# AI Based Directory Discovery Attack and Prevention of the Medical Systems

Ying He[1], Cunjin Luo[2 3], Jiyuan Zheng[1], Kuanquan Wang[4], Henggui Zhang[5]

[1]University of Nottingham, Nottingham, UK
[2]University of Essex, Colchester, UK
[3]Southwest Medical University, Luzhou, China
[4]Harbin Institute of Technology, Harbin, China
[5]The University of Manchester, Manchester, UK

## Abstract

*The medical system has been targeted by the cyber attackers, who aim to bring down the health security critical infrastructure. This research is motivated by the recent cyber-attacks happened during COVID 19 pandemics which resulted in the compromise of the diagnosis results. This study was carried to demonstrate how the medical systems can be penetrated using AI-based Directory Discovery Attack and present security solutions to counteract such attacks. We then followed the NIST (National Institute of Standards and Technology) ethical hacking methodology to launch the AI-based Directory Discovery Attack. We were able to successfully penetrate the system and gain access to the core of the medical directories. We then proposed a series of security solutions to prevent such cyber-attacks.*

## 1. Introduction

Medical systems have renovated healthcare industry. It has been used to improve the quality of healthcare services by enabling early disease detection and diagnosis which is crucial to the success rate of treatment. Other auxiliary medical services such as personalised medicine also have impact on health and wellbeing. However, the increasing use of technology, fast-increasing health data aggregation and the need of intelligent medical systems in the healthcare industry also led to a series of cyber security issues. It is vital to protect the medical data since compromised data may lead to wrong diagnosis results.

The NHS has experienced and is a victim of the high profile ransomware attack, WannaCry, which has affected a large number of organisations around the globe. After the attack, the healthcare organisations started taking security actions to defend against cyber-attacks [1]. This situation has been compounded by the COVID 19, which further challenged the security of medical systems. Healthcare organisations such as the US Department of Health and Human Services, Brno University Hy juospital in Czech Republic, the World Health Organization (WHO) and its Partners have suffered from cyber-attacks during COVID 19 pandemics [2]. It is imperative to protect information in medical systems [3-8].

Ethical hacking which is also called pen-testing can help identify the weaknesses of the system and demonstrate security breaches. It is usually performed following a systematic methodology such as NIST or OWASP in a secure environment. There is existing work demonstrating cyber-attacks towards medical system and presenting security protection solutions [9-14],

In a previous study, we have launched brute force and dictionary attack [14,15]. We have successfully performed the dictionary attack and identified a vulnerability related to authentication “A2 2017-Broken Authentication”, one of the listed OWASP Top 10 vulnerability. In another study, we proposed three attacking activities where AI algorithms are applicable, including password cracking, face recognition and directory discovery [14].

This paper builds on and extends our previous work by proposing a trained algorithm to directory discovery, which is an important type of attacks in ethical hacking.

## 2. Related Work

### 2.1. Healthcare Cyber Security

Medical information can be divided into two different types, sensitive and non-sensitive. The information that is relevant to patients such as diagnosis information is classified as sensitive information. For example, the Electrocardiogram (ECG) waveform data and is classified as sensitive information. Such information has been targeted by attackers using various attack vectors. Previous research investigated into ethical hacking in medical system [14, 15], however, AI based ethical hacking has not been well studied in healthcare

### 2.2. AI Based Ethical Hacking

AI algorithms can help automate the traditional ethical hacking. The key advantage of AI based ethical hacking are more accurate, can be deployed on multiple systems at a time, it does not reply on human experts, and it is more

ISSN: 2325-887X DOI: 10.22489/CinC.2022.401

suitable for repeatable tests [16].

## 3. Methods

### 3.1. NIST Pen-Testing Methodology

The ethical hacking carried out adopted the NIST pen-testing methodology. A complete ethical hacking covers four stages including information gathering, discovery, attack and reporting. [16]. Figure 1 below describes the work flow of NIST pen-testing methodology.

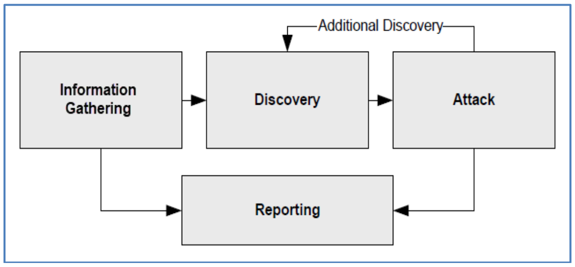

Figure 1. NIST Pen-testing Methodology [16]

### 3.2. The Simulated Medical System

To set up the simulated environment, VirtualBox is used as the virtualization tool. We set up a simulation system on the virtual platform through implementing an open-source medical system, OpenEMR. The context of the study is based upon this simulation platform.

## 4. AI based Ethical Hacking

In a previous study, we have launched brute force and dictionary attack [14, 15]. We have successfully performed dictionary attack and identified a vulnerability related to authentication “A2 2017-Broken Authentication”, which is one of the listed OWASP Top 10 vulnerability. In another study, we proposed three attacking activities where AI algorithms are applicable, including password cracking, face recognition and directory discovery [15].

This section proposes applying a trained algorithm to directory discovery.

Directory discovery is an important process in ethical hacking [16]. Spidering can discover directories of web applications by sending HTTP requests and attempting to access other directories by extracting hyperlinks from the response HTML file. Traditional method for directory discovery is the brute force method. In this method, pen-testers create a wordlist and pick URLs in order from the wordlist and then send HTTP requests to the targeted web application. If the response code does not belong to error codes, then the directory is discovered successfully.

This paper proposes a trained algorithm for directory discovery by adopting semantic clustering of sentences. By doing this, directories with the same naming conventions and file extensions are likely to be grouped together so that if one directory within a certain cluster is valid, then other directories within the same cluster are highly likely to be valid. Figure 2 shows the global view of the algorithm.

The wordlist used for directory discovery in this project is composed of absolute paths of directories from three different web applications which are openEMR, DVWA, and Joomla. Firstly, we carried out data-preprocessing process to transform the collected data into sentences. Then, we transferred these into numerical representation by sentence embedding. Finally, the encoded data is fed into the clustering algorithm to create semantic clusters.

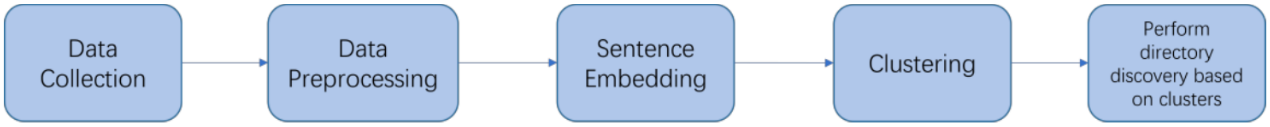

Figure 2. Global view of the algorithm

## 5. Security Solutions

An important strategy is to enhance the authentication of openEMR. Currently, openEMR directories can be accessed directly by unauthenticated users which can lead to sensitive information disclosure. He, et al, identified that the openEMR session and cookies management is poorly designed which enables cookie stealing [15]. It is advised that openEMR should enforce a secure authentication and session management mechanism which restrict passwords' minimum size and complexity; authentication is required for every directory; users' session should be protected, and session should expire when the browser is closed. For wherever possible, multi-factor authentication is always recommended.

## 6. Conclusion and Future Work

Built on previous research on the ethical hacking of the OpenEMR system, this study presented AI-based directory discovery attack and its cyber defence solutions. It has implications for further research into the AI based cyber security attacks and its defence in healthcare. Since the OpenEMR has been widely used in health organisations worldwide, the findings have implications for security professionals in healthcare to prevent against AI based security attacks.

Future work will consider integrating the simulation platform with an intelligent cardiac diagnosis systems with complex computational models such as the arrhythmia detection and classification in ECG data [21-24] and see how the core AI component of diagnosis models can be penetrated and protected. Future work will also consider adopting novel security solutions from other industries to counteract cyber-attacks in healthcare [25-30].

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 61803318).

## References

[1] Martin, Guy, Saira Ghafur, James Kinross, Chris Hankin, and Ara Darzi. "WannaCry—a year on." (2018): k2381.

[2] He, Ying, Aliyu Aliyu, Mark Evans, and Cunjin Luo. "Health care cybersecurity challenges and solutions under the climate of COVID-19: Scoping review." Journal of Medical Internet Research 23, no. 4 (2021): e21747.

[3] Luo, Cunjin, Kuanquan Wang, and Henggui Zhang. "Modeling the effects of amiodarone on short QT syndrome variant 2 in the human ventricles." In 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 4273-4276. IEEE, 2017.

[4] Luo, Cunjin, Kuanquan Wang, and Henggui Zhang. "Modelling the effects of quinidine, disopyramide, and E-4031 on short QT syndrome variant 3 in the human ventricles." Physiological Measurement 38, no. 10 (2017): 1859.

[5] Luo, Cunjin, Kuanquan Wang, and Henggui Zhang. "Modelling the effects of chloroquine on KCNJ2-linked short QT syndrome." Oncotarget 8, no. 63 (2017): 106511.

[6] Luo, Cunjin, Kuanquan Wang, and Henggui Zhang. "In silico assessment of the effects of quinidine, disopyramide and E-4031 on short QT syndrome variant 1 in the human ventricles." PloS One 12, no. 6 (2017): e0179515.

[7] Luo, Cunjin, Kuanquan Wang, and Henggui Zhang. "Effects of amiodarone on short QT syndrome variant 3 in human ventricles: a simulation study." Biomedical Engineering Online 16, no. 1 (2017): 69.

[8] Luo, Cunjin, Kuanquan Wang, and Henggui Zhang. "Modeling the effects of amiodarone on short QT syndrome variant 2 in the human ventricles." In 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 4273-4276. IEEE, 2017.

[9] Evans, Mark, Ying He, Leandros Maglaras, and Helge Janicke. "Heart-is: A novel technique for evaluating human error-related information security incidents." Computers & Security 80 (2019): 74-89.

[10] He, Ying, Ruben Suxo Camacho, Cunjin Luo, and Henggui Zhang. "Security Defense Strategy for Cardiac Medical Diagnosis System (CMDS)." In 2019 Computing in Cardiology (CinC), pp. Page-1. IEEE, 2019.

[11] Evans, Mark, Ying He, Cunjin Luo, Iryna Yevseyeva, Helge Janicke, and Leandros A. Maglaras. "Employee perspective on information security related human error in healthcare: Proactive use of IS-CHEC in questionnaire form." IEEE Access 7 (2019): 102087-102101.

[12] Evans, Mark, 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 (2019): 109-119.

[13] He, Ying, and Chris Johnson. "Challenges of information security incident learning: an industrial case study in a Chinese healthcare organization." Informatics for Health and Social Care 42, no. 4 (2017): 393-408.

[14 He, Ying, 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, no. 11 (2015): 941-949.

[15] He, Ying, Kun Ni, and Cunjin Luo. "Attacking Pathways of Health Information System (HIS)." In 2021 Computing in Cardiology (CinC), vol. 48, pp. 1-4. IEEE, 2021.

[16] Baloch, Rafay. Ethical hacking and penetration testing guide. CRC Press, 2017.

[17] Wang, Ding, and Ping Wang. "Offline dictionary attack on password authentication schemes using smart cards." In Information security, pp. 221-237. Springer, Cham, 2015.

[18] Zuo, Chaoshun, Qingchuan Zhao, and Zhiqiang Lin. "Authscope: Towards automatic discovery of vulnerable authorizations in online services." In Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, pp. 799-813. 2017.

[19] Wang, Ding, and Ping Wang. "Two birds with one stone:

Two-factor authentication with security beyond conventional bound." IEEE Transactions on Dependable and Secure Computing 15, no. 4 (2016): 708-722.

[20] Gao, Haichang, 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, no. 7 (2017): 1671-1685.

[21] Luo, Cunjin, Kuanquan Wang, and Henggui Zhang. "Effects of amiodarone on short QT syndrome variant 3 in human ventricles: a simulation study." Biomedical Engineering Online 16, no. 1 (2017): 69.

[22] Luo, Cunjin, Kuanquan Wang, and Henggui Zhang. "Effects of island-distribution of mid-cardiomyocytes on ventricular electrical excitation associated with the KCNQ1-linked short QT syndrome." In 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 3684-3687. IEEE, 2017.

[23] Luo, Cunjin, Kuanquan Wang, and Henggui Zhang. "Functional effects of island-distribution of mid-cardiomyocytes on Re-entrant excitation waves in the KCNQ1-linked short QT syndrome." In 2016 Computing in Cardiology Conference (CinC), pp. 933-936. IEEE, 2016.

[24] Luo, Cunjin, Kuanquan Wang, Ming Yuan, Zhili Li, Qingjie Wang, Yongfeng Yuan, Qince Li, and Henggui Zhang. "Effects of amiodarone on ventricular excitation associated with the KCNJ2-linked short QT syndrome: Insights from a modelling study." In 2015 Computing in Cardiology Conference (CinC), pp. 1093-1096. IEEE, 2015.

[25] Tzokatziou, Grigoris, Leandros A. Maglaras, Helge Janicke, and Ying He. "Exploiting SCADA vulnerabilities using a human interface device." Int J Adv Comput Sci Appl (2015): 234-241.

[26] Ayres, Nicholas, Leandros A. Maglaras, Helge Janicke, Richard Smith, and Ying He. "The mimetic virus: a vector for cyberterrorism." International Journal of Business Continuity and Risk Management 6, no. 4 (2016): 259-271.

[27] Zamani, Efpraxia, Ying He, and Matthew Phillips. "On the Security Risks of the Blockchain." Journal of Computer Information Systems (2018): 1-12.

[28] Wood, Andy, Ying He, Leandros Maglaras, and Helge Janicke. "A security architectural pattern for risk management of industry control systems within critical national infrastructure." (2017).

[29] Evans, Mark, Leandros A. Maglaras, Ying He, and Helge Janicke. "Human behaviour as an aspect of cybersecurity assurance." Security and Communication Networks 9, no. 17 (2016): 4667-4679.

[30] He, Ying, Chris Johnson, Yu Lu, and Yixia Lin. "Improving the information security management: An industrial study in the privacy of electronic patient records." In 2014 IEEE 27th International Symposium on Computer-Based Medical Systems (CBMS), pp. 525-526. IEEE, 2014.

Cunjin Luo, PhD
School of Computer Science and Electronic Engineering
University of Essex,
Colchester, CO4 3SQ
UK
cunjin.luo@essex.ac.uk