Highlights

Attacking and Defence Pathways for Intelligent Medical Diagnosis System (IMDS)

Ying He,Ruben Suxo Camacho,Hasan Soygazi,Cunjin Luo

- This research identified a lack of a systematic research into the attacking and defence pathways of the Intelligent Medical Diagnosis System (IMDS);
- This research developed an IMDS simulation platform to simulate a realistic medical system, using the OpenEMR and an added Cardiac Diagnosis Component;
- This research demonstrated ethical hacking into the IMDS diagnosis records following the NIST ethical hacking method and identified four major vulnerabilities of the OpenEMR;
- This research identified the hacking pathway into the embedded cardiac diagnosis component and presents a set of cyber security strategies tailored for IMDS to prevent diagnosis records compromise.

Attacking and Defence Pathways for Intelligent Medical Diagnosis System (IMDS)*

Ying He[a], Ruben Suxo Camacho[b], Hasan Soygazi[b] and Cunjin Luo[c,d,*]

[a]School of Computer Science, University of Nottingham, Nottingham NG8 1BB, United Kingdom
[b]School of Computer Science and Informatics, De Montfort University, Leicester LE1 9BH, United Kingdom
[c]School of Computer Science and Electronic Engineering, University of Essex, Colchester CO4 3SQ, UK
[d]Key Lab of Medical Electrophysiology, Ministry of Education, Institute of Cardiovascular Research, Southwest Medical University, Luzhou 646000, China

ARTICLE INFO

Keywords:
Ethical Hacking
OpenEMR System
Intelligent Medical Diagnosis System (IMDS)
Cardiac Diagnosis Records
Cyber Defence Strategies

ABSTRACT

Background: The Intelligent Medical Diagnosis System (IMDS) has been targeted by the cyber attackers, who aim to damage the Healthcare Critical National Infrastructure (CNI). This research is motivated by the recent cyber attacks happened worldwide that have resulted in the compromise of medical diagnosis records. This study was conducted to demonstrate how the IMDS could be attacked and diagnosis records compromised (i.e. heart disease) and suggest a list of security defence strategies to prevent against such attacks.
Methods: This research developed an IMDS simulation platform by implementing the OpenEMR system. A Cardiac Diagnosis Component is then added to the IMDS. The IMDS is fed with the ECG data (retrieved from the PhysioNet/Computing in Cardiology Challenge 2017). This research then launched systematic ethical hacking, which was tailored to target IMDS diagnosis records. The systematic hacking was based on the NIST ethical hacking method and followed an attack pathway, starting from identifying the entry points of the medical websites, then propagating to gain access to the server, with the ultimate aim of modifying the heart disease diagnosis records.
Results: The hacking was successful. Four major vulnerabilities (i.e. broken authentication, broken access control, security misconfiguration and using components with known vulnerabilities) were identified in the simulated IMDS and the cardiac diagnosis records were compromised. This research then proposed a list of security defence strategies to prevent such attacks at each possible attacking points along the attacking pathway.
Conclusions: This research demonstrated a systematic ethical hacking to the IMDS, identified four major vulnerabilities and proposed the security defence pathways. It provided novel insights into the protection of IMDS and will benefit researchers in the community to conduct further research in security defence of IMDS.

1. Introduction

In the last decade, technology has become an important part of the healthcare industry and one of the main targets for cyber terrorists. According to the SANS Institute report, 94% of the organisations in healthcare have been targeted by the cyber attackers [1]. Recent research shows that over the past five years time, thousands of healthcare related data breaches have been reported, affecting more than 154 million health records in total [2]. The Intelligence Medical Diagnosis System (IMDS), such as implantable cardioverter defibrillator (ICD) [3] has revolutionised the way of diagnosing diseases. The IMDS has been created to improve the quality of services in health care and the usage of such system has increased rapidly in recent years [4]. IMDS enables early detection and diagnosis, providing opportunities for early diseases intervention, hence increasing the successful rate of treatment [5, 6]. The IMDS provides medical diagnosis to the patients using its intelligent diagnosis component based on exiting medical records and the data collected from medical devices. It helps reducing the workload of the doctors and the medical cost. However, any advancement of the IMDS will be in vain if the diagnosis records are compromised. IMDS has been targeted by the cyber attackers, who aim at damaging the Critical National Infrastructure (CNI). This paper is motivated by the incidents happened worldwide that have resulted in the compromise of diagnosis results [2, 7]. This creates challenges to the sustainability of the IMDS, which falls in the category of global sustainability of health and well being.

Recent research show that IMDS has significant security vulnerabilities [8, 9]. McMahon et al. [10] identified that almost 10% of medical devices have critical security risks. The findings reveals that the most common vulnerability was in Drop bear SSH Server, which can be exploited by executing malicious code. Another vulnerability was found in the MS17-010 security update. Similar to other type of vulnerabilities, attackers can run commands and execute codes remotely to compromise the system. Williams and Woodward [1] explained that the futuristic use of implanted medical devices at one time was demonstrated by successful attacks on devices such as insulin pumps and pacemakers. The lack of security controls in the devices is a significant concern.

*This document is the results of the research project funded by National Natural Science Foundation of China (NSFC) under Grant No. 61803318 (Cunjin Luo).
*Corresponding author: cunjin.luo@essex.ac.uk
ying.he@nottingham.ac.uk (Y. He); P2446462@my365.dmu.ac.uk (R.S. Camacho); P17241568@alumni365.dmu.ac.uk (H. Soygazi); cunjin.luo@essex.ac.uk (C. Luo)
ORCID(s): 0000-0003-2023-5547 (Y. He)

Existing research shows that problems such as access to infusion pump via web interfaces, default coded passwords, internet access to devices, are the common security issues detected in medical devices [11]. Communications on web services without authentication and encryption are also common vulnerabilities in medical devices as the attackers can access systems remotely [1].

The governments and research communities have realised the importance in protecting their IMDS [12, 13, 10, 14]. Current cyber security research in healthcare mainly focuses on maintaining or protecting medical devices [12, 13, 10, 14, 15, 16, 3, 17]. For example, Kainz, et al. developed a security system simulator in order to help manufacturers of medical devices or security systems identify incompatible combinations of medical devices [18]. However, those strategies cannot be applied to protect IMDS diagnosis records. There is also research in proposing data encryption mechanisms [19] and the combination with scrambling techniques [20] to protect wavelet-based ECG data both in transit and in storage. It does not provide a systemic view on how the attackers manage to gain access to medical systems and ultimately research the ECG data.

There is existing research attempting to simulate a medical platform and assess its vulnerabilities in a systematic manner [21]. However, the simulated environment is high level and does not reflect the comprehensive functionalities of a realistic medical system. Accordingly, this research was undertaken to develop an IMDS simulation environment with comprehensive medical operational processes by implementing an open source system, OpenEMR. This research then launched a series of ethical hacking to show how IMDS diagnosis records can be compromised and presented a list of cyber defence strategies to prevent such compromise. This study makes the following contributions,

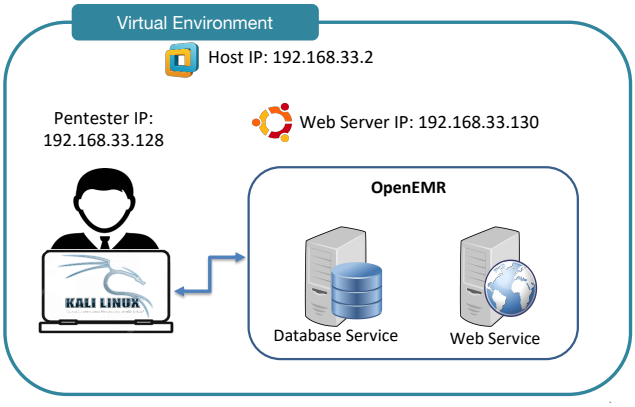
- develops an IMDS simulation platform to simulate a realistic medical system, using OpenEMR and an added Cardiac Diagnosis Component;
- demonstrates ethical hacking to the IMDS diagnosis results following the NIST ethical hacking method;
- identifies four major vulnerabilities for OpenEMR and identifies the pathway of hacking into the embedded cardiac diagnosis component;
- presents a set of cyber security strategies tailored for IMDS to prevent diagnosis records compromise.

2. Methods

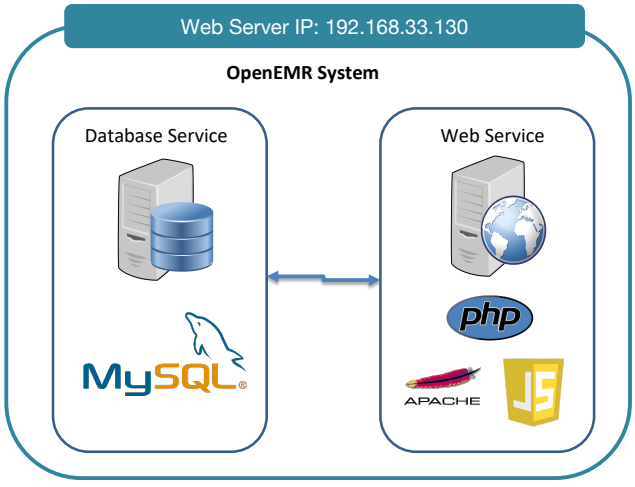
2.1. IMDS Simulation Platform

This section introduces the IMDS simulation platform which was developed to carry out the ethical hacking research. IMDS simulation platform is required as it is unrealistic to launch ethical hacking towards real world medical systems due to its impact on medical business operations. In this case, the IMDS simulation platform implemented the OpenEMR system, which is an open source medical system, and added a Cardiac Diagnosis Component to simulate a realistic IMDS. This research then fed the ECG data from the PhysioNet/Computing in Cardiology (CinC) Challenge 2017 into the simulated IMDS. The rest of this section introduces the details of the development of the IMDS platform.

,



(a) Ethical Hacking Simulation Platform



(b) OpenEMR System Simulation Platform

Figure 1: Virtual Simulation Platform

OpenEMR Implementation. The OpenEMR was selected among a list of open source medical systems as it accomplishes with the HIPAA and NIST standards and it is also the most widely used medical system by different healthcare organisations [22]. OpenEMR supports a comprehensive security risk management scheme which is based on the HIPAA and NIST standards [23]. OpenEMR contains main medical business functionalities including practice management, electronic medical records, electronic billing, scheduling, free support and it can be integrated with other medical systems and devices. It runs over PHP, MySQL and JavaScript, and is compatible with Linux, Windows and Mac OS. Also, the system is ONC certified, which is an IT certification for Health programs. OpenEMR contains a database component which has more than 100 interconnected tables. The OpenEMR version used for building this simulation environment is OpenEMR 5.0.1.3. The VMware [24] was used

to create a virtual environment and the infrastructure of the virtual environment is shown in Figure 1(b). The host operation system of the VMware is Window 10. The Linux (*Ubuntu* 18) operating system was installed on VMware and the OpenEMR was implemented on Linux (*Ubuntu* 18).

Ethical Hacking Virtual Environment. This research set up a virtual environment where an ethical hacking can be launched. Linux (*Kali*) was installed on VMware, which will be used for launching attacks. More details is illustrated in Figure 1(a). This research then set up the networking connection between the OpenEMR and the Ethical Hacking VM. There are two types of configurations, NAT and Bridge. NAT creates a virtual network adapter that works as the gateway where all the virtual machines with the same configuration can be connected with each other creating a small network. Bridge configuration installs a feature into the physic network adapter, where the VMs take the network configuration similar to the host computer. Unlike NAT configuration, Bridge connects virtual machines with the rest of computers. Both can enable the virtual machines to access the internet. In this case, the NAT configuration was chosen using the IP 192.168.33.130.

Cardiac Diagnosis Component Development. This research developed and implemented a cardiac diagnosis component into the OpenEMR. To do this, the internal code was modified to integrate this component and display the ECG records inside OpenEMR. To achieve this, additional tables were created in order to hold the ECG data. The OpenEMR database can then be accessed through using phpMyAdmin and created a table called "*ecg_records_data*". The ECG data from the PhysioNet/Computing in Cardiology (CinC) Challenge 2017 was fed into the table "*ecg_records_data*". After configuring the database, the system source code has to be modified to display the ECG results. Moving to the folder "*/interface/patient_file/summary/*" found the file "*demographics.php*". Between the style and head mark, a java script code was added to graphic the database results. We moved to the folder "*interface/patient_file/summary*" and looked for the file "*vitals_fragment.php*". A script was added to create containers of the graphics. The SQL code was then created to retrieve ECG data from the database. This research then set colors configuration of graphics and created code to graphic each line in the graphics.

Classification of Heart Diseases. As the ethical hacking is targeting the ECG diagnosis data, our ultimate aim is to modify the heart disease diagnosis records. In order to make meaningful changes, we need to understand different types of arrhythmia diseases. There are mainly two types of arrhythmias, tachycardia meaning heartbeat is greater than 100 beats in one minute and bradycardia meaning heartbeat is slow which is less than 60 beats in one minute. Moreover, these have sub-categories which are atrial fibrillation, atrial flutter, supraventricular tachycardia, ventricular tachycardia, ventricular fibrillation, long QT syndrome, sick sinus syndrome and conduction block [25, 26].

2.2. Systematic Ethical Hacking to IMDS

Ethical hacking, which is also called penetration testing, is a set of tests that are launched on information systems to identify vulnerabilities emulating being a real attacker using the same tools and techniques [27]. Widely used ethical hacking methodologies includes but not limited to the ones from National Institute of Standards and Technology (NIST) [28], Penetration Testing Execution Standard (PTES) and Open Web Application Security Project (OWASP). Some organisations such as EC-council or Offensive Security have developed their own methodologies [27].

A systematic ethical hacking include four main stages, which are information gathering, discovery, attacking and reporting. At the information gathering stage, the tester performs reconnaissance and collects all possible information about the target before starting the attack. At the discovery stage, the tester attempts to understand the structure of the system analysing the paths and directories. At the attacking stage, the tester identifies the vector to attack, usually based on the results of the vulnerability scanner and tests them against the OWASP Top 10 vulnerabilities [29]. At the reporting stage, the tester takes the evidence of the last stages to write the report with the principal findings. This research launched a series of ethical hacking aiming to exploit the vulnerabilities of the IMDS. The ultimate goal was to compromise the ECG records and its waveform data. To achieve this, this research identified the entry points to the system, gained a foothold and then explored the vulnerabilities within the system, and finally found opportunities to modify the diagnosis records. This ethical hacking was carried out at the network level through the virtual network as depicted in Figure 1. This research followed the NIST ethical hacking methodology to perform a systematic ethical hacking [28].

3. Results

3.1. Information Gathering Stage

In this stage, the goal was to collect information from public sources in order to understand the background of the target and the infrastructure where is based on. This stage was not intrusive, and the information gathering was a reconnaissance based on the review of the publicly available information. The tester had found the OpenEMR user guide and its database structure, which could help identify the attacking pathways in the discovery stage. The tester then checked connectivity with target 192.168.33.130 using *ping* command and received a successful communication respond. The tester used the tool "whatweb" to check for more details on the web technologies. The tester was able to find the version of part of the framework where OpenEMR was running, i.e. the versions of Apache server and PHP.

3.2. Discovery Stage

In this stage, the tester applied the tools including Web Crawler, Port Scanner and Vulnerability Scanner. *Web Crawler* was used to identify directories and location that may contain vulnerabilities or forbidden location of externals. This

technique uses a dictionary to test the URL with different directories. The process is recursive, which means if the tool finds a directory, the process starts again from that point. Most of the directories found were able to get access without authentication. These identified directories will be used in the attacking stage. *Port Scanner* was used to scan ports and services to obtain valuable information about the structure of the server and version of the services provided. The tester found that the server was using the program $ProFTP$ to transfer files. The version of $Apache$ 2.4.39 over *port* 80. *Port* 443 was not configured as this environment is local, but in real environment this should be the port for web service. The tester also found a database using *port* 3306. *Vulnerability Scanner* is an automated tool that can run different scripts to identify vulnerabilities in the website. The tester used "Nikto" [30, 31] which is an in-built tool of Kali Linux. "Nikto" analyses the services and their version to identify vulnerabilities based on a set of well-known vulnerabilities. Additionally, this tool contains a set of tests, which are executed to verify certain vulnerabilities. The command was "$nikto - hosthttp : //192.168.33.130/openemr/$". Most of the vulnerabilities found were based on the default configuration of the server. This information will be co-related with the findings in the attacking stage.

3.3. Attacking Stage

This stage is based on the OWASP Top 10 vulnerabilities which are the 10 most common vulnerabilities for web applications [29]. The tester used various tools to test against the Top 10 vulnerabilities and the tester were able to identify four major vulnerabilities. Due to page limits, the tester only report the procedure to discover these four vulnerabilities.

A2:2017-Broken Authentication. Some functions of applications related to authentication and session management are frequently applied erroneously, allowing attackers to compromise passwords, keys, session tokens, or to exploit other implementation flaws and pretend to be another user, which is extremely vulnerable to Brute Force attack. The tester used "Burp Suite" to launch a Brute Force attack. A step by step instruction on using the Burp Suite can be found in [32]. The finding of this attack was the system administrator user name "$admin$" and the password "$admin$" .

A5:2017-Broken Access Control. Restrictions on what authenticated users are allowed to do are often not properly enforced. Attackers can exploit these flaws to access unauthorised functionalities and/or data, such as accessing other users' accounts, viewing sensitive files, modifying other users' data, changing access permissions, etc. To start with, the tester accessed the system with a user, in this case as administrator. Then, the tester obtained the cookie of that session using any add-on about cookie manager like "Open cookie manager" for Firefox. In this case, the value of cookie obtained was "$104684c91029014d427ef a0d124 f 4ec1$". The tester then tried to access the system from another computer using another user. In this case the tester used a common user named "user1". In this computer, start any cookie manager add-on. Then copy the value taken from the computer where the administrator was logged in, change the value of the current cookie with this one and save the cookie. After a refresh of the website, the cookie has been stolen. The tester is now able to navigate as administrator without administrator login credentials.

A6:2017-Security Misconfiguration. Security misconfiguration is the most commonly seen issue. This is due to insecure default configurations, incomplete or ad-hoc configurations, open cloud storage, misconfigured HTTP headers, and verbose error messages containing sensitive information. Not only must all operating systems, frameworks, libraries, and applications be securely configured, but also they must be patched and upgraded in a timely fashion. Within the OpenEMR system, there was an option to upload files in the "Administration" tab and "File". The website is showing the path where the images can be located. Before choosing a file, locate a reverse shell in Kali Linux. Then, choose any of them. After that, the file was opened with a text editor to set it. In the IP variable, insert the Kali Linux IP and the port can be left by default and save it. Then, the file was chosen and uploaded. Now, save to complete this process. Following the path where the images are stored, the tester can see the uploaded file in the list. This file contained a code to give access to the server from a remote computer. After verifying that the uploaded file was there. Open a terminal in Kali Linux to configure the listener. Use the command "$nc - lnvp1234$". This command is waiting for any connection where the destination is Kali through the *port* 1234. Then, click on the uploaded file called "$php - reverse - shell.php$" and the connection was established. Now, it is possible to execute command lines remotely like "ipa" or "$uname - a$". To check the privileges inside. It was possible to access the file system "$/etc/passwd$" where it stored the users' details without privilege escalation.

A9:2017-Using Components with Known Vulnerabilities. Components such as libraries, frameworks, website, and other software modules, run with the same privileges as the application. If a vulnerable component is exploited, such an attack can facilitate serious data loss or server takeover. Applications and APIs using components with known vulnerabilities may undermine application defences and enable various attacks and impacts. In this task, the "searchsploit" [33] tool was used to explore public exploits. The command for this exploit, is "*python 45161.py http://192.168.33.130 /openemr -u user1 -p password -c'bash -i >& /dev/tcp/192// .168.33.128/1234 0>&1*". According to the usage of the exploit, it required a username and password. This vulnerability might be combined with a previous Brute Force attack to obtain the login credentials. Before executing this command, the computer of the tester should listen that port. After a successful execution, it was possible to execute commands remotely. This exploit got remote access to the server. After gaining access, the tester was able to execute commands and code, or escalate privileges to root permissions.

3.4. Reporting Stage

,

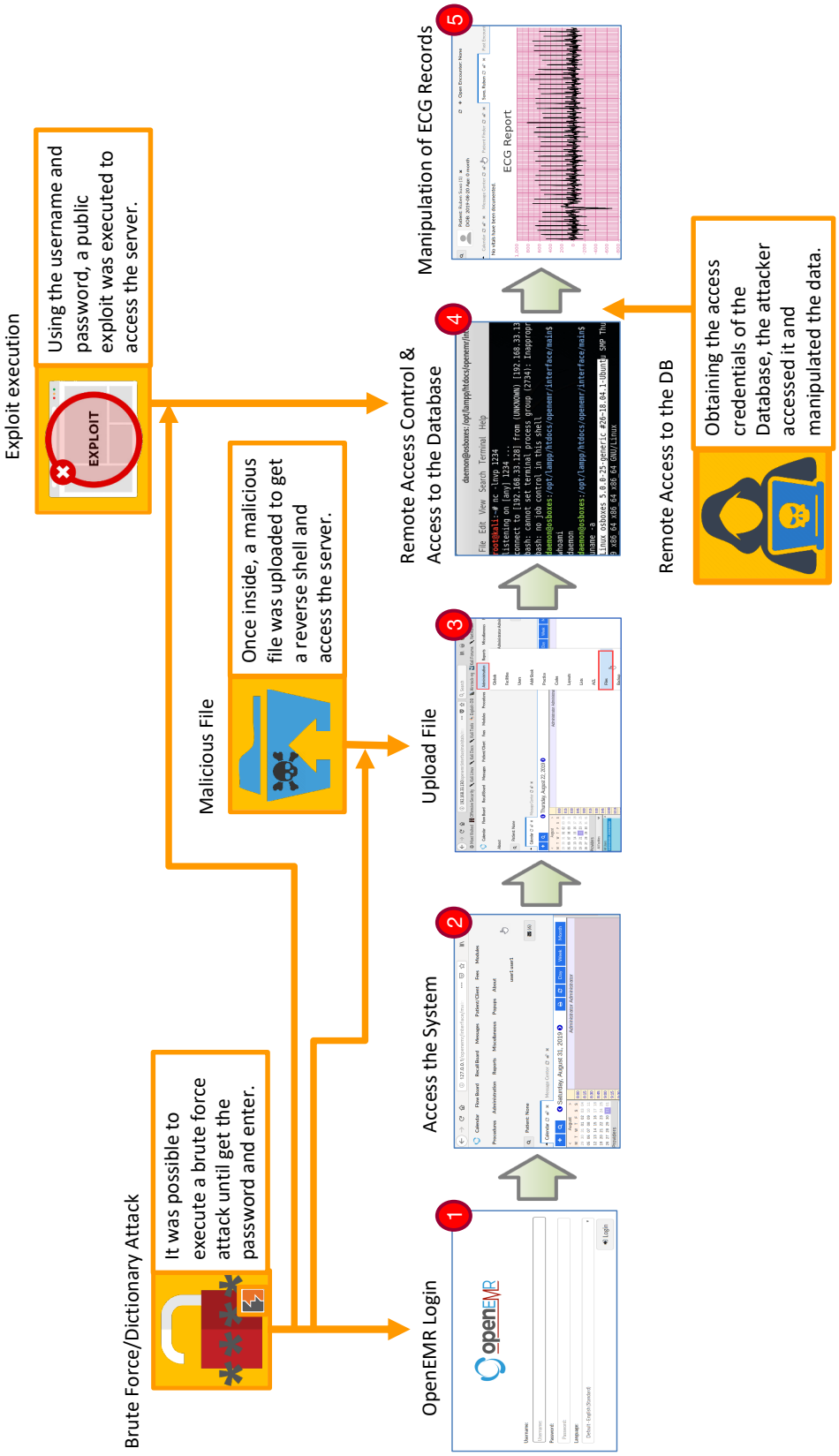


Figure 2: Ethical Hacking Pathway Workflow

Table 1
Identified OpenEMR Vulnerabilities - Based on OWASP Top 10 Vulnerabilities

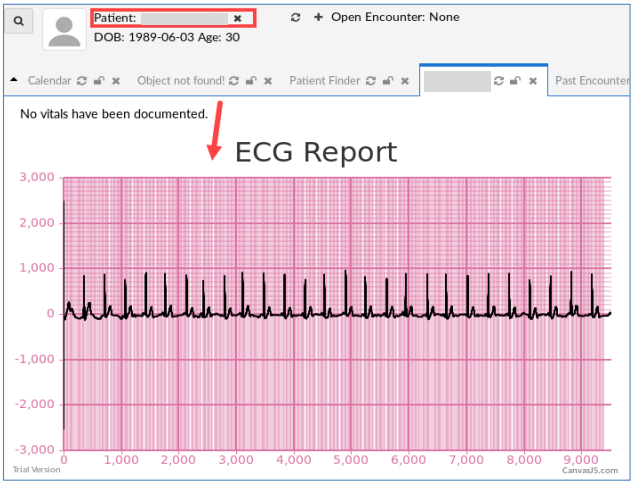
Identified Vulnerabilities	Vulnerabilities Description
A2 Broken Authentication	This vulnerability can be exploited using Brute Force attack. It is possible to try infinite authentication attempts until the correct username and password are obtained.
A5 Broken Access Control	To exploited this vulnerability, the attacker needs to get access to the computer or to manipulate the user with social engineering in order to obtain the cookies. The system allows two sessions opened at the same time. This means that the system is not managing the session and cookies properly. A normal user is able to steal the cookie of an admin user and can navigate as administrator without admin login credentials.
A6:Security Misconfiguration	To exploit this vulnerability, the attacker should gain access to the system and have the permission to upload files in the system. After the exploit, the system can receive any files without checking the extension of the file. In this case, it was possible to upload a web shell file, which contained a code to give access to the server from a remote computer.
A9 Using Components with Known Vulnerabilities	There is a public exploit that can exploit a vulnerability in the OpenEMR system. This exploit can get remote access to the server. After accessing the system, the attacker was able to execute commands and code, and escalate privileges to root permission. When analysing the exploit, the file "$/interface/super/edit\_globals.php$" was able to rewrite forms, and with those forms, it was able to take advantage on "$/interface/main/daemon\_frame.php$" to execute commands remotely.

Table 2
Security Defence Strategies for the Identified OpenEMR Vulnerabilities

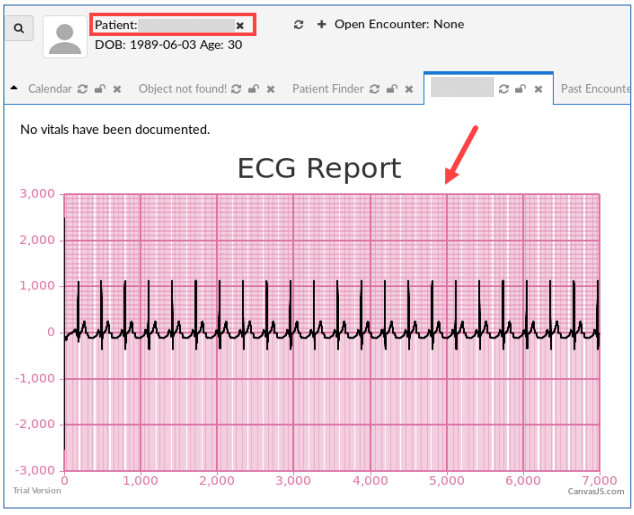
Identified Vulnerabilities	Security Defence Strategies
A2 Broken Authentication	It is recommended to apply a Captcha [34] or two-factor authentication [35] for login. Another option is to limit the number of attempts and block the user when wrong passwords are attempted.
A5 Broken Access Control	It is recommended to apply secure cookie management [36], for example accepting only one session per user. Additionally, the cookie should attach more data like the version of the web browser or the version of the operating system to complete the value of the cookie. Doing this, one cookie will only work for one browser.
A6:Security Misconfiguration	It is recommended to check the file extension and check the header of the file, all of which should be controlled by the source code. Additionally, only allow images in this case and no other types of files can be uploaded to the medical system [37]. There are also tools to automatically detect PHP-based unrestricted file upload vulnerabilities [38].
A9 Using Components with Known Vulnerabilities	It is recommended to analyse the privilege of "$edit\_globals.php$" and "$daemon\_frame.php$" and disable the modification of different forms or parameters as well as disable the command execution from them.

This stage summaries the vulnerability identified (as shown in Table 1) and attacking pathways (Figure 2). The ethical hacking tester started with the exploitation of "A2:2017-Broken Authentication" by using a Brute Force attack until getting the correct password. As an alternative, the tester can also access the website by exploiting the "A5: 2017-Broken Access Control" vulnerability. Once on the web page, the tester started looking for vulnerabilities, where the tester found a tab to upload files. On the web page the system showed a message explaining that the permitted file extensions were JPG, PNG and PDF. The tester was then able to exploit the "A6:2017-Security Misconfiguration" vulnerability to upload and execute a PHP file. The PHP file contained a code to give access to the server from a remote computer. Another way to gain access to the server was to use a public exploit, in this case, the exploitation of "A9:2017-Using Components with Known Vulnerabilities". After executing the exploit, the tester got access to the server. Once inside, the aim was to look for valuable files or other passwords. The database connection file was found and it stored user login credentials. The tester then got access to the database and was able to modify the data. In this case, the password of the database was empty as the setting was by default. Even if the password was formed by more than 16 characters using numbers and special characters combined, this exploitation permits to see the password in plain text. After knowing the password, it was possible to connect directly with the database service. Inside was found the "$ecg\_records\_data$" that contained the ECG records of patients. The records of a patient was then checked and later successfully modified by using an SQL query. In this case, the ECG database was modified, changing the ECG records of a patient from normal to one with arrhythmia. Figure 3 shows the comparison results before and after Ethical Hacking.

,



(a) ECG Records Before Ethical Hacking



(b) Compromised ECG Records After Ethical Hacking

Figure 3: ECG Records Before and After Ethical Hacking

3.5. Cyber Defence Pathway

This section presents the cyber defence strategies on how the identified vulnerabilities of the IMDS can be addressed and the security solutions that can stop the attack at each attacking points. Table 2 summaries a list of security recommendations. Figure 4 illustrates the vulnerabilities on the attack pathway and how the defence mechanisms can be applied to reduce the probability of a successful attack.

The attack started with gaining access to the system. The scheme shows two different solutions to protect it from password cracking through Brute Force attack, implementing a two-factor authentication or a Captcha. Two-factor authentication method forces the user to present at least two different pieces of evidence or factors to access the system. Those factors should be about something the user knows (a password), something the user possesses (a PIN) or something the user is (his fingerprint). A Captcha is a program or code that tries to identify if the user is a computer or a human by little random puzzles or questions that requires manual interactions. Either use two-factor authentication or a Captcha or both to block automated login access attempts.

When an attacker enters a server, he/she can make changes. A recommended solution is the system administrator back up data regularly. Once the system administrator knows that the data has been compromised, he/she can recover the original data. Depending on the internal security policies, the backup could be daily, weekly, monthly or yearly. It is also important to enable the logging and auditing functionalities to track changes." and auditing functionalities to track change.

The attacker was then able to upload a file and execute it. The website does not have a mechanism to detect the extension of the file and check the header of the file as well to verify the veracity of its format. This point could be fooled, but the same procedure should be made in the server, verifying the veracity of the file in the website and the server to ensure that it is not a malicious file [37]. There are also tools to automatically detect PHP-based unrestricted file upload vulnerabilities [38].

When the attacker was inside, he/she discovered a configuration file that contained credentials in plain text. The solution is to separate the web server from the database server. The database should not accept remote connections and the queries should come from an authenticated web server by using an internal API that requires a separate authentication by another user. In this way, even when the attacker obtains the credentials to access the database, the connection would be denied [37].

If the attacker cannot be stopped at earlier stages, they would ultimately reach the patients' medical records, the ECG records, in this study. At this stage, the system administrator needs to detect what data was changed and identify the data to be recovered. The recommendation is to store the hashes of each record into another table before the attack. Then, when the program requires a record, take the current hash of the records by the website and compare it with the hash stored in the database. If the hashes are equal, it means that the data has not been compromised, otherwise the data may have been compromised. This concept was taken from the "salt password store" [39] process but modified to protect the integrity of data. Once the compromised data is discovered, the system should alert the medical users that the files have been compromised [40, 41]. The system administrator should also take immediate actions to recover the data from the backups and inform the medical users when the data is available to use.

,

4. Discussion

In previous sections, this research demonstrated ethical hacking to exploit four vulnerability identified from the simulated IMDS. This research also provided bottom-line security defence strategies to prevent these exploit. Table 3 provides an expanded version of security defence strategies if

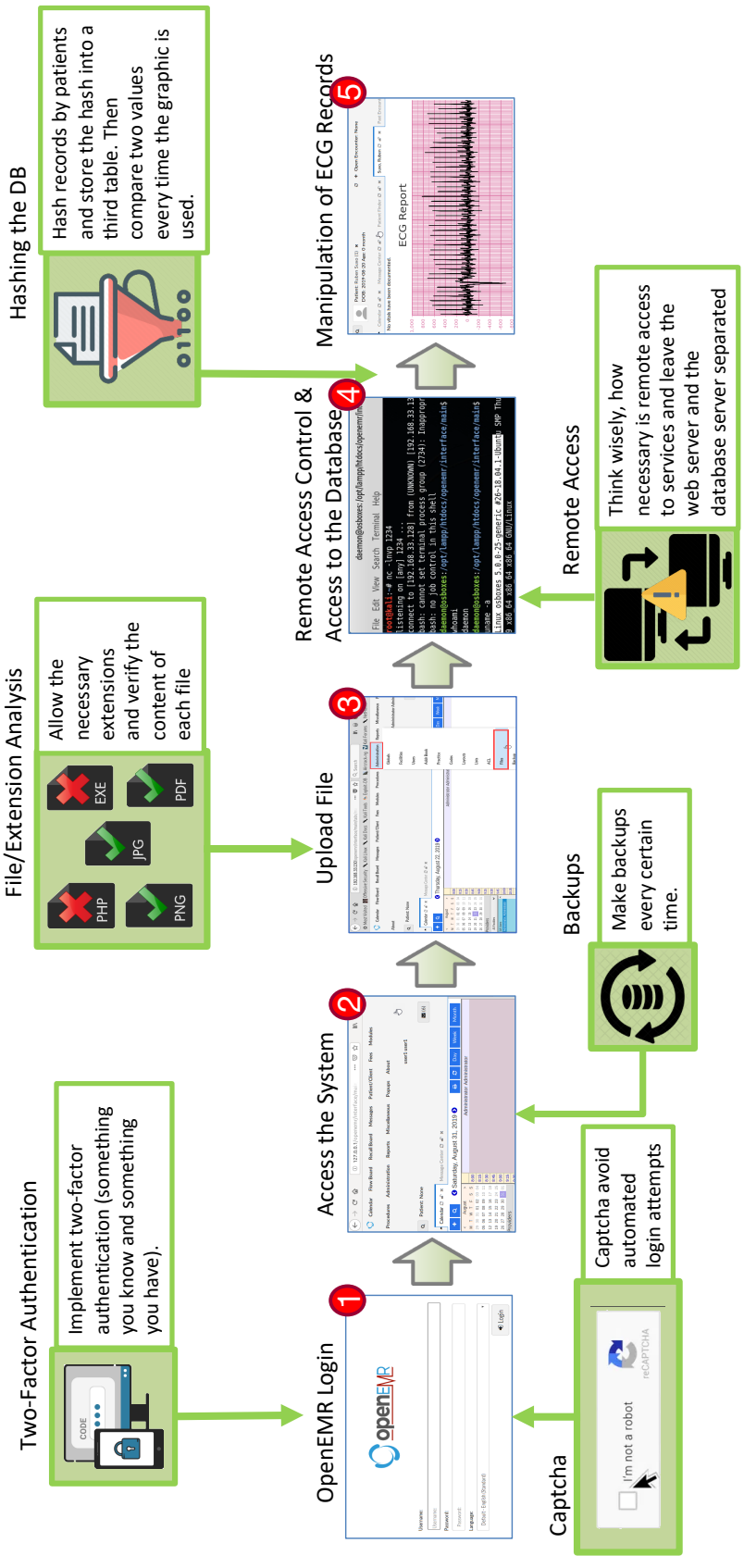


Figure 4: Security Defence Pathway Workflow

Table 3
Enhanced Security Defence Strategies

Vulnerabilities	Security Defence Strategies
Code Injection	It is recommended that prepared statements should be used and input should be validated. Error messages should be disabled and database credentials should be separated and encrypted.
Sensitive Data Exposure	It is recommended that all sensitive data such as passwords should be encrypted with strong, robust and up-to-date algorithms.
Broken/Weak Authentication	It is recommended that, two-factor authentications should be used. Furthermore, complexity and length of passwords should be restricted to avoid weak password generation. Login activities should be limited after the failure of login.
XML External Entities	It is recommended that XML processors and libraries should be patched and up-to-date.
Broken Access Control	It is recommended that strong and complex password with two-factor authentication should be used.
Security Misconfiguration	It is recommended that all unnecessary features should be disabled and design by security principles should be considered.
Cross-Site Scripting (XSS)	It is recommended to validate and filter the inputs as strictly as possible based on what is expected.
Using Component with Known Vulnerabilities	It is recommended that features and components that are not used should be removed from the system. The component should be updated regularly and while using components vulnerability databases such as CVE should be considered.
Insufficient Logging & Monitoring	It is recommended that all activities should be monitored, the system should be tested regularly and log files should be stored properly. Moreover, error messages should be clear and adequate.
Backup Data Exposure	It is recommended that, data should be stored in an encrypted format or should be destroyed if it is unused.

a thorough protection of the IMDS is needed depending the cyber security budget, resource availability and data protection priorities.

The data related to the patients especially the diagnosis data is categorised as sensitive data. Electrocardiogram (ECG) data includes both patient information and the ECG waveform data and is classified as sensitive data. The intelligent medical diagnosis system (IMDS) has been created to improve the quality of services in health care. IMDS has gained popularity and the usage of such system starts to increase rapidly [4]. IMDS is of high importance in healthcare as it can help detect disease at an earlier stage [5, 6], hence increasing the successful rate of treatment. The IMDS consists of artificial intelligence components such as experts systems and knowledge-based systems as well as database management systems that are used to store medical records. Expert systems are designed for resolving complex cases by using decision-making capability of human beings whereas knowledge-based systems use knowledge from different data sources to solve complex problems.

Existing cyber security research focuses on the protection of medical devices [12, 13, 10, 14, 15, 16, 3, 17] and medical data [20]. Kainz, et al. developed a security system simulator in order to help manufacturers of medical devices or security systems identify incompatible combinations of medical devices [18]. Ma, et al. proposed the data encryption mechanisms [19] combined with scrambling techniques [20] to protect wavelet-based ECG data both in transit and in storage. Edemacu et al. proposed a novel access control scheme to protect the security of the shared health data in collaborative ehealth systems[42]. Das et al. developed a lightweight authentication protocol for wearable devices environment[43]. Sánchez-Guerrero et al. proposed a privacy-aware profile management approach to manage the privacy of patient profiles [44]. Hosseini, et al. adopted machine learning techniques to improve the security of patients' portals and websites [45]. Seepers et al. investigated and developed mechanisms to enhance heart-beat-based security [46, 47]. There are also research efforts attempting to address the security protection in a systematic manner, mainly in the area of security risk assessment, security behavior analysis and security management frameworks [48, 49, 50, 51, 52, 53]. Evans et al. proposed a systematic approach to analyse human error related security incidents and implemented it in both public sector and private sector healthcare organisations [48]. He and Johnson developed a systematic approach to enhance the sharing and exchanging of security lessons in healthcare organisations across different countries [49]. Fernández-Alemán et al. performed an empirical study to analyse the health professional security behaviors [50]. Seepers et al. proposed a lightweight security framework to protect the medical data collected using mobile health solutions [51]. Entzeridou et al. investigated into the public and physician's security risk perception of electronic health records [52]. Yasqoob et al. proposed an integrated safety, security, and privacy (ISSP) risk assessment framework to assess the risks of medical devices and the required security controls [53]. Yin et al. proposed real-time monitoring and control of industrial cyberphysical systems framework that allows the plant-wide performance supervised monitor-

ing and control as well as fault detection [54]. Jiang et al. developed MATLAB toolbox data based key-performance-indicator oriented fault detection toolbox (DB-KIT) for process monitoring and fault diagnosis [55].

However, there is little research in assessing the attacking and defence pathways of the medical systems in a systematic manner [21]. Our research created a simulated environment reflecting the comprehensive functionalities of a realistic medical system and identified key vulnerabilities and provided defence strategies. It has implications for the IT professionals in healthcare to protect their IMDS.

5. Conclusion and Future Work

This research developed an IDMS simulation platform by implementing OpenEMR and an added Cardiac Diagnosis Component. This research demonstrated how the IMDS diagnosis results can be compromised through exploiting four vulnerabilities identified from the OpenEMR by following the NIST ethical hacking methodology. We then presented a set of cyber security strategies tailored to IMDS to prevent such compromise. This research then discusses security defence strategies in general in a wider scope to protect IMDS. This research demonstrated a systematic ethical hacking to the OpenEMR and provided novel insights into the protection of IMDS. It will benefit researchers in the community to conduct further research in security defence of IMDS.

From a security defence perspective, future work will comprehensively apply and extensively test the security mechanisms proposed in our cyber defence strategies. A follow-up systematic ethical hacking will also be carried out in order to test the effectiveness of the cyber defence strategies. Different security mechanisms will be compared taking into consideration the key performance indicator (KPI), metrics and reporting of cyber security [56]. The security defence should also take into account advanced ethical hacking techniques towards IMDS such as the AI based vulnerability assessment and ethical hacking. We will also attempt advanced security defence strategies, such as Security Information and Event Management (SIEM), Orchestration Automation and Response (SOAR) [57], and Security Operations Center (SOC). From a medical diagnostics perspective, future work will consider using a more mature IMDS, such as the arrhythmia detection and classification in ambulatory ECG proposed by Andrew Y. Ng [58]. Future work will also focus on expanding the data set to include data collected from different medical devices such as MCG and MRI.

Acknowledgment

We would like to thank the reviewers for their insightful comments. This work was supported by National Natural Science Foundation of China (Grant No. 6180-3318).

Competing interests

We declare there is(are) no conflict(s) of interest accociated with this research.

Author's contributions

Conceptualisation: YH and CL; Methodology: YH, RSC, HS and CL; Results: YH, RSC, HS and CL; Writing - Original Draft: YH, RSC, HS and CL. Writing - Revision: YH, RSC, HS and CL; Supervision: YH and CL.

Summary points

What was already known on the topic.

- The Intelligent Medical Diagnosis System (IMDS) has been targeted by the cyber attackers;
- There is an increasing number of cyber-attacks happened worldwide that have resulted in the compromise of medical diagnosis records;
- There is a lack of a systematic research into the attacking and defence of the Intelligent Medical Diagnosis System (IMDS).

What this study added to our knowledge,

- This study developed an IMDS simulation platform using the OpenEMR and an added Cardiac Diagnosis Component for ethical hacking purpose;
- This study identified the hacking pathway into the IMDS diagnosis records and identified four major vulnerabilities of the OpenEMR;
- This study presents a defence pathway and a comprehensive set of cyber security strategies tailored for IMDS to prevent diagnosis records compromise.

Appendix

Table 4
Abbreviations

Abbreviations	Description
CNI	Critical National Infrastructure
CVE	Common Vulnerabilities and Exposure
ECG	Electrocardiogram
HIPAA	Health Insurance Portability and Accountability Act
ICD	Implantable Cardioverter Defibrillatorl
IMDS	Intelligent Medical Diagnosis Systemn
MCG	Magnetocardiography
MRI	Magnetic Resonance Imagingn
NIST	National Institute of Standards and Technology
OpenEMR	Open-Source Electronic Medical Record
OWASP	Open Web Application Security Project
PTES	Penetration Testing Execution Standard
XSS	Cross-Site Scripting

References

[1] Patricia AH Williams and Andrew J Woodward. Cybersecurity vulnerabilities in medical devices: a complex environment and multifaceted problem. *Medical Devices (Auckland, NZ)*, 8:305, 2015.

[2] Jay G Ronquillo, J Erik Winterholler, Kamil Cwikla, Raphael Szymanski, and Christopher Levy. Health it, hacking, and cybersecurity: national trends in data breaches of protected health information. *JAMIA Open*, 1(1):15–19, 2018.

[3] Karen Sandler, Lysandra Ohrstrom, Laura Moy, and Robert McVay. Killed by code: Software transparency in implantable medical devices. *Software Freedom Law Center*, pages 308–319, 2010.

[4] Jonathan Guo and Bin Li. The application of medical artificial intelligence technology in rural areas of developing countries. *Health equity*, 2(1):174–181, 2018.

[5] Rahul Kala, Anupam Shukla, and Ritu Tiwari. Hybrid intelligent systems for medical diagnosis. In *Intelligent Medical Technologies and Biomedical Engineering: Tools and Applications*, pages 187–202. IGI Global, 2010.

[6] Arve Kjoelen, Marc J Thompson, Scott E Umbaugh, Randy Hays Moss, and William V Stoecker. Performance of ai methods in detecting melanoma. *IEEE Engineering in Medicine and Biology Magazine*, 14(4):411–416, 1995.

[7] Daniel B Kramer, Matthew Baker, Benjamin Ransford, Andres Molina-Markham, Quinn Stewart, Kevin Fu, and Matthew R Reynolds. Security and privacy qualities of medical devices: An analysis of fda postmarket surveillance. *PLoS One*, 7(7), 2012.

[8] Zhaowen Lin, Xinglin Xiao, Yi Sun, Yudong Zhang, and Yan Ma. A privacy-preserving intelligent medical diagnosis system based on oblivious keyword search. *Mathematical Problems in Engineering*, 2017, 2017.

[9] Clemens Scott Kruse, Brenna Smith, Hannah Vanderlinden, and Alexandra Nealand. Security techniques for the electronic health records. *Journal of Medical Systems*, 41(8):127, 2017.

[10] Emma McMahon, Ryan Williams, Malaka El, Sagar Samtani, Mark Patton, and Hsinchun Chen. Assessing medical device vulnerabilities on the internet of things. In *2017 IEEE International Conference on Intelligence and Security Informatics (ISI)*, pages 176–178. IEEE, 2017.

[11] Tehreem Yaqoob, Haider Abbas, and Mohammed Atiquzzaman. Security vulnerabilities, attacks, countermeasures, and regulations of networked medical devices—a review. *IEEE Communications Surveys & Tutorials*, 21(4):3723–3768, 2019.

[12] Axel Wirth. Cybercrimes pose growing threat to medical devices. *Biomedical Instrumentation & Technology*, 45(1):26–34, 2011.

[13] Meng Zhang, Anand Raghunathan, and Niraj K Jha. Trustworthiness of medical devices and body area networks. *Proceedings of the IEEE*, 102(8):1174–1188, 2014.

[14] Mandeep Khera. Think like a hacker: Insights on the latest attack vectors (and security controls) for medical device applications. *Journal of Diabetes Science and Technology*, 11(2):207–212, 2017.

[15] Guanglou Zheng, Rajan Shankaran, Mehmet A Orgun, Li Qiao, and Kashif Saleem. Ideas and challenges for securing wireless implantable medical devices: A review. *IEEE Sensors Journal*, 17(3):562–576, 2016.

[16] Johannes Sametinger, Jerzy Rozenblit, Roman Lysecky, and Peter Ott. Security challenges for medical devices. *Communications of the ACM*, 58(4):74–82, 2015.

[17] Shelby David Kobes. Security implications of implantable medical devices. 2014.

[18] Wolfgang Kainz, Jon P Casamento, Paul S Ruggera, Dulciana D Chan, and Donald M Witters. Implantable cardiac pacemaker electromagnetic compatibility testing in a novel security system simulator. *IEEE Transactions on Biomedical Engineering*, 52(3):520–530, 2005.

[19] Tao Ma, Pradhumna Lal Shrestha, Michael Hempel, Dongming Peng, Hamid Sharif, and Hsiao-Hwa Chen. Assurance of energy efficiency and data security for ecg transmission in basns. *IEEE Transactions on Biomedical Engineering*, 59(4):1041–1048, 2012.

[20] Ayman Ibaida and Ibrahim Khalil. Wavelet-based ecg steganography for protecting patient confidential information in point-of-care systems. *IEEE Transactions on Biomedical Engineering*, 60(12):3322–3330, 2013.

[21] Cunjin Luo, Hasan Soygazi, Helge Janicke, and Ying He. Security defense strategy for intelligent medical diagnosis systems (imds). In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 3454–3457. IEEE, 2019.

[22] Beatriz Sainz de Abajo and Agustín Llamas Ballestero. Overview of the most important open source software: analysis of the benefits of openmrs, openemr, and vista. In *Telemedicine and e-health services, policies, and applications: Advancements and developments*, pages 315–346. IGI Global, 2012.

[23] Maryam Farhadi, Hisham Haddad, and Hossain Shahriar. Compliance checking of open source ehr applications for hipaa and onc security and privacy requirements. In *2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC)*, volume 1, pages 704–713. IEEE, 2019.

[24] ZHOU Xiang-ying. Building virtual computer network experiment based on vmware [j]. *Research and Exploration in Laboratory*, 7, 2006.

[25] Mayo Clinic. Heart arrhythmia - symptoms and causes, 2019.

[26] Ridhi Saini, Namita Bindal, and Puneet Bansal. Classification of heart diseases from ecg signals using wavelet transform and knn classifier. In *International Conference on Computing, Communication & Automation*, pages 1208–1215. IEEE, 2015.

[27] Rafay Baloch. *Ethical hacking and penetration testing guide*. Auerbach Publications, 2017.

[28] Karen Scarfone, Murugiah Souppaya, Amanda Cody, and Angela Orebaugh. Technical guide to information security testing and assessment. *NIST Special Publication*, 800(115):2–25, 2008.

[29] Dave Wichers and Jeff Williams. Owasp top-10 2017. *OWASP Foundation*, 2017.

[30] Gilberto Najera-Gutierrez and Juned Ahmed Ansari. *Web Penetration Testing with Kali Linux: Explore the methods and tools of ethical hacking with Kali Linux*. Packt Publishing Ltd, 2018.

[31] Trupti Bhosale, Shraddha More, and SN Mhatre. Testing web application using vulnerability scan. 2019.

[32] Sunny Wear. *Burp Suite Cookbook: Practical recipes to help you master web penetration testing with Burp Suite*. Packt Publishing Ltd, 2018.

[33] Corey P Schultz and Bob Perciaccante. *Kali Linux Cookbook*. Packt Publishing Ltd, 2017.

[34] Haichang Gao, Mengyun Tang, Yi Liu, Ping Zhang, and Xiyang Liu. Research on the security of microsoft's two-layer captcha. *IEEE Transactions on Information Forensics and Security*, 12(7):1671–1685, 2017.

[35] Ding Wang and Ping Wang. Two birds with one stone: Two-factor authentication with security beyond conventional bound. *IEEE Transactions on Dependable and Secure Computing*, 15(4):708–722, 2016.

[36] Vamsi Motukuru, Vikas Pooven Chathoth, and Vipin Anaparakkal Koottayi. Cookie based session management, January 9 2018. US Patent 9,866,640.

[37] Mikael Willberg. Web application security testing with owasp top 10 framework. 2019.

[38] Jin Huang, Yu Li, Junjie Zhang, and Rui Dai. Uchecker: Automatically detecting php-based unrestricted file upload vulnerabilities. In *2019 49th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*, pages 581–592. IEEE, 2019.

[39] Loic Ferreira. Method for authenticating with a password comprising a salt, August 8 2019. US Patent App. 16/311,445.

[40] Ahmed Ghoneim, Ghulam Muhammad, Syed Umar Amin, and Brij Gupta. Medical image forgery detection for smart healthcare. *IEEE Communications Magazine*, 56(4):33–37, 2018.

[41] Yuuki Watanabe. Image data alteration detection device, image data alteration detection method, and data structure of image data, May 14 2020. US Patent App. 16/743,062.

[42] Kennedy Edemacu, Beakcheol Jang, and Jong Wook Kim. Collaborative ehealth privacy and security: An access control with attribute revocation based on obdd access structure. *IEEE Journal of Biomedical and Health Informatics*, 2020.

[43] Ashok Kumar Das, Mohammad Wazid, Neeraj Kumar, Muhammad Khurram Khan, Kim-Kwang Raymond Choo, and YoungHo Park. Design of secure and lightweight authentication protocol for wearable devices environment. *IEEE Journal of Biomedical and Health Informatics*, 22(4):1310–1322, 2017.

[44] Rosa Sánchez-Guerrero, Florina Almenárez Mendoza, Daniel Diaz-Sanchez, Patricia Arias Cabarcos, and Andrés Marín López. Collaborative ehealth meets security: Privacy-enhancing patient profile management. *IEEE Journal of Biomedical and Health Informatics*, 21(6):1741–1749, 2017.

[45] Nafiseh Hosseini, Fatemeh Fakhar, Behzad Kiani, and Saeid Eslami. Enhancing the security of patients' portals and websites by detecting malicious web crawlers using machine learning techniques. *International Journal of Medical Informatics*, 132:103976, 2019.

[46] Robert Mark Seepers, Wenjin Wang, Gerard De Haan, Ioannis Sourdis, and Christos Strydis. Attacks on heartbeat-based security using remote photoplethysmography. *IEEE Journal of Biomedical and Health Informatics*, 22(3):714–721, 2017.

[47] Robert M Seepers, Christos Strydis, Ioannis Sourdis, and Chris I De Zeeuw. Enhancing heart-beat-based security for mhealth applications. *IEEE Journal of Biomedical and Health Informatics*, 21(1):254–262, 2015.

[48] Mark Evans, Ying He, Leandros Maglaras, Iryna Yevseyeva, and Helge Janicke. Evaluating information security core human error causes (is-chec) technique in public sector and comparison with the private sector. *International Journal of Medical Informatics*, 127:109–119, 2019.

[49] Ying He and Chris Johnson. Improving the redistribution of the security lessons in healthcare: An evaluation of the generic security template. *International Journal of Medical Informatics*, 84(11):941–949, 2015.

[50] José Luis Fernández-Alemán, Ana Sánchez-Henarejos, Ambrosio Toval, Ana Belén Sánchez-García, Isabel Hernández-Hernández, and Luis Fernandez-Luque. Analysis of health professional security behaviors in a real clinical setting: An empirical study. *International Journal of Medical Informatics*, 84(6):454–467, 2015.

[51] Marcos A Simplicio, Leonardo H Iwaya, Bruno M Barros, Tereza CMB Carvalho, and Mats Näslund. Secourhealth: a delay-tolerant security framework for mobile health data collection. *IEEE Journal of Biomedical and Health Informatics*, 19(2):761–772, 2014.

[52] Eleni Entzeridou, Evgenia Markopoulou, and Vasiliki Mollaki. Public and physician's expectations and ethical concerns about electronic health record: Benefits outweigh risks except for information security. *International Journal of Medical Informatics*, 110:98–107, 2018.

[53] Tahreem Yasqoob, Haider Abbas, and Narmeen Shafqat. Integrated security, safety, and privacy risk assessment framework for medical devices. *IEEE Journal of Biomedical and Health Informatics*, 2019.

[54] Shen Yin, Juan J Rodriguez-Andina, and Yuchen Jiang. Real-time monitoring and control of industrial cyberphysical systems: With integrated plant-wide monitoring and control framework. *IEEE Industrial Electronics Magazine*, 13(4):38–47, 2019.

[55] Yuchen Jiang and Shen Yin. Recent advances in key-performance-indicator oriented prognosis and diagnosis with a matlab toolbox: Db-kit. *IEEE Transactions on Industrial Informatics*, 15(5):2849–2858, 2018.

[56] Cyril Onwubiko and Austine Onwubiko. Cyber kpi for return on security investment. In *2019 International Conference on Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA)*, pages 1–8. IEEE, 2019.

[57] Chadni Islam, Muhammad Ali Babar, and Surya Nepal. A multi-vocal review of security orchestration. *ACM Computing Surveys (CSUR)*, 52(2):1–45, 2019.

[58] Awni Y Hannun, Pranav Rajpurkar, Masoumeh Haghpanahi, Geoffrey H Tison, Codie Bourn, Mintu P Turakhia, and Andrew Y Ng. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nature Medicine*, 25(1):65, 2019.