e-ISSN:2582-7219

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

Volume 7, Issue 7, July 2024

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707064 |

Cyber Threat Intelligence Mining for Proactive Cybersecurity Defense a Survey Perspectives

Prof. Rajesh N, Karthik P Naidu

Assistant Professor, Department of MCA, AMC Engineering College, Bangalore, India

Student, Department of MCA, AMC Engineering College, Bangalore, India

ABSTRACT: Cybersecurity is increasingly vital in the face of sophisticated and persistent cyber assaults. Traditional reactive approaches are inadequate for modern cyber defense. This initiative examines the development of a proactive cybersecurity defense system leveraging Cyber Threat Intelligence (CTI) mining. The system incorporates advanced data mining techniques to analyze threat data, predict potential attacks, and enhance security posture. We review extant CTI frameworks, propose a novel CTI mining system, and detail its modular architecture. Experimental results demonstrate the system's efficacy in anticipating and mitigating cyber threats, emphasizing its potential to enhance cybersecurity resilience.

I. INTRODUCTION

The proliferation of cyber threats, including ransomware, phishing, and advanced persistent threats (APTs), necessitates more robust cybersecurity strategies. Traditional reactive measures, which respond after an attack occurs, are increasingly ineffective against sophisticated adversaries. Cyber Threat Intelligence (CTI) mining presents a proactive approach, enabling organizations to anticipate and counteract cyber threats. By analyzing enormous quantities of threat data, CTI mining provides actionable insights, allowing organizations to preemptively strengthen their defenses. This project seeks to develop a CTI mining framework that leverages data mining and machine learning techniques to identify emergent threats and inform proactive defense strategies.

II. LITERATURE REVIEW

A. Cyber Threat Intelligence (CTI)

CTI involves the accumulation, analysis, and dissemination of information regarding prospective cyber hazards. Taxonomies by Barnum (2014) provide a structured approach to CTI, categorizing threat data into indicators, tactics, techniques, and procedures (TTPs) . Research by Chismon and Ruks (2015) outlines the operational, tactical, and strategic levels of CTI, emphasizing its role in informing security operations . Li et al. (2019) review various CTI sharing platforms and standards, such as STIX and TAXII, emphasizing their importance in facilitating information exchange .

B. Data Mining for Cybersecurity

Data mining techniques, such as clustering, classification, and anomaly detection, are critical for analyzing cybersecurity data. A survey by Buczak and Guven (2016) examines the application of data mining in intrusion detection systems (IDS), emphasizing its efficacy in identifying patterns indicative of attacks . Abawajy et al. (2014) review machine learning algorithms for detecting malware and network intrusions, noting the challenges of scalability and data diversity . Kim et al. (2020) discuss the integration of deep learning in cybersecurity, particularly for analyzing large-scale threat data .

C. Proactive Cybersecurity Defense

Proactive defense strategies involve predicting and mitigating hazards before they materialize. Hutchins et al. (2011) introduce the Cyber Kill Chain model, which details the phases of a cyber attack and emphasizes opportunities for proactive defense . Sommer and Paxson (2010) discuss the limitations of signature-based detection and advocate for anomaly-based approaches to identify novel threats . Shackleford (2016) analyzes the use of threat hunting techniques, which involve actively scanning for indicators of compromise (IOCs) within networks .

D. Existing CTI Frameworks

Several frameworks have been devised to facilitate CTI operations. MITRE's ATT&CK framework, described by Strom et al. (2018), provides a comprehensive matrix of adversarial techniques, facilitating in threat detection and response . The ThreatStream platform, reviewed by Gao et al. (2018), incorporates threat intelligence inputs to provide real-time insights into threat landscapes . Fransen et al. (2018) discuss the Collective Intelligence Framework (CIF), which facilitates the aggregation and analysis of threat data from various sources .

III. EXISTING SYSTEM

A. Traditional Security Measures
Traditional cybersecurity measures primarily concentrate on reactive responses, including firewalls, antivirus software, and IDS. Scarfone and Mell (2007) outline the architecture and deployment of IDS, emphasizing their role in detecting known attack signatures . While effective for known threats, these systems struggle to detect new or evolving attack patterns, leading to breaches in security .

B. Conventional CTI Systems
Conventional CTI systems aggregate threat data from various sources and provide intelligence to security teams. Ahmed et al. (2017) review the limitations of these systems, including challenges in data integration and real-time analysis . Alperovitch (2017) critiques the reliance on manual analysis in traditional CTI systems, which limits their ability to scale and respond to emergent threats effectively .

IV. PROPOSED SYSTEM

A. Overview
The proposed system leverages CTI mining to enhance proactive cybersecurity defense. It incorporates data mining and machine learning techniques to analyze threat data, predict potential attacks, and provide actionable insights. The system comprises of several modules: data acquisition, preprocessing, threat analysis, and visualization.

B. System Architecture
Data Collection: Aggregates threat data from diverse sources, including CTI feeds, network records, and social media. Utilizes APIs and web scraping tools to gather real-time data .

Preprocessing: Cleans and normalizes the data, eradicating noise and ensuring consistency. Implements techniques such as tokenization and stemming for text data .

Threat Analysis: Applies clustering algorithms to identify patterns in the data. Uses supervised learning models to classify threats and predict potential attacks .

Visualization: Provides a user-friendly interface for visualizing threat intelligence. Implements dashboards that display threat trends, risk scores, and recommended actions .

V. MODULE DESCRIPTION

A. Data Collection Module
Collects threat data from multiple sources, including structured CTI inputs, unstructured web data, and network records. Ensures the data is up-to-date and exhaustive, encompassing numerous threat vectors and indicators .

B. Preprocessing Module
Processes unprocessed data to eradicate inconsistencies and prepare it for analysis. This includes data cleansing, normalization, and feature extraction. For text data, implements natural language processing (NLP) techniques to extract pertinent entities and relationships .

C. Threat Analysis Module
Utilizes clustering techniques, such as k-means and DBSCAN, to group similar hazard indicators. Applies classification algorithms, including decision trees and neural networks, to categorize hazards and predict their impact. Incorporates anomaly detection to identify peculiar patterns that may indicate emerging threats .

D. Visualization Module

Presents the analyzed data through interactive interfaces. Visualizations include thermal maps of threat activity, timelines of attack patterns, and network graphs of threat relationships. Enables security analysts to rapidly interpret the data and take informed actions .

VI. RESULTS

The proposed system was validated using real-world threat data. Evaluation metrics included detection accuracy, false positive rate, and response time. The system demonstrated a high accuracy in identifying potential hazards and provided actionable insights for proactive defense. The visualization module effectively communicated complex threat patterns, enhancing situational awareness for security analysts.

VII. CONCLUSION

The proposed CTI mining system represents a significant advancement in proactive cybersecurity defense. By leveraging data mining and machine learning, it provides timely and actionable insights into potential threats. The system's modular architecture assures scalability and flexibility, making it suitable for diverse organizational contexts. Future work will focus on enhancing the system's real-time capabilities and integrating additional data sources to enhance its threat prediction accuracy.

REFERENCES

1. Barnum, S. (2014). "Standardizing cyber threat intelligence information with the Structured Threat Information Expression (STIX)."
2. Chismon, D., & Ruks, M. (2015). "Threat intelligence: Collecting, analyzing, and evaluating."
3. Li, J., et al. (2019). "Cyber threat intelligence sharing: A survey."
4. Buczak, A. L., & Guven, E. (2016). "A survey of data mining and machine learning methods for cyber security intrusion detection."
5. Abawajy, J., et al. (2014). "Big data security: Challenges and strategies."
6. Kim, Y., et al. (2020). "Deep learning-based threat detection."
7. Hutchins, E. M., et al. (2011). "Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains."
8. Sommer, R., & Paxson, V. (2010). "Outside the closed world: On using machine learning for network intrusion detection."
9. Shackleford, D. (2016). "Threat hunting: Open season on the adversary."
10. Strom, B. E., et al. (2018). "MITRE ATT&CK: Design and philosophy."
11. Gao, Y., et al. (2018). "Threat intelligence sharing platform."
12. Fransen, F., et al. (2018). "Improving incident response with threat intelligence."
13. Scarfone, K., & Mell, P. (2007). "Guide to intrusion detection and prevention systems (IDPS)."
14. Stallings, W., & Brown, L. (2018). "Computer security: Principles and practice."
15. Ahmed, M., et al. (2017). "Survey on network security intrusion detection system using machine learning."
16. Alperovitch, D. (2017). "The limitations of cyber threat intelligence."
17. Undercoffer, J., et al. (2003). "A model for threat and risk assessment in cyber security."
18. Aggarwal, C. C. (2015). "Data mining: The textbook."
19. Chen, H., et al. (2017). "Big data mining for the internet of things."
20. Thomas, D. R. (2019). "Deep learning for cybersecurity."
21. Conti, M., et al. (2018). "Blockchain-based data provenance for the internet of things."
22. Das, S., & Kant, K. (2019). "Threat intelligence sharing in cybersecurity."
23. Lee, R. M. (2018). "Threat intelligence in practice."
24. Zuech, R., et al. (2015). "Big data analytics for detecting cyber attacks."
25. Sadique, F. (2020). "Cyber threat intelligence: Enhancing cybersecurity."

INNO SPACE
SJIF Scientific Journal Impact Factor

INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

निस्केयर
NISCAIR

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com