Cyber Security Tools and Techniques suggested for mitigating insider attacks in the healthcare system

Dr.Mervin Retnadhas Mary[1,2,] [a], *Dr.Pawan Whig*[3,] [b]

[1]Postdoctoral Researcher, Lincoln University College, Petaling Jaya-47301, Selangor, Malaysia.
[2]Assistant Professor, Department of Information Technology, Saudi Electronic University, Saudi Arabia.
[3]Dean Research, Vivekananda Institute of Professional Studies-TC, New Delhi
mervinvijay@gmail.com , pawan.whig@vips.edu

Abstract: Cybersecurity plays a vital role today. Due to the increase in digitalization, numerous threats are possible nowadays. Many fields, like education, the military, the banking sector, etc., need to protect a huge amount of data. The health sector is one of the major fields where the patient's data must be secured in a protective way. All the threats and vulnerabilities in the health sector must be identified, and the researchers must give more importance to mitigating these threats. It also focused on the insider attack, which is considered a harmful attack for the organisation. The protection of the data can be achieved through cybersecurity tools. This research paper briefly identifies the cybersecurity tools and techniques that help to mitigate and provide more support to protect our data. This review paper addresses the tools, techniques, challenges, and state of the art of cybersecurity. Also, it addresses the collaboration with Artificial Intelligence (AI), Natural Language Processing (NLP), and Machine Learning (ML) techniques to detect and automate insider attacks in the health sector. After the analysis of tools in the different fields, the implementation was started with the Support Vector Machine, the Decision tree algorithm, Random Forest, AdaBoost, and Naive Bayes.

Keywords: Cybersecurity, Health Care, Cyberattacks, Insider Attacks, Health Recommendation Systems, Health Care System, Countermeasures, Prevention techniques, Mitigation and Taxonomy, Artificial Intelligence and Machine Learning.

---

a https://orcid.org/0000-0002-8587-0624
b https://orcid.org/0000-0003-1863-1591

Introduction

Data protection and privacy are great concerns in the digital world, and the concern of protecting data and privacy is increasingly growing[2]. Due to the numerous cyberattacks in the various fields, Cybersecurity has playing major role in identifying the vulnerabilities and threats. Many researchers, academicians, and organisations are working together to find the threats and focus on maintaining the CIA (Confidentiality, Integrity and Availability). Particularly, more attention is required for the insider

attacks, which are considered to be increasingly harmful attacks. An insider attack is possible in any organisation, and the research focuses on the healthcare system[4]. If the records have been modified, then it may be a threat to human life. To protect our system from insider attacks, cybersecurity is the most important solution. We are trying for the solution not only with cybersecurity, but with the collaboration with Artificial Intelligence (AI), Natural Language Processing (NLP) and Machine Learning (ML) techniques. These collaborated model allow for a systematic and reliable vulnerability discovery process and ensure high security in the field. By tailoring support to individuals and healthcare providers, various intelligence technologies incorporated to ensure patient acceptance and trust necessitate addressing substantial privacy and security concerns inherent in the healthcare system.

The collected dataset requires preprocessing due to the missing values, noise, inconsistency and errors. When the preprocessing is executed on the data, it helps to clean the data, and the data will become an organised and standardised format. We can improve the cleansing with the support of NLP and ML techniques, and the required feature will be extracted from the dataset. Furthermore, it can be analysed using various tools and techniques. The data can be modified, altered or removed by an insider of the organisation, and the researcher has to pay more attention to the insider attack. There are 'n' number of forensic tools available to identify the threats, and tools can be selected by the investigator based on the requirement.

Related work

Adeyinka Ayodeji Mustaphaa et al. [3] summarised the Current Trends and Innovations in Cybersecurity Technologies: A Comprehensive Review. They gave the in depth analysis on the recent cybersecurity trends, tools, and best practices. This review paper also found solutions for the cybersecurity threats by integrating Machine Learning (ML), blockchain, cloud security, and Artificial Intelligence (AI) techniques. They focused on quantum computing, Internet of Things security, and deepfakes for strengthening cybersecurity postures in today's digital era.

Juswin Sajan John et al[11] elaborated their research on Categorizing Mental Stress: A Consistency-Focused Benchmarking of ML and DL Models for Multi-Label, Multi-Class Classification via Taxonomy-Driven NLP Techniques. They discussed various techniques, including Support Vector Machines (SVM), Random Forest (RF) and Long Short-Term Memory (LSTM) algorithms incorporating various feature combinations involving Term Frequency-Inverse Document Frequency (TF-IDF) and Latent Dirichlet Allocation (LDA). This model utilizes Support Vector Machines (SVM) with TF-IDF (Term Frequency-Inverse Document Frequency) representations as feature vectors.

Aswini et al. [9] described various tools and techniques used in the field of cybersecurity. They mainly focused on the Rootkit malware, which resides in the kernel of the operating system. They are suggesting Anti-rootkit tools include Spy DLL Remover, Sanity Check, and Root Repeal to handle different kinds of rootkit malware.

Key Contribution

The research focused on identifying insider attacks in the healthcare system by integrating the cybersecurity tools, NLP, ML and AI techniques. The collaboration of the following tools and techniques mitigates the insider attack in a better way. TF-IDF (Term Frequency-Inverse Document Frequency) - It is the most important statistical method in Natural Language Processing to analyse the importance of a word in a document to the available corpus. It also helps remove stop words automatically. IDF reduces the weight of common words while increasing the weight of the rare words. Equations 1 and 2 defines the formula for finding out the similarity of TF-IDF.

$$TF(t,d) = \frac{No.of\ times\ 't' appears\ in\ the\ documnet\ 'd'}{Total\ no\ of\ terms\ in\ the\ document\ 'd'}$$

...eqn (1)

$$IDF(t,D) = \frac{Total\ no\ of\ douments\ in\ the\ corpus'D'}{Number\ of\ documnets\ containing\ the\ term\ 't'}$$

...eqn (2)

Method, Experiments, and Results

The two datasets from CERT and the Government hospital dataset released by the Saudi Red Crescent Authority have been taken for the analysis, including the logon activity of the user and the heart diseases of the patient. The dataset was preprocessed, and the pattern was analysed in the logon dataset to check for the vulnerability of insider attacks. The heart disease dataset has been analysed with SVM and a Decision Tree method, and checked for accuracy. SVM predicted correctly with 82.4 %, and the decision tree prediction was 75.6%, which was reflected in the figure 1.

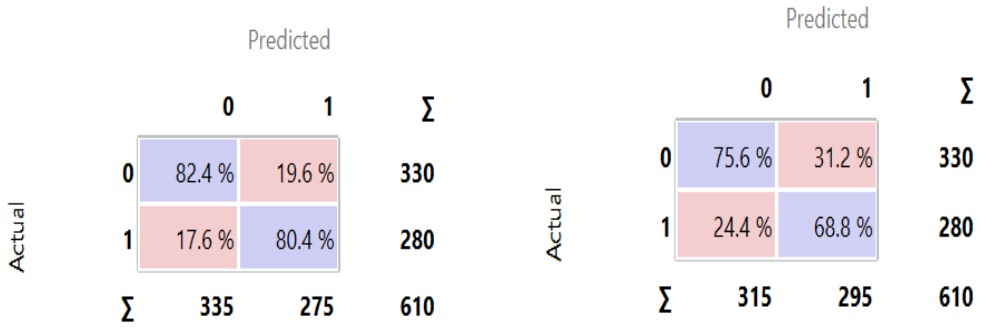

Figure 1. Analysis based on SVM and Decision Tree

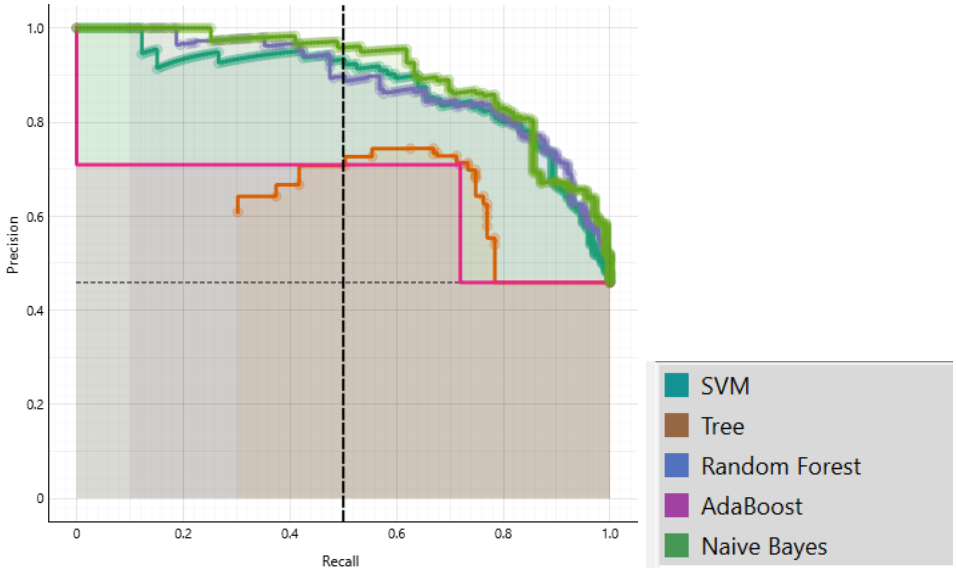

Figure 2. Graphical Representation of the Performance Analysis

Finally, the performance metrics were analysed for both the techniques, SVM and the decision tree. The results were reflected in the table 2, and the corresponding graph is implemented with precision and recall. Figure 2 shows the graphical representation of this performance analysis.

Conclusions

Cybersecurity is the process of protecting and safeguarding computer systems, networks, and data from cyber threats that attack the confidentiality, integrity, and accessibility of information systems. It is essential for protecting the safety of individuals and organisations since they highly depend on digital technologies[2]. In this paper, we explore a new angle of a collaborative approach to identify the threats and vulnerabilities in healthcare systems. Based on the research, the tools and techniques have been identified in the field of Cybersecurity, Machine Learning, Natural Language Processing and Cyber Forensics. The datasets are tested with the Support Vector Machine and the Decision Tree method. The results were analysed with the performance metrics. Further more investigation has to be implemented to remove all vulnerabilities in the healthcare system. We plan to reconstruct a broader class of attacks in future work, and the implementation has to be extended to achieve the goal.

References

1. Dr. Suneel Pappala, Dr. Kanigiri Suresh, " Advancements in Ethical Hacking Techniques and Tools: Strengthening Cybersecurity Resilience", International Journal for Multidisciplinary Research (IJFMR), Volume 7, Issue 2, March-April 2025. Pp.1-12.
2. Adeyinka Ayodeji Mustaphaa, Rabbiat Jumai Alhassanb, Thomas Anafeh Ashic, " Current Trends and Innovations in Cybersecurity Technologies: A Comprehensive Review", Journal of Scientific and Engineering Research, 2024, 11(5):100-112.

3. Dr.P.Venkadesh, Dr.S.V.Divya, B.Sudarson, S.Sowmiyan,S.Sebin,N.Vishal," A Comprehensive Analysis Of Cyber Crimes And Cyber Security Tools", International Journal of Creative Research Thoughts (IJCRT), Volume 12, Issue 11 November 2024, pp.599-608.
4. Rachid Ait Maalem Lahcen, Bruce Caulkins, Ram Mohapatra and Manish Kumar, "Review and insight on the behavioral aspects of cybersecurity", Springer Cybersecurity, 2020. pp. 1-18. https://doi.org/10.1186/s42400-020-00050-w.
5. Fabian Böhm , Florian Menges and Günther Pernul," Graph-based visual analytics for cyber threat intelligence" , Springer Cybersecurity, 2020. pp. 1-19. https://doi.org/10.1186/s42400-018-0017-4
6. Mohammad Wazid ,Amit Kumar Mishra ,Noor Mohd ,Ashok Kumar Das," A Secure Deepfake Mitigation Framework: Architecture, Issues, Challenges, and Societal Impact", Cyber Security and Applications, 2024, 100040,pp 1- 9, https://doi.org/10.1016/j.csa.2024.100040.
7. Muthukumar Manickam,Ganesh Gopal Devarajan ∗," A three-factor mutual authentication scheme for telecare medical information system based on ECC", Cyber Security and Applications, 2024, 10031,pp 1- 9. https://doi.org/10.1016/j.csa.2024.100035.
8. Ashwini D. Mate and Dr.D.R.Ingle, "Cybersecurity Tools and Methods", International Conference Proceeding ICGTETM Dec 2017, pp. 1-6, http://doi.one/10.1727/IJCRT.17175.
9. Liang Zhang, Andrei Lobov, "Semantic Web Rule Language-based approach for implementing Knowledge-Based Engineering systems", Advanced Engineering Informatics, 62(2024,102587, pp. 1-17, https://doi.org/10.1016/j.aei.2024.102587.
10. Juswin Sajan John , Boppuru Rudra Prathap , Gyanesh Gupta , Jaivanth Melanaturu , " Categorizing Mental Stress: A Consistency-Focused Benchmarking of ML and DL Models for Multi-Label, Multi-Class Classification via Taxonomy-Driven NLP Techniques", Natural Language Processing Journal, pp. 1-14.2025. https://doi.org/10.1016/j.nlp.2025.100162