International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com)
doi : https://doi.org/10.32628/CSEIT2390351

Analysing Cyber Threats: A Comprehensive Literature Review on Data-Driven Approaches

Ms. Ayushi Monani*, Mr. Omkar Bhusnale, Mr. Kunal Borade, Mrs. Rucha Madali
Department of Computer Engineering, Dr D. Y. Patil Institute of Technology, Pimpri, Pune, Maharashtra, India

ARTICLEINFO

Article History:

Accepted: 01 May 2023
Published: 20 May 2023

Publication Issue
Volume 9, Issue 3
May-June-2023

Page Number
188-193

ABSTRACT

This literature review paper aims to examine and analyze the existing research on prediction models for different types of cyber-attacks. Four key research papers have been selected as the base for this review: "A Prediction Model of DoS Attack's Distribution Discrete Probability," "Apriori Viterbi Model for Prior Detection of Socio-Technical Attacks," "Cyber Attacks Prediction Model Based on Bayesian Network," and "Applying Data Science to Cybersecurity Network Attacks & Events." An overview of the value of prediction models in cybersecurity and their function in reducing potential threats come first in the review. The methodology section outlines the search strategy used to identify relevant literature and the selection criteria for the base papers. The subsequent sections provide an overview of the field, highlighting the historical development and key theories or frameworks related to cyber-attack prediction. The themes or subtopics identified in the literature are discussed, focusing on the discrete probability distribution model of DoS attacks, the Apriori Viterbi model for detecting socio-technical attacks, the Bayesian network-based prediction model, and the application of data science in analyzing network attacks and events. The review critically evaluates the selected papers, analyzing their methodologies, findings, and limitations. It identifies gaps, controversies, and conflicting findings in the literature, paving the way for further research in the field. The synthesis and interpretation section integrates the findings from the different studies, compares various perspectives, and discusses the implications and significance of the literature for cyber-attack prediction. In conclusion, this literature review paper provides a comprehensive analysis of prediction models for cyber-attacks, based on the selected base papers. It highlights the strengths and weaknesses of existing approaches, identifies research gaps, and offers recommendations for future studies. This review contributes to the advancement of knowledge in the field of cybersecurity

and aids in the development of more effective prediction models to combat evolving cyber threats.

Keywords: Prediction model, DoS attack, Discrete probability distribution, Apriori Viterbi model, Socio-technical attacks, Cyber-attack prediction, Bayesian network, Data Science, Cybersecurity, Network attacks, Events, Literature review.

I. INTRODUCTION

In the realm of cybersecurity, the ability to predict and detect cyber-attacks has become increasingly crucial. As the frequency and sophistication of malicious activities continue to rise, organizations and individuals alike face significant challenges in safeguarding their digital assets and networks. To address these challenges, researchers have developed various prediction models aimed at anticipating and mitigating potential cyber threats. This literature review paper explores and analyses existing research on prediction models for cyber-attacks, focusing on four base papers: "A Prediction Model of DoS Attack's Distribution Discrete Probability," "Apriori Viterbi Model for Prior Detection of Socio-Technical Attacks," "Cyber Attacks Prediction Model Based on Bayesian Network," and "Applying Data Science to Cybersecurity Network Attacks & Events."

The objective of this review is to gain a comprehensive understanding of the current state of the field and to identify gaps, trends, and opportunities for further research. By examining and critically evaluating the selected base papers, we aim to contribute to the advancement of knowledge in the area of cyber-attack prediction.

The proliferation of distributed denial-of-service (DoS) attacks poses a significant threat to network infrastructure and system availability. The base paper titled "A Prediction Model of DoS Attack's Distribution Discrete Probability" investigates a prediction model that utilizes discrete probability distributions to forecast the distribution of DoS attacks. This paper offers insights into the modelling techniques employed and evaluates the effectiveness of such an approach in predicting and mitigating DoS attacks.

Socio-technical attacks, which exploit vulnerabilities stemming from the interaction between human behaviour and technical systems, represent another area of concern. The base paper titled "Apriori Viterbi Model for Prior Detection of Socio-Technical Attacks" proposes a novel prediction model that combines the Apriori algorithm and the Viterbi algorithm to detect and prevent socio-technical attacks. By studying this paper, we can gain valuable insights into the prior detection of attacks that leverage human factors, allowing for proactive measures to mitigate their impact.

Furthermore, the base paper titled "Cyber Attacks Prediction Model Based on Bayesian Network" explores the use of Bayesian networks as a predictive tool for cyber-attacks. By leveraging probabilistic inference, this model offers a systematic approach to predicting and identifying potential attack scenarios. Understanding the strengths and limitations of Bayesian network-based models can inform future advancements in cyber-attack prediction.

Lastly, the base paper titled "Applying Data Science to Cybersecurity Network Attacks & Events" highlights the application of data science techniques in analysing network attacks and events. By leveraging data mining, machine learning, and statistical analysis, this paper demonstrates the potential of data-driven approaches

to improve the accuracy and efficiency of cyber-attack prediction.

Through this literature review, we aim to synthesize the key findings, methodologies, and limitations of the selected base papers. By critically evaluating and integrating the insights from these studies, we can identify gaps in the existing literature and propose recommendations for future research. Ultimately, this review seeks to contribute to the development of more effective prediction models, enhancing our ability to anticipate and mitigate cyber threats in an increasingly interconnected world.

II. METHODS AND MATERIAL

To conduct this literature review, a systematic search strategy was employed to identify relevant research papers related to prediction models for cyber-attacks. The following steps were followed in the methodology:

A. Identification of Databases and Search Terms:

Key databases in the field of cybersecurity, such as IEEE Xplore, ACM Digital Library, and Scopus, were selected. Search terms and keywords were identified to capture the relevant literature, including phrases like "prediction model," "cyber-attacks," "security," "network attacks," and "data science."

B. Search Execution:

The selected databases were searched using the identified search terms. The search was limited to academic papers, conference proceedings, and journal articles written in English.

C. Inclusion and Exclusion Criteria:

Papers were included if they focused on prediction models for cyber-attacks. Papers addressing different types of attacks, such as DoS attacks, socio-technical attacks, or network attacks, were considered relevant. Only papers published within a specific timeframe were included to ensure the currency of the literature.

D. Screening and Selection:

The search results were screened based on the titles and abstracts to identify potentially relevant papers. Full-text versions of the selected papers were obtained and thoroughly reviewed to determine their suitability for inclusion.

E. Base Paper Selection:

From the pool of relevant papers, four base papers were chosen based on their significance, relevance, and contribution to the field. The selected base papers were "A Prediction Model of DoS Attack's Distribution Discrete Probability," "Apriori Viterbi Model for Prior Detection of Socio-Technical Attacks," "Cyber Attacks Prediction Model Based on Bayesian Network," and "Applying Data Science to Cybersecurity Network Attacks & Events."

F. Data Extraction and Analysis:

The selected base papers were analyzed in detail to understand their methodologies, findings, and limitations. Key information such as the prediction models used, data sources, evaluation metrics, and key results were extracted for each paper.

G. Synthesis and Interpretation:

The extracted data and findings from the selected base papers were synthesized and interpreted. Common themes, trends, and gaps in the literature were identified and discussed.

H. Writing the Literature Review:

The literature review paper was structured based on the identified themes and subtopics. The findings from the selected base papers were integrated with the broader literature to provide a comprehensive analysis.

By following this systematic methodology, a comprehensive overview of the literature on prediction models for cyber-attacks was obtained. The methodology ensured the inclusion of relevant and

high-quality research papers, allowing for an informed and rigorous analysis in this literature review.

III.RESULTS AND DISCUSSION

The results and discussion section of this literature review paper provides an analysis and synthesis of the selected base papers on prediction models for cyber-attacks. It highlights the key findings, methodologies, and limitations of each study while also identifying common themes, trends, and gaps in the literature.

A. "A Prediction Model of DoS Attack's Distribution Discrete Probability":

The base paper proposes a prediction model that utilizes discrete probability distributions to forecast the distribution of DoS attacks. The study demonstrates the effectiveness of this approach in predicting and mitigating DoS attacks. The methodology involves analysing historical attack data to estimate the probability distribution parameters, which are then used to predict future attack patterns.

Limitations of the model include the assumption of stationarity in attack patterns and the need for continuous updating of the probability distribution parameters.

B. "Apriori Viterbi Model for Prior Detection of Socio-Technical Attacks":

This base paper presents a novel prediction model that combines the Apriori algorithm and the Viterbi algorithm to detect and prevent socio-technical attacks. The model focuses on identifying attack patterns that leverage human factors and social engineering techniques. By analysing historical attack data and user behaviour patterns, the model generates prior probabilities of different attack scenarios and uses the Viterbi algorithm for real-time attack detection. The study demonstrates promising results in detecting socio-technical attacks and highlights the importance of considering human behaviour in prediction models. However, challenges remain in accurately modelling complex human behaviour and adapting the model to evolving attack strategies.

C. "Cyber Attacks Prediction Model Based on Bayesian Network":

The base paper proposes a prediction model based on Bayesian networks for cyber-attacks. The model leverages probabilistic inference to predict and identify potential attack scenarios. By considering the relationships between attack variables, the model provides a systematic approach to attack prediction. The study demonstrates the potential of Bayesian network-based models in enhancing prediction accuracy and assisting in decision-making processes. However, challenges exist in modelling the complex dependencies and uncertainties in cyber-attack scenarios, as well as in obtaining accurate prior probabilities for the Bayesian network.

D. "Applying Data Science to Cybersecurity Network Attacks & Events":

This base paper explores the application of data science techniques in analyzing network attacks and events. The study emphasizes the importance of utilizing data mining, machine learning, and statistical analysis to extract insights from large-scale security data. By applying data science methods, the paper highlights the potential for improving the accuracy and efficiency of cyber-attack prediction.

The results demonstrate the usefulness of data-driven approaches in identifying patterns, detecting anomalies, and enhancing situational awareness.

However, challenges remain in handling the vast amount of security data, ensuring data quality and privacy, and developing robust models that can adapt to evolving attack techniques.

Overall, the selected base papers provide valuable insights into different approaches and methodologies for predicting cyber-attacks. They highlight the importance of considering diverse factors such as attack patterns, human behaviour, probabilistic relationships, and data-driven analysis. While each

model presents promising results, limitations exist in terms of assumptions, scalability, and adaptability to evolving threats.

The synthesis of the selected base papers reveals several common themes and trends. These include the integration of different data sources, the importance of feature selection and dimensionality reduction, the need for continuous model updating, and the potential for combining multiple prediction techniques for enhanced accuracy.

However, gaps and challenges in the literature also become apparent. These include the need for more realistic modelling of complex attack scenarios, incorporating real-time data and dynamic updates, considering contextual information, and developing robust models that can handle uncertainties and adversarial behaviours. This analysis contributes to the understanding of prediction models

IV.CONCLUSION

In conclusion, based on the chosen base papers, this literature review offers a thorough analysis of prediction models for cyber-attacks. The results emphasize the significance of creating reliable prediction models to foresee and reduce cyber threats in a world that is becoming more interconnected.

The base papers under review present various methodologies and approaches for anticipating various cyber-attacks, such as DoS attacks, socio-technical attacks, and network attacks. To improve prediction accuracy and support decision-making, these models make use of techniques like discrete probability distributions, Apriori and Viterbi algorithms, Bayesian networks, and data science methods.

While the results in the chosen papers are encouraging, there are still a number of issues that need to be resolved. Common issues include assumptions about stationarity, modelling human behaviour, uncertainty in attack scenarios, and the scalability of models. Addressing these issues and further developing the field of cyber-attack prediction should be the main goals of future research.

Overall, by offering insights into the current prediction models for cyberattacks, this literature review adds to the body of knowledge in cybersecurity. It highlights the significance of integrating various factors, incorporating real-time data, and developing robust models that can adapt to evolving threats. It also identifies gaps, trends, and opportunities for future research.

This review provides a thorough understanding of the state of prediction models for cyber-attacks by analyzing and synthesizing the results from the chosen base papers. It helps in the creation of more accurate prediction models and improves our capacity to foresee and mitigate cyber threats in the digital environment, making it a valuable resource for researchers, practitioners, and policymakers in the field of cybersecurity.

V. REFERENCES

[1]. JunZhaoabXudongLiuabQibenYancBoLiabMinglaiShaoabHaoPe-ngabLichaoSund https://www.sciencedirect.com/science/article/pii/S0167404820304259?ref=pdf_download&fr=RR-2&rr=74b2e83f0c4f85d

[2]. https://brainyx.co/journal/journal9/

[3]. https://scholar.google.com/scholar_lookup?title=The%20impact%20of%20artificial%20intelligence%20on%20autonomous%20cyber%20defense&author=Crawford&publication_year=2017

[4]. https://developer.lookingglasscyber.com/reference/getdocs

[5]. https://warden.cesnet.cz/en/index

[6]. https://lookingglasscyber.com/solutions/scoutthreat/

[7]. https://f.hubspotusercontent30.net/hubfs/7412038/Data%20Sheets/Data%20Sheet%20_scoutTHREAT%202021%20_vFINAL.pdf

[8]. https://lookingglasscyber.com/resources/case-studies/

[9]. https://ieeexplore.ieee.org/document/9725445/metrics#metrics
[10].https://pure.port.ac.uk/ws/portalfiles/portal/13360330/ICISSP_2019_24_CR_1_.pdf
[11].https://pyattck.readthedocs.io/_/downloads/en/latest/pdf/
[12].http://ceur-ws.org/Vol-3056/paper-05.pdf
[13].https://bi-survey.com/big-data-security-analytics
[14].https://juice-shop.herokuapp.com/#/
[15].https://github.com/greenbone/openvas-scanner
[16].https://greenbone.github.io/docs/latest/background.html
[17].https://github.com/naserdamer/SMDD-Synthetic-Face-Morphing-Attack-Detection-Development-dataset

Cite This Article :

Ayushi Monani, Omkar Bhusnale, Kunal Borade, Rucha Madali, "Analysing Cyber Threats: A Comprehensive Literature Review on Data-Driven Approaches", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 9, Issue 3, pp.188-193, May-June-2023. Available at doi : https://doi.org/10.32628/CSEIT2390351
Journal URL : https://ijsrcseit.com/CSEIT2390351