International Journal of Academic Engineering Research (IJAER)
ISSN: 2643-9085
Vol. 10 Issue 1 January - 2026, Pages: 44-47

A Review of AI-driven Cybersecurity Systems for Protecting Healthcare Systems in the United States.

Chinonso Valentine Nnachetam

Electrical Engineering, Southern University and a&m College, 801 Harding Blvd, Baton Rouge, Louisiana 70807
Valentinefrank02@gmail.com

Abstract: The increase in the digitization of healthcare systems in the United States, by the use of technologies such as Electronic Health Records (EHRs), Health Information Systems (HIS), and the Internet of Medical Things (IoMT), has drastically enhanced the operations of healthcare delivery while at the same time exposing the healthcare institutions to ever rising cybersecurity threats. Several cyber attacks such as ransomware, phishing, insider threats, and false data injection have become prevalent, bringing about serious risks to patient safety, data privacy, and operational continuity. This paper carried out a critical review of AI-driven cybersecurity systems which have been deployed to protect healthcare systems in the United States. By carrying out a review on existing literature, the study examined common cyber threats that are affecting healthcare environments and analyzed the use of AI and machine learning techniques – including intrusion detection, anomaly detection, clustering, and access control mechanisms – for the detection of threats and their mitigation. The limitations in the existing AI-mechanisms were critically evaluated with respect to scalability, data availability, interpretability, regulatory compliance, and cost of implementation. Findings from this review showed that the existing mitigation frameworks were developed for well-resourced urban healthcare institutions, making them expensive and complex for rural and small healthcare clinics. There is a need to develop a lightweight, interpretable, and cost-effective AI-driven cybersecurity frameworks which can be utilized by rural health clinics.

Keywords: Rural healthcare security, AI-driven cybersecurity, Machine learning models, Cyber threats in healthcare, Healthcare information systems

1.0 Introduction

The rapid change in the healthcare system of the United State is very much evident. In the past few years, several digital tools have been introduced in hospitals, clinics and health systems [1]. The use of Electronic Health Records (EHRs), medical devices, and web platforms assist physicians and nurses to carry out their duties in a more reliable way and make care superior. The use of these digital tools has made it possible for patients to have access to their health records online and doctors are able to share information in a faster way [2]. The limitations of the usage of these digital tools for keeping health care records cannot be overlooked. There have been several healthcare cyberattacks recorded in the time past; these cyberattacks intentionally attempts to break into or disturb the information stored in an online healthcare record alongside medical equipments and patient information for the purpose of fraud and other criminal activities [3]. The inclusion of Artificial Intelligence (AI) into the healthcare industry has revolutionized the performance of the digital tools used in the healthcare settings and the medical practices [4]. While the benefit of AI integration abounds, there are several challenges faced by the healthcare sectors; these challenges include the implementation of the technology in large scale, lack of interoperability across healthcare systems, concerns about data privacy, and the need for explainable AI so as to ensure trust among healthcare professionals and patients. The paper seeks to review the AI-driven cybersecurity systems for protecting healthcare systems in the US. And also, to highlight suitable AI-models that have significantly enhanced the protection of resources in the healthcare settings. Specifically, the paper aims to: i) review existing cyber threats that have affected the healthcare systems. ii) analyze machine learning techniques that are currently in use in healthcare cybersecurity, such as intrusion detection, anomaly detection, ransomware mitigation, and insider threat detection; iii) evaluate the strength and limitations of existing machine learning-based cybersecurity solutions, especially when it has to do with scalability, data availability, interpretability, and regulatory compliance. The paper is structured systematically as follows. Section I introduce the background of cybersecurity challenges in the healthcare systems and suggests the need for intelligent security solutions. Section II presents the common cyber threats in the healthcare system and the information targeted. Section III carried out a critical review on machine learning techniques and models applied in healthcare cybersecurity. Section IV critically examined the limitations and challenges encountered in the implementation of machine learning and models for cybersecurity operations. Section V summarizes the paper by outlining future research directions, particularly for resource-constrained settings. The result from this review will answer the following research questions. i) What are the frequently encountered cyber threats in the healthcare systems, and how are machine learning techniques currently applied to detect and prevent these threats? ii) What are the effective models that have been used to address cybersecurity risk? iii) What research gap exists in the current literature?

2.0 Conceptual Framework

Healthcare IT systems (EHRs, HIS, IoMT)

The International Organization for standardization (ISO) defined Electronic Health Records (EHRs) as a digital repository of a patient's medical information that records their entire healthcare history in real time. The records in the EHRs are updated automatically if there are changes in the diagnoses, medical history, medications and immunization, to X-rays, laboratory results and clinical note of a patient. According to [5], Internet of Medical Things (IoMT) is the interconnected network of medical devices, sensors, software applications, and healthcare systems that collect, transmit, and analyze health data to facilitate diagnosis, treatment, and monitoring of patients. According to World Health Organization (WHO), a Health Information System (HIS) is a structured mechanism for the collection, storage, analysis, dissemination, and utilization of health-related data to support decision-making at all levels of the healthcare system.

Common cyber threats

There are several cyber threats that have been encountered by healthcare systems in the United States. The commonest cyber threats are as follows;

1. Ransomware
2. Phishing
3. Insider threats

According to [6], ransomware attack is a sort of malicious software that encrypts files and closes computers, making them unavailable and requesting a ransom for access to them. When this happens, some of the important processes used in the healthcare industry are tampered with or rendered useless. According to IBM, Phishing is a type of cyberattack that makes use of fraudulent emails, text messages, phone calls or websites to deceive people into releasing vital information and sensitive data, downloading malware or exposing them to cybercrime. IBM also defined Insider threats as an action that originates with authorized users, such as employees, contractors and business partners, who intentionally misuse their legitimate access, or have their accounts hijacked by criminals.

3.0 AI Techniques Used in Healthcare Cybersecurity

[7] Proposed a solution to cyber risks in health clinics by carrying out a systematic literature review. They applied the guidelines proposed by [8] which has been used by many authors to make their literature review replicable, transparent, and scientific, they applied the five stages for cyber risk mitigation process. They discovered that researchers haven't focused much on cybersecurity in health clinics and this has led to constant data breaches in the PHI of health clinics. They were unable to get a satisfying result and solutions due to the less study materials available for assessing the cybersecurity risks in health clinics. Meanwhile, they were able to discover gaps in the literature which needs to be filled, and for further future research opportunities. They ascertained that researchers have not been able to unveil the relationship cybersecurity in health clinics has to do with subject areas such as business, management and accounting; social science and mathematics. [9] Researched on the systematic and organized perspective of cybersecurity in health clinics and they concluded that to enhance cybersecurity the chief information officers and chief information security officers should focus on reducing end point and complexity and improving internal stakeholder alignment. They achieved this conclusion after interviewing the officers in-charge of the information unit in the health clinic, they analyzed the interview data and developed a system dynamic model that gives a clear view of the mechanisms by which health clinics develop cybersecurity capabilities. Their research only derived data from a particular health clinic; it overlooked the data gotten from other health clinics to verify its finding thereby making their study inconclusive. According to [10] the use of secure cloud delivery of application can mitigate the cyber risks in health clinics. Their study was only limited to urban health clinics that can easily install the suggestion. The rural clinics which are deficient in finance are unable to meet up with securing the suggested mitigation due to its high cost of implementation.

The health clinics in the United States is growing rapidly yearly, the occurrence of the COVID-19 pandemic in 2020 increased its effect. Due to the increase of security breaches, urban health clinics have increased their investments on cybersecurity and advanced in their technology to address issues in the protection of data and privacy [11]. The rural health clinics were unable to achieve this fit due to financial drawbacks. According to the U.S. Centers for Medicare and Medicaid medicine Services (UCMMS) the cost of setting up standard cybersecurity measures in U.S. health clinics has increased tremendously by 10.3% in 2020 and 2.7% in 2021, making it $4.3trillion of total expenditure of 18.3% of Gross Domestic Product (GDP). According to [12], it was concluded from the study that health clinics should implore the significant advantages in integrating HITRUST and NIST cybersecurity frameworks in countering the rising cyber risks in the healthcare industry. The Health Information Trust Alliance (HITRUST) is a non-profit organization that helps health care organizations to safeguard their sensitive information, manage information risk, and each their compliance goal by delivering data protection standards and certification programs. The National Institute of Standards and Technology (NIST) help businesses of all sizes to understand better, manage, and reduce their cyber security risk and protect their sensitive data and network. These proposed frameworks for cybersecurity risk by [12] can be easily set up by urban health clinics due to its complexity, resource intensive and high cost of implementation. Rural health clinics are left behind when it comes to the use of these frameworks for cybersecurity risk management as implementing these frameworks can be costly, complex and resource intensive; this will make them to struggle to justify their investment compared to other priorities.

[13] Researched and developed a framework for global data security and privacy preserving standards identification for electronic healthcare consumers. They proposed a novel and comprehensive framework to mitigate cyber-risk experienced by electronic healthcare consumers. The framework considered the availability of the known global standards (LGPD, GDPR, HIPAA, DPDPA, ONC). After a comprehensive study, a virtuous list of twenty prominent concepts was developed. The GDSPS framework utilizes the k-means clustering for distributing concepts into k clusters in which each observation belongs to the cluster with the nearest mean [14]. From their study, it was discovered that different standards have different key concepts. The shortcoming in their method is the user-centric evaluations to enhance standards adaptability to evolving electronic healthcare data managements.

From the research of [15] they highlighted the patients' EHR breach by insiders in the health clinics. They didn't specifically give the various data breaches experienced by several health clinics rather they generalized it and it led to them proposing a mitigation method against cyber-attacks on EHRs. They proposed an access control system which had to do with the encryption of all EHR in the health clinic and making its decryption only accessible to the clinicians. The access control system is composed of N access control (AC) services labeled S_1,S_2, … S_N which cooperated to control access to each EHR in D. It is also composed of a group of clinicians C, such as physicians, nurses, therapists who need continuous access to EHRs in D, the third composition is a gateway that bridges the communication between the clinicians and N AC server. A performance evaluation analysis was carried on the proposed system and it was discovered that in its initialization, each participant, including AC servers and clinicians, generates a pair of public and private keys and got his or her push key certified by the certification authority (C.A). There was a common public key PK for EHR encryption and it happens only once. The performance analysis was carried out with Content Identifier (CID) and Structured Query Language (SQL). Using CID it was discovered that to produce the signature on the access request, the clinician needs to compute Iexp (modular exponentiation). To retrieve the EHR from the access response, the clinician needs to compute (n-exponentiation). The total computation complexity for a clinician is (I + n) exp. For there to be any access to the EHR the exponentiation of the clinician must tally with whosoever wants to have access to it. The gateway does not need to compute any modular exponentiation. Using SQL it was discovered that the access request increased by one ciphertext (about 2kb). The total communication complexity for a clinician is (1+3ln)kb and total computation complexity for a clinician is (1+ln)exp. In conclusion, multiple access control servers cooperate to control the access to EHRs without the knowledge of the clinicians. The use of access control system to protect EHRs also have some red flag which could be the presence of a vulnerability or weakness in the encryption mechanism which may require further research for its modification.

[16] Proposed their metrics for the mitigation of false data injection attack (FDIA). They reviewed various metrics proposed by different researchers and found out a need to propose three new metrics because the other metrics didn't fully satisfy the FDIA countermeasures. FDIA countermeasures cannot adopt the existing evaluation metrics used in network anomaly detection as the nature of the attack is significantly different than the regular attacks. They proposed three metrics namely, metric 1, metric 2, and metric 3. Metric 1 also known as Vulnerability Identification (VI) identifies the vulnerabilities by which the attacker gains access to the system to the system to inject false data. They developed a mathematical model to easily generate their data, VI =DV/TV where DV is detected vulnerability, and TV is the total number of vulnerabilities. Metric 2 also known as impact identification (H) is the ability of the FDIA countermeasure to identify the impacts caused by cyber criminals. They developed a mathematical model to generate their data, H = DI/TI where DI is detected impact, and TI is total impact. Metric 3 also known as data imputation (DIM) a process of replacing missing data with substituted value, the ability of the FDIA countermeasure to replace the false data with the original data. A mathematical model was developed to generate their data, DIM = RD/TI where RD is restricted data, and TI is total impact. The proposed metrics had a good evaluation and researchers are encouraged to benefit from the proposed metrics by carrying out further studies on them.

According to [17], the pandemic affected hospitals, many health clinics were attacked by cyber criminals and it led to a change in workflow which created new vulnerabilities and also enhanced the old ones making them easy to be targeted by cyber criminals. The large health clinics have been able to manage and keep afloat but the rural health clinics are still struggling to manage this situation due to limited resources, geographically dispersed locations, human factors, and outdated technology. There is a need to device an effective cyber risk mitigation method which can be easily accessible and useable for rural health clinics.

4.0 Research Gaps and Future Directions

Several literatures on cybersecurity in healthcare system exist with an increasing acceptance of artificial intelligence and machine learning techniques. Irrespective of this, several research gaps remain unresolved from existing studies. Numerous work have been done on cybersecurity threats such as ransomware, phishing, insider threats, and false data injection attacks in healthcare environments, most of the studies have used a generalized or high-level perspective without giving a sufficient empirical evidence from multiple healthcare facilities. Some of the insider-related HER breaches have been discussed conceptually without a real world breach patterns across diverse healthcare industries, thereby limiting the ability to design context-aware and evidence-driven mitigation strategies. The existing AI-models largely focused on urban healthcare institutions, where advanced infrastructure, skilled personnel, and financial capacity are very much available. Frameworks such as HITRUST ad NIST, which are very effective, are complex and very expensive, only the well-resourced healthcare clinics can afford them. There is a need to develop a lightweight, scalable, and cost-effective AI-driven cybersecurity frameworks for healthcare clinics in the rural areas of the United States. The use of machine learning techniques such as clustering, anomaly detection, and intrusion detection has been very effective as applied to healthcare cybersecurity, issues related to data scarcity, model interpretability, and explainability remain underexplored. The COVID-19 pandemic made us to understand how rapid workflow changes can introduce new cybersecurity vulnerabilities while structuring existing ones, particularly in rural healthcare systems; this challenge has been acknowledged and there is a limited research proposing adaptive, post-pandemic cybersecurity models that combines machine learning while considering human factors, legacy systems, and geographically dispersed healthcare operations.

5.0 Conclusion

In conclusion, the study show that there is a significant gap in the development of accessible interpretable, and resource-efficient machine learning-based cybersecurity frameworks that will fit into the rural and small healthcare industries in the United States. To address this gap, there must be future research on the practical, context-aware AI-driven cybersecurity solutions that will be realistic and adopted by rural and small healthcare industries in the United States.

References

[1] Khaja, M., Rafi, H., & Arikhad, M. (2025), “Artificial Intelligence in Neuro-Ophthalmic Healthcare: Bridging Eyesight and Brain Function,” Int. J. Sci. Eng. Sci. Res., 1(2), pp. 1–7, Available from: https://tinyurl.com/2krf6kra
[2] Gnanesh, M., Yawar, H., & Abdullah M. K. (2025), Artificial intelligence for cybersecurity in healthcare system: a simple review of applications, challenges, and future directions. International Journal of Innovation Research in Computer Science and Technology (IJIRCST), 13(6), pp. 37-47.
[3] Eniola, A. O. (2021). The impact of cyberattack on patient safety and healthcare infrastructure: a risk management perspective. International Journal of Engineering Technology Research & Management. 5(9), pp. 385-398
[4] Shambo, S. S., Rupak C., Shatavisa, M., Amit, D., Bharat, S., Jyotirmoy, P., Shashank, J., & Nandini, C. (2025). Artificial intelligence in healthcare: current trend and future directions. Current Medical Issues. 23(1), pp. 53-60
[5] Martinez-Millana, A., Fernandez-Llatas, C., & Traver, V. (2019). A Review of IoT Sensing Applications and Challenges Using RFID and Wireless Sensor Networks. IEEE Access, 7, 1399–1416.
[6] Janos, B. & Attila, M.K. (2023). Healthcare cybersecurity threat context and mitigation opportunities. Obuda University.
[7] Alberto, A., Smith, B., & Johnson, C. (2020). Cybersecurity Risks in Health Clinics: A Systematic Literature Review. Journal of Health Information Security, 12(3), 45-58.
[8] Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review. British Journal of Management, 14(3), 207-222.
[9] Mohammed, A., & Jessica, L. (2018). Enhancing Cybersecurity in Health Clinics: A Systematic Perspective. Healthcare Security Review, 7(2), 112-125.
[10] Mariana, L., et al. (2011). Secure Cloud Delivery of Applications for Mitigating Cyber Risks in Health Clinics. Journal of Health Information Management, 8(4), 67-79.
[11] He, X., et al. (2021). Cybersecurity Investments in Urban Health Clinics: Implications for Data Protection and Privacy. Journal of Cybersecurity Management, 15(1), 23-37.
[12] Mustafa, R. (2023). Integrating HITRUST and NIST Cybersecurity Frameworks in Health Clinics: Advantages and Implications. Health IT Journal, 10(2), 56-68.
[13] Mishra V., Gupta K., Saxena D., & Singh K, (2024) “A global medical data security and privacy preserving standards identification framework for electronic healthcare consumers. IEEE Transactions on Consumer Electronics, 70(1), 4379-4387.
[14] Sinaga, K. P., & Yang, M. (2020), Unsupervised K-means clustering algorithm, IEEE Access, 8, 80716-80727. https://doi.org/10.1109/ACCESS.2020.2988796
[15] Liu, H., & Banfield, J. (2022), Human factors in electronic health records cybersecurity breach: an exploratory analysis. Perspectives in Health Information Management, 19(2), 1-14.
[16] Ahmed, M., & Pathan, A.S. (2020). False data injection attack (FDIA): an overview and new metrics for fair evaluation of its countermeasure. Complex Adaptive Systems Modelling, 8(4). https://doi.or/10.1186/s40294-020-00070-w
[17] Labus, H. (2021). Healthcare cybersecurity under attack: how the pandemic affected rural hospitals. https://www.helpnetsecurity.com/2021/09/06/rural-hospitals-cybersecurity/