively relevant during the past fifteen years as a consequence of the continued evolution of personal mobile devices and medical sensors. It has become feasible to gather the personal health information data through a minimally obtrusive application of proper mobile sensors and devices. The collected data, which can be assimilated to the realm of big data, imply administrative and legal aspects. The administrative aspect mostly pertains to the extraction of relevant medical knowledge, while the legal aspect is connected to the mandatory constraint to observe the full preservation of the personal health information data privacy during its entire lifecycle.

This article presents an integrated personal health information data management system, which implements all the necessary constraints. It is compatible with the vast majority of the current and, with a substantial probability, future client-side mobile data collection devices. The system’s data transportation channels are implemented using secure 5G data channels. The validity and efficiency of the system are evaluated using a comprehensive field trial that considers 750 enrolled participants. Thus, it is demonstrated that the deployed integrated system is able to efficiently and scalably accommodate the involved fully homomorphic encryption data processing tasks. This is a relevant achievement, considering that it is one of the few existing approaches that implements a fully functional integrated personal health information data management system, which specifies complete data privacy mechanisms considering all stages of the data management process: acquisition, transportation, processing, and storage. This contribution has also the merit to describe a system, which uses virtualized secure 5G data channels, which add an additional layer of security relative to the fully homomorphically encrypted data management routines. It is interesting to note that this data security model can be applied to other relevant use cases, such as the organization of chess tournaments.

The architecture and logical specification of the virtualized 5G data infrastructure will be improved, together with the data processing components that are based on the fully homomorphic encryption routines. This planned optimization effort has the role to relieve the load on the computational resources that are necessary in order to transfer, store, and process the collected personal health information data. The thorough field trial demonstrates that the system is capable to manage the intended use case scenarios. Consequently, the necessary planning has been made for its continued development and maintenance.

Author Contributions: Conceptualization, R.B.; methodology, R.B., A.V. and D.M.D.I.; software, R.B.; validation, R.B., A.V. and D.M.D.I.; formal analysis, A.V. and D.M.D.I.; investigation, R.B.; resources, R.B., A.V. and D.M.D.I.; data curation, R.B.; writing, R.B.; writing—review and editing, R.B., A.V. and D.M.D.I.; supervision, R.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research work received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.
Acknowledgments: The research work that is reported in this paper has been conducted using the IBM Cloud infrastructure.

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

References


34. Thakur, N.; Han, Chia Y. An Ambient Intelligence-Based Human Behavior Monitoring Framework for Ubiquitous Environments. Information 2021, 12, 81. [CrossRef]


57. Idoga, P.E.; Toycan, M.; Nadiri, H.; Çelebi, E. Factors affecting the successful adoption of e-health cloud based health system from healthcare consumers’ perspective. IEEE Access 2018, 6, 71216–71228. [CrossRef]

