Multi-Layer Privacy-Preserving (MPP) Mechanism for Protected Health Information (PHI) In A Health Campaign Management System

Syarulnaziah Anawar[1†] Muhammad Hafiz Jamil[1†], Zakiah Ayop[1†], Nurfadzilah Othman [1†], Norharyati Harum[1†], Erman Hamid[1†], and Suzana Zambri[2†]

[1†]Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka, Malaysia
[2†]Faculty of Computer and Mathematical Sciences , Universiti Teknologi MARA Malaysia

Summary
Health campaigns can be an excellent method for health care systems to empower people and communities to a healthy lifestyle. At the same time, it can also reduce non-communicable disease cases. A health campaign management system involves various parties, such as campaign participants, campaign managers, health professionals, and stakeholders. The main drawback of a health campaign management system is the difficulty of the organizer to manage data for campaign participants in large groups. Specifically for data privacy, the organizer does not have a guarantee to protect the participant information, particularly protected health information (PHI), according to the different types of users in the system. Therefore, this paper proposes multi-layer privacy-preserving (MPP) mechanism to protect PHI data in a health campaign management system. The proposed MPP mechanism integrates multiple models of access control, which are role-based access control (RBAC) and attribute-based access control (ABAC) with dynamic data masking architecture. The integration of ABAC, RBAC, and data masking will overcome the difficulty of the dynamic data masking architecture in separating the authorized users to provide full access to the real data. The proposed MPP mechanism will provide a holistic privacy-preserving mechanism for the health campaign management system.

Key words:
Privacy; Data masking; Access control; Health system; RBAC; ABAC

1. Introduction

According to the 2019 National Health and Morbidity Survey (NHMS), about 3.9 million people in Malaysia aged 18 years and above have diabetes. This includes 6.4 million people in Malaysia who have hypertension. Both diseases are categorized as non-communicable diseases (NCD). The government has taken the initiative to conduct various health campaigns to reduce NCD cases in Malaysia through a 10-year National Strategic Plan for Non-Communicable Diseases (NSP-NCD) 2016-2025. The health campaign is to empower people and communities to a healthy lifestyle and at the same time, reduce NCD cases in Malaysia.

For an individual effort to increase self-care to prevent NCD, many apps can be used to monitor and track an individual's health progress. The individual also can increase their health awareness, because some of these apps can calculate food calories, monitor blood pressure, and have much more functionalities. In recent years, there have also been health campaigns created by organizations that offer their services to the public in Malaysia. Services that they usually offer are healthy meal tips and weight-loss programs. The health campaign is usually communicated via social networking sites (SNS), such as Facebook, Instagram, and WhatsApp. By using the social networking site (SNS) as a platform, a participant that participates in the campaign will join a private group created by the organizer. In this group, the organizer will manage the participants. The organizers also can post any updates in this group to disseminate information to the participants. This SNS platform also has notification features. It can alert participants when the organizer posts an update in that group.

The main drawback of a health campaign in the SNS platform is the difficulty of the organizers to manage the data of campaign participants in large groups. Specifically for data privacy, the organizer does not have a guarantee to protect the participant's information, particularly the protected health information (PHI), according to different types of users in the system. A health campaign management system involves various parties, such as campaign participant, campaign manager, health professionals, and stakeholders. PHI is the health data that is generated and provided in connection with healthcare provision, healthcare operations, and healthcare payment services under the Health Insurance Portability and Accountability Act of 1996 [1], a statute in the United States. However, in Malaysia, data protection under the Personal Data Protection Act 2010 (PDPA) does not specifically identify health information to be protected in health entities and business associates. Under the PDPA, health data can only be processed by health professionals.

To date, works in privacy-preserving mechanism in health systems that combine both access control and dynamic data masking are scarce, particularly in the context of health systems [2]. In a health campaign management system, specifying the role to users is important to give different authorities to implement resource access control.

Manuscript received October 5, 2021
Manuscript revised October 20, 2021
https://doi.org/10.22937/IJCSNS.2021.21.10.29

Therefore the novelty of the work in this paper is the integration of role-based access control (RBAC) and attribute-based access control (ABAC) with dynamic data masking to protect protected health Information (PHI) data. To the best of our knowledge, our multi-layer privacy-preserving (MPP) mechanism is the first work that integrates multiple types of access controls, which are RBAC and ABAC with dynamic data masking architecture. The closest study is the work done by Peng et al. [12], which combines attribute-based access control (ABAC) with data masking. The MPP mechanism will provide a holistic privacy-preserving mechanism for the health campaign management system.

This paper begins with Section 1, which explains the motivation of the study and reviews some related work. In Section 2, the overview in data privacy techniques, PHI, and related work on data privacy in a health system is presented. In Section 3 describes the prototyping methodology used in this study. In Section 3 and 4, the design of the health campaign management system and the multi-layer privacy-preserving (MPP) mechanism is presented. In Section 6, we present the proof of concept implementation of our proposed mechanism. This paper is summarized in the last section.

2. Literature Review

2.1 Overview in Data Privacy Techniques

According to Abouelmehdi et al. [3], there are four types of data privacy techniques, namely encryption, authentication, data masking, and access control.

Encryption: Encryption is an efficient way of preventing unauthorized access to sensitive information. It also protects and maintains the authenticity of the information when a user is using the system [2]. For example, when an unauthorized person accesses the encrypted information, it cannot gain any useful information. Some of the commonly used encryption techniques are advanced encryption standard (AES), which encrypts fixed data blocks (of 128 bits) at one time, and Rivest-Shamir-Adleman (RSA), which is based on the factorization of the product of two large prime numbers.

Authentication: Authentication is the act of confirming real and authentic credentials made by or concerning the subject [3]. Two types of authentications are user authentication and data authentication [3]. User authentication can be defined as how users prove their authenticity to the system, while data authentication is used to authenticate the message sender and to ensure that all sensitive message information in transit has not been modified [4]. A commonly used information for user authentication is username or identity (ID) combined with password.

Data masking: Data masking works by replacing the sensitive data with a random value that looks like the original data and after masking, it cannot return to the original value [5]. This technique should preserve the information from leaking to public or unauthorized entities. If the masked information is leaked to an unauthorized person, it will not lead to the information owner. There are three categories for data masking, namely pseudonymization [3], anonymization, and de-identification [6]. Pseudonymization transforms and replaces the original data with a pseudonym that is impossible to relate with the original information without additional information [3a]. In the anonymization technique, information that is personally identifiable to the individual will be modified or removed.

Access Control: [3] states that an access control policy is a user access management and monitoring that can significantly help to ensure the confidentiality and integrity of protected health information. After a user is authenticated, the user can access the health system, but their access will be controlled according to the access control policy. This policy is usually developed based on the privilege and right of each system user, and must be authorized by the data owner [3].

2.2 Protected Health Information (PHI)

Protected health information (PHI) is defined as "individually identifiable health information that is transmitted or maintained in any form or medium by a covered entity or its business associate" [7]. PHI is used in the health system because health information is both sensitive and legally covered. Therefore, caution must be taken if such information is transmitted or delivered electronically. Consequences that can happen if the health information is leaked to an unauthorized person or hacker may include the hacker selling the information illegally and making a profit. The buyer of the information may use the information to create fake IDs to purchase medical equipment or drugs or even to file a false insurance claim.

Nearly all developed countries have laws that set standards for privacy and security to ensure the protection of health information [8]. In Malaysia, data protection is defined under the Personal Data Protection Act 2010 (PDPA). However, the act does not specifically identify health information to be protected in health entities and business associates. Figure 1 provides the data that have been included in PHI. There are 18 data attributes for PHI [7].

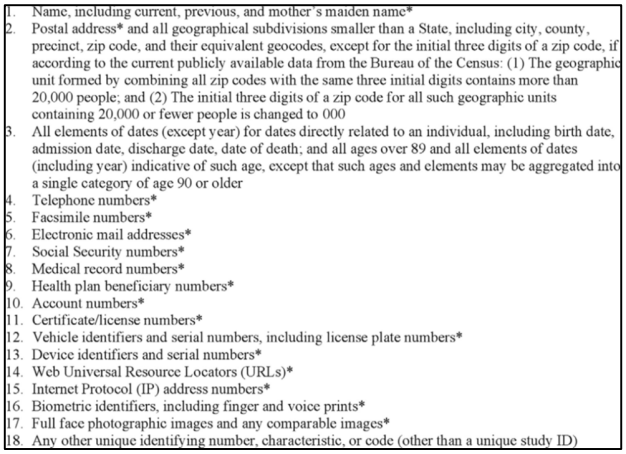
1. Name, including current, previous, and mother's maiden name*
2. Postal address* and all geographical subdivisions smaller than a State, including city, county, precinct, zip code, and their equivalent geocodes, except for the initial three digits of a zip code, if according to the current publicly available data from the Bureau of the Census: (1) The geographic unit formed by combining all zip codes with the same three initial digits contains more than 20,000 people; and (2) The initial three digits of a zip code for all such geographic units containing 20,000 or fewer people is changed to 000
3. All elements of dates (except year) for dates directly related to an individual, including birth date, admission date, discharge date, date of death; and all ages over 89 and all elements of dates (including year) indicative of such age, except that such ages and elements may be aggregated into a single category of age 90 or older
4. Telephone numbers*
5. Facsimile numbers*
6. Electronic mail addresses*
7. Social Security numbers*
8. Medical record numbers*
9. Health plan beneficiary numbers*
10. Account numbers*
11. Certificate/license numbers*
12. Vehicle identifiers and serial numbers, including license plate numbers*
13. Device identifiers and serial numbers*
14. Web Universal Resource Locators (URLs)*
15. Internet Protocol (IP) address numbers*
16. Biometric identifiers, including finger and voice prints*
17. Full face photographic images and any comparable images*
18. Any other unique identifying number, characteristic, or code (other than a unique study ID)

Fig. 1 Identifiers considered HIPAA protected health information [7]

2.3 Related Work On Privacy Mechanism in Health Systems

Previously, several studies have proposed various methods for privacy protection using a single or combination of the previously mentioned data privacy techniques. In the context of authentication, the authenticated key agreement is usually used to establish cryptographic keys between the authorized user and the server to safeguard data confidentiality and integrity in a health system. Zhang [10] proposed a dynamic biometric authentication three-factor key agreement protocol at the server side that preserves user privacy, because the transmitted message is untraceable. Besides biometric, a physiological-feature-based key agreement could be used to preserve data privacy in wearable health devices, as demonstrated by Tang [11].

One of the important research areas for privacy protection mechanism in health systems is defining the suitable solution to protect PHI according to different types of users in the system. Several solutions have been proposed to address different types of users in the system of user access control. Some of the commonly used access-control models for health systems are role-based access control (RBAC), attribute-based access control (ABAC), and capability-based access control (CapBAC). There has been a significant interest in the use of RBAC for access management in health systems through user roles. Despite its popularity, RBAC is not without limitations in which access permissions can be assigned only to user roles, thus making it inflexible when they are used alone. Therefore, RBAC is usually extended or combined with other models to provide flexibility in access control. For instance, recently Bouadjemi and Abdi [12] have extended the core RBAC model by incorporating obligation to manage exception situations. In the notion of obligation, permission can be assigned dynamically to extend role permissions in exceptional situations, whereby the user may be associated with the same obligation or different obligations. Pal et al. [13,14] have designed a policy-based access control policy for IoT-based healthcare, which combines all ABAC, RBAC, and CapBAC models. The proposed policy provides authorized users with access to specific services by utilizing attributes for role management, capabilities for access right implementation, and attributes for fine-grained policy decisions based on capabilities.

Data masking is one of the emerging techniques in privacy-preserving mechanism in health systems. Ali and Ouda [15] have reviewed some irreversible and key-based reversible data masking techniques that can be used by a health system, which include substitution, masking out, shuffling, hashing, pseudonymization, and modulus-based masking. There are two types of architectures for data masking, namely dynamic data masking and static data masking. However, the former architecture is not widely implemented in most of the papers that employ data masking in health systems [16]. This may be due to the difficulty in separating the authorized users to provide full access to the real data [17]. One of the possible solutions to combine data masking with the access-control technique in dynamic data masking. On the other hand, using solely data masking for privacy-preserving mechanism in a health system is quite risky as a patient can be re-identified after the data are masked by linking some attributes in the anonymized data to other publicly available data sources. This is demonstrated through an experiment by Mvule et al. [18] that developed a software service that utilized a set of masking algorithms to modify the PHI of a patient's records before they are released to the stakeholders.

The closest benchmark for the privacy mechanism used in this study is the work done by Peng et al. [19]. In their study, Peng et al. combined attribute-based access control (ABAC) with data masking. They proposed a differential attribute desensitization system (DADS) method, whereby the data owner can define the level of sensitive attributes, and later DADS will automatically identify and mask different types of sensitive information. However, their work does not explore a variety of data masking techniques that are suitable to mask protected health information (PHI) under dynamic data masking architecture.

In short, various efforts have been made to overcome privacy issues in health system implementation, but the current solutions still have weaknesses that need to be addressed to provide a holistic privacy-preserving mechanism for PHI in a health system.

3. Methodology

The health campaign management system is developed following the prototyping model, which consists of six phases: requirement gathering, quick design, building

prototype, customer evaluation, refining prototype, and engineer product. The activity flow of each phase is illustrated in Figure 2. This paper only discusses activities up to phase 3. The activities in each phase are described in Table 1.

Table 1: Description of Methodology

Phase	*Activities*
Requirement Gathering	• Identify required hardware and software, such as Visual Studio Code, Web Browser, Laragon, DBeaver, Desktop, Smartphone, Cloud server). • Identify modules/functions that need to be embedded in the prototype
Quick Design	• Determine the design of the prototype based on hardware, software, modules, and function of the health campaign system. The discussion is elaborated in Section 4. • Determine the design of multi-layer privacy-preserving mechanism based on the four privacy techniques, namely password encryption, user authentication, user access level, and data masking. The user access level is designed using a combination of role-based access control (RBAC) and attribute-based access control (ABAC) approach. The discussion is elaborated in Section 5.
Building Prototype	• Build a proof-of-concept prototype based on determining design • The Multi-layer Privacy-Preserving Mechanism testing is explained in Section 6.

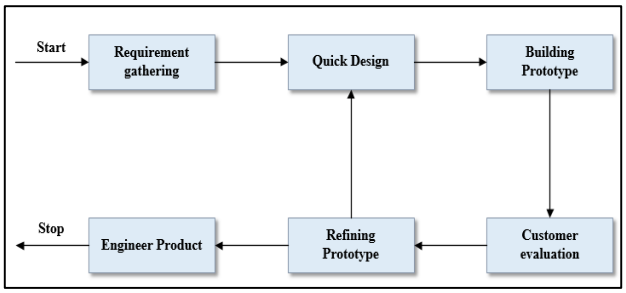

Fig. 2 Prototyping Model

4. Design of Health Campaign Management System

4.1 System Architecture

The health campaign management system is part of our mHealth participatory sensing system, myCommHealth [20] which is a work in progress that aims to implement IoT cloud-based mHealth participatory sensing that incorporates privacy-preserving mechanism and healthcare analytics. myCommHealth allows campaign organizers to recruit participants and monitor participants' health-related activities, and coaching. The system will use the cloud as an open platform that allows communities and stakeholders to collect, analyze, and share health information.

Our health campaign management system is designed as an effective and integrated health monitoring campaign management system that can help an organization efficiently manage and conduct health monitoring campaigns. Using this system, any organization, hospital, or other medical institution, etc., who wishes to be a health campaign organizer can register the health campaign information into this system. Figure 5 shows our system architecture and information flow. The system can be accessed from any computer connected to the Internet using a standard browser for a web-based application, or from a mobile phone for better portability, ubiquity, and connectivity.

Figure 3 shows the system architecture for our health campaign management system. For this project, the LEMP stack has been used to set up the system environment, which consists of Nginx, PHP, MySQL, and Linux. The system will be installed on Linux Debian 10 and connected to the API, which acts as a back end to the system. On this project, the package that will be used for the front end is Laravel-AdminLTE. The system will request the API when the user runs any module, and the API will authenticate the request from the valid user, or not. The API request will then be forwarded to ProxySQL, which acts like a proxy server. If the proxy server detects this request from an unauthenticated user, the proxy server will forward the request to the MySQL database through port 4006 and MySQL will process the request as an unauthenticated user. The result returned by MySQL would mask sensitive information that was configured on the proxy server.

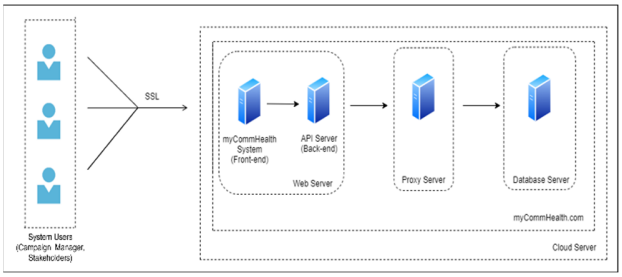

Fig. 3 System Architecture

4.2 System Design and Functionality

The system offers several features for its users (participant and campaign manager), which include campaign setup, participant management, and fee payment. The system decomposition diagram is shown in Figure 4.

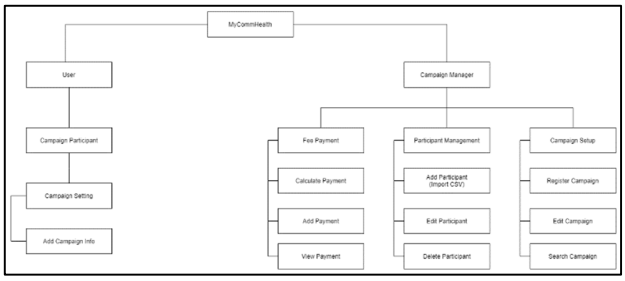

Fig. 4 Decomposition Diagram

In this paper, the discussion of the system design and functionality is limited to the module under the campaign manager only:

Campaign management: There are four tasks that a campaign manager can accomplish in Campaign Setup. Users can register, edit, delete, and search for a campaign. For campaign registration, the campaign manager must fill in the campaign details, such as campaign name, campaign description, and campaign duration. Once completed, the user will be redirected to the payment module. If the payment is successful, the campaign manager can continue to manage participants and stakeholders. Lastly, the campaign manager can publish the newly created campaign.

Participant management: Under this module, the campaign manager can add new participants using three methods: import participants from previous campaigns, upload a CSV file that contains participant details, or add participants manually using the form in the system. Participants will receive an email to receive information on privacy policy and terms of use upon successful registration.

Fee Payment: For the payment module, after the campaign manager creates the campaign, they must complete the full payment before the campaign can be published. The total amount of the system fee is based on the duration of the campaign. Upon fee confirmation, the campaign manager will be redirected to a third-party secure payment gateway. All bank validations will be made through the payment gateway. A receipt will be provided as proof of payment through email.

5. Design of Multi-layer Privacy-preserving Mechanisms

The multi-layer privacy-preserving (MPP) mechanism describes the PHI data protection process to secure a health campaign management system. MPP for PHI in a health campaign management system is implemented through four steps as illustrated in Figure 5.

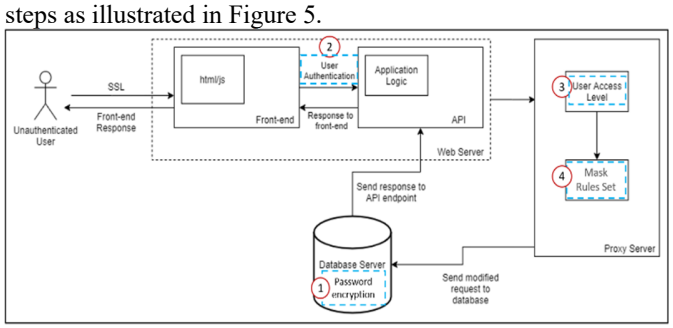

Fig. 5 Proposed Multi-layer Privacy-Preserving Mechanism

5.1 Password Encryption and User Authentication

The user will use both an email and a password as credentials to access the system. For security purposes, the user's password will be encrypted before being stored in the database, because a plaintext password will be easily read and may harm the user account if the hacker or unauthorized person gets access to the database or API endpoint. Password encryption is constructed using the AES algorithm and Laravel hash bcrypt to encrypt the password.

The second process is user authentication, whereby the authentication process will be connected to the API, which acts as a back end to the system. The health system will request the API when the user runs any module, and the API will perform the authentication. The API request will then be forwarded to ProxySQL, which acts like a proxy server. If ProxySQL detects this request from an unauthenticated user, ProxySQL will forward the request to the database through port 4006 and MySQL will process the request as the unauthenticated user. The result returned by MySQL would mask sensitive information that was configured in the mask rule.

5.2 User Access Level

Implementation of user access level on the proxy server is to verify the user permission when they request from the API to retrieve data from the database according to their

privilege. If the user is authorized to access the data from the database, specific database users that have permission to read and write will be used and forwarded to the database server. If the user is unauthenticated, it will go through the query rules set whereby the confidential data will be masked when the query has been forwarded to the database server. This implementation will be applied to the ProxySQL packages.

The access management control is implemented using a combination of role-based access control (RBAC) and attribute-based access control (ABAC) approach. Our access control precisely defines the customization to allow or deny access to the PHI based on the role inputs that it receives, and the attribute of the health campaign ID. To preserve the privacy of campaign participant records, a few access control rules are defined in this use case as follows: campaign participant records are split into four categories: (i) demographic, (ii)campaign, (iii) health, and (iv) billing, where:

a) All users are not allowed to delete any records. All data has a retention age of 3 years.

b) Health professionals are allowed to read records of campaign participants within demographic and health categories who are under the designated campaign.

c) The campaign manager is allowed to read and modify records within the demographic, campaign, and billing categories.

d) Stakeholders are only allowed to read the demographic

5.3 Data Masking

This study employs dynamic data masking architecture, where the rules of anonymizing are implemented as part of the method of transferring the data from the database to application users. The data in the database is never present in an unmasked form. To application users, the API Server is just a gateway to transfer the database request through and the application users cannot determine if they are connected directly to the database or via the API.

After the proxy server verifies the user access level, it will go through the set of rules. Here, the request will proceed as a modified query. The data type that has been listed as PHI data will be masked with specified data masking techniques. The data masking techniques used to protect the PHI are masking out and number and date variance. Masking out works by removing sensitive content from the query result while still maintaining the look and feel. For the number and date variance technique, it will change the original date to another date with the specified range that has been configured in the masking rules. This will not change the look of the data and at the same time, obfuscates the original data.

The proxy comes with features that provide enhanced security, such as encryption of data in flight and masking of sensitive end-user data. These features are suitable to be implemented in the web-based health system, because this type of system contains a generous amount of sensitive information. The overall data masking process is depicted in Figure 6. To perform data masking efficiently, we use a masking filter, where the value of a certain column in the database returned by a query can be masked with multiple data masking techniques at the same time.

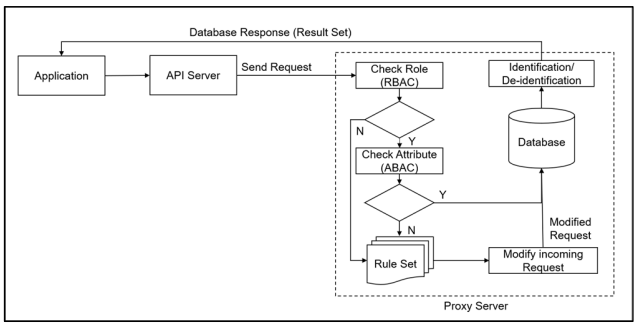

Fig. 6 Access Control and Data Masking Process Flow

Next, we create a query rule for specific queries using the ProxySQL feature, then apply a mask to the health data with appropriate data masking techniques for a different type of PHI. A query rule (see Figure 7) is created that matches the query that will be run by the client. The proxy will then use regex to verify the query from the API endpoint. Regex is a collection of letters and symbols representing a logical sequence. Figure 8 shows the query of mask rules that were applied on the proxy. The column "match_pattern" is where the regex will verify if the query is requested to the table that contains PHI data. If the query is matched with the regex pattern, it will replace them with a masking regex pattern that responds with the modified query. The column "replace_pattern" is where the regex pattern will modify the query with specified masking techniques to protect the confidential data.

```
Admin>INSERT INTO mysql_query_rules (rule_id,active,username,match_pattern,replace_pattern,apply)
 VALUES (1,1,'hafiz','^[sS][eE][lL][eE][cC][tT] (.*)p_name([ ,\n])(.*)', 'SELECT \1CONCAT(LEFT(p_nam
e,2),REPEAT('X',10)) p_name\2\3',0);_
```

Fig. 7 Query Rule

6. Proof of Concepts

The system will be installed on Linux Debian 10 and connected to the API, which acts as a back end to the system. The system will request the API when the user runs any module, and the API will authenticate the request from the valid user, or not. The API request will then be forwarded to ProxySQL, which acts like a proxy server. If the proxy server detects this request from an unauthenticated user, the proxy server will forward the request to the MySQL database through port 4006 and MySQL will process the

request as the unauthenticated user. The result returned by MySQL would mask sensitive information that was configured on the proxy server.

In the proof of concept, a database user is used for the unauthenticated user when they request the API. The result of the query will mask the confidential data and respond to the API request. Figure 8 shows a response from the API endpoint and it can be seen that the PHI information, such as email, name, and user location, have been masked successfully. On the contrary, Figure 9 shows the response from the API endpoint for an authorized user and it shows that the confidential information is in its original look.


```
{
  "p_id": 37,
  "p_password": "$2y$10$BLLE7Vh17LFBBL0mnlHsquqYNSmp7wyyV6a6pobMliTNhTQ.9DLSu",
  "p_status": "Pending",
  "p_initial_weight": null,
  "p_initial_height": null,
  "p_target weight": null,
  "p_sharing": "",
  "confirmed": 0,
  "confirmation_code": "jKSdIGj9IXEjC0bi0XCDgGHAM8NxNC”,
  "p_email": "XXXXXXXXXX",
  "p_name": "XXXXXXXXXX",
  "p_address": "XXXXXXXXXX",
  "p_postcode": "XXXXXXXXXX",
  "p_city": "XXXXXXXXXX",
  "p_state": "XXXXXXXXXX",
  "p_contact": "XXXXXXXXXX"
}
```


Fig. 8 Non-authorized Access Privacy Testing


```
{
  "p_id": 37,
  "p_password": "$2y$10$BLLE7Vh17LFBBL0mnlHsquqYNSmp7wyyV6a6pobMliTNhTQ.9DLSu",
  "p_status": "Pending",
  "p_initial_weight": null,
  "p_initial_height": null,
  "p_target weight": null,
  "p_sharing": "",
  "confirmed": 0,
  "confirmation_code": "jKSdIGj9IXEjC0bi0XCDgGHAM8NxNC”,
  "p_email": "hafiz.jamil@gmail.com",
  "p_name": "hafiz jamil",
  "p_address": "ttu",
  "p_postcode": "75450",
  "p_city": "ayer keroh",
  "p_state": "melaka",
  "p_contact": "012-3456789"
}
```


Fig. 9 Authorized Access Privacy Testing

7. Conclusion

In summary, this paper proposes multi-layer privacy-preserving (MPP) mechanism to protect protected health information (PHI) data in a health campaign management system. The proposed MPP mechanism integrate multiple models of access control, which are RBAC and ABAC with dynamic data masking architecture, that will provide a holistic privacy-preserving mechanism for the health campaign management system. In this paper, we show how the current privacy-preserving mechanism in health systems that combine both ABAC access control and data masking is extended by incorporating the RBAC model for access management by specifying the role to users to give different authority to implement a fine-grained resource access control. The integration of ABAC, RBAC, and data masking will overcome the difficulty of dynamic data masking architecture in separating the authorized users to provide full access to the real data.

Currently, the proposed MPP mechanism only implements masking out and number and date variance data masking techniques used to protect PHI, because in dynamic data masking architecture, any error in the masking process will automatically disrupt the data transfer. Therefore, future extensions of our work may explore a wider variety of data masking techniques that are suitable to mask protected health information (PHI) under dynamic data masking architecture. Exploration of other variety of data masking techniques will allow the most effective solution without degrading the performance of the health systems.

Acknowledgments

This research is funded by Prototype Research Grant Scheme, Ministry of Higher Education, Malaysia (PRGS/2018/FTMK-CACT/T00020). A high appreciation to Fakulti Teknologi Maklumat dan Komunikasi, Universiti Teknikal Malaysia Melaka (UTeM) for supporting the work done in this paper.

References

[1] L. A. Thompson, E. Black, W. P. Duff, N. P. Black, H. Saliba, & K. Dawson, *Protected health information on social networking sites: ethical and legal considerations*, Journal of medical Internet research, vol. 13(1), pp. 73-80 (2011)

[2] K.J. Hui, S. Anawar, N.F. Othman, Z. Ayop, and E. Hamid, *User privacy protection behavior and information sharing in mobile health application*, International Journal of Advanced Trends in Computer Science and Engineering, vol. 9, no. 4, pp. 5250 – 5258 (2020)

[3] K. Abouelmehdi, A. Beni-Hssane, H. Khaloufi, and M. Saadi, *Big data security and privacy in healthcare: a review*, Procedia Computer Science, vol. 113, pp. 73-80 (2017)

[4] J. L. Fernández-Alemán, I. C. Señor, P. Á. O. Lozoya, and A. Toval, *Security and privacy in electronic health records: A systematic literature review*, Journal of biomedical informatics, vol. 46, no. 3, pp. 541-562 (2013)

[5] S. Salsano, L. Veltri, and D. Papalilo, *SIP security issues: the SIP authentication procedure and its processing load*, IEEE network, vol. 16, no. 6, pp. 38-44 (2002)

[6] G. K. Ravikumar, T. N. Manjunath, R. S. Hegadi, and I. M. Umesh, *A Survey on Recent Trends, Process and*

Development in Data Masking for Testing, International Journal of Computer Science Issues (IJCSI), vol. 8, no. 2, p. 535 (2011)

[7] G. S. Nelson, *Practical implications of sharing data: a primer on data privacy, anonymization, and de-identification*, In SAS Global Forum Proceedings, pp. 1-23 (2015)

[8] D. T. Fetzer, and O. C. West, *The HIPAA privacy rule and protected health information: implications in research involving DICOM image databases*, Academic radiology, vol. 15, no. 3, pp. 390-395 (2008)

[9] B. C. Drolet, J. S. Marwaha, B. Hyatt, P. E. Blazar, and S. D. Lifchez, *Electronic communication of protected health information: privacy, security, and HIPAA compliance*, The Journal of hand surgery, vol. 42, no. 6, pp. 411-416 (2017)

[10] L. Zhang, Y. Zhang, S. Tang, and H. Luo, *Privacy protection for e-health systems by means of dynamic authentication and three-factor key agreement*, IEEE Transactions on Industrial Electronics, vol. 65(3), pp. 2795-2805 (2017)

[11] W. Tang, K. Zhang, J. Ren, Y. Zhang, and X. Shen, *Flexible and efficient authenticated key agreement scheme for BANs based on physiological features*, IEEE Transactions on Mobile Computing, vol. 18(4), pp. 845-856 (2018)

[12] A. Bouadjemi and M. K. Abdi, *Towards an Extension of RBAC Model*, International Journal of Computing and Digital Systems, vol. 10, pp. 1-11 (2020)

[13] S. Pal, M. Hitchens, V. Varadharajan, and T. Rabehaja, *On design of a fine-grained access control architecture for securing IoT-enabled smart healthcare systems*, 14th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services , pp. 432-441 (2017)

[14] S. Pal, M. Hitchens, V. Varadharajan, and T. Rabehaja, *Policy-based access control for constrained healthcare resources in the context of the Internet of Things*, Journal of Network and Computer Applications, vol. 139, pp. 57-74 (2019)

[15] O. Ali and A. Ouda, *A classification module in data masking framework for business intelligence platform in healthcare,* 2016 IEEE 7th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), pp. 1-8 (2016)

[16] K. Siddartha and G. K. Ravikumar, *Analysis of Masking Techniques to Find out Security and other Efficiency Issues in Healthcare Domain*, Third International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC), pp. 660-666 (2019)

[17] A. I. Baranchikov, A. Y. Gromov, V. S. Gurov, N. N. Grinchenko, and S. I. Babaev, *The technique of dynamic data masking in information systems*, 5th Mediterranean Conference on Embedded Computing (MECO), pp. 473-476 (2014)

[18] K. Mivule, S. Otunba, and T. Tripathy, *Implementation of data privacy and security in an online student health records system*, Department of Computer Science, Bowie State University, Tech. Report (2014)

[19] J. Peng, X. Huang, M. Li, J. Zhang, Y. Zhang, and N. Gao, *Differential Attribute Desensitization System for Personal Information Protection*, 2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing &Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation, pp. 1243-1248 (2019)

[20] S. Anawar, Y. C. Kuan, Z. Ayop, and E. Hamid, *INTEGRATED IoT-CLOUD BASED mHealth PARTICIPATORY SENSING*, Internet of Things: Smart Systems and Application, p. 11 (2019)

Syarulnaziah Anawar holds her Bachelor of Information Technology (UUM), Msc in Computer Science (UPM), and PhD in Computer Science (UiTM). She is currently a Senior Lecturer at the Faculty of Information and Communication Technology, UTeM. She is a member of the Information Security, Digital Forensic, and Computer Networking (INSFORNET) research group. Her research interests include human-centered computing, participatory sensing, mobile health, usable security and privacy, and societal impact of IoT.

Muhammad Hafiz Jamil received diploma of ICT from Universiti Teknikal Malaysia Melaka in 2019 and currently is a pursuing his Bachelor of Science Computer (Computer Networking) in UTeM. His research interest is in computer systems, data management, and networking.

Zakiah Ayop holds BSc. in Computer Science (2000) from UTM and MSc in Computer Science (2006) at UPM. Currently she is a senior lecturer in Faculty of Information and Communication Technology (FTMK), Universiti Teknikal Malaysia Melaka (UTeM). She is a member of the Information Security, Digital Forensic, and Computer Networking research group. Her research interest are Information System, Internet of Things (IoT) and Network and Security.

Nur Fadzilah Othman received a degree in Computer Engineering in 2008 and master's in educational technology in 2011 at Universiti Teknologi Malaysia (UTM). In 2017, she obtained her PhD in the field of Information Security at Universiti Teknikal Malaysia Melaka (UTeM). She started her career as a senior lecturer at the Faculty of Information Technology and Communication, UTeM from March

2018. She is an active researcher and has been written and presented a number of papers in conferences and journals.

Norharyati Harum holds her bachelor's in engineering (2003), MSc. in Engineering (2005) and PhD in Engineering (2012) from Keio University, Japan. She is currently a senior lecturer at Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka (UTeM). Her interests in research area are Internet of Things (IoT), Smart Applications, Embedded System, Wireless Sensor Network, Next Generation Mobile Communication, Radio Frequency Planning and Signal Processing. She is an accomplished inventor, holding patents to radio access technology, copyrights of products using IoT devices.

Erman bin Hamid is Senior Lecturer at the Department of Computer Systems and Communications, Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka (UTeM). Having educational backgrounds in Information Technology and Computer Networking, motivate him to explore potentials and opportunities offered by both fields. He enjoys doing research in the area of Network Management, Network Visualization and Internet of Things (IoT).

.

Suzana Zambri holds her bachelor's in Information Technology (2001) and MSc. in Information Technology (2003) from Universiti Utara Malaysia. She is currently a senior lecturer at Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA (UiTM). Her interests in the research areas are Digital Entrepreneurship, Digital Marketing, Digital Content Creation. She is a member of the Industrial Informatics Research Group. She is an accomplished inventor, holding copyrights of products using IoT devices, mobile apps, and web-based systems. She is an active researcher and has written and presented a number of papers in conferences and journals.