ASIAN BULLETIN OF BIG DATA MANAGMENT
Vol. 3. Issue 2 (2023)
DOI: https://doi.org/ 10.62019/abbdm.v3i2.85

Machine Learning Applications in Cyber Threat Intelligence: A Comprehensive Review

Iqra Naseer*

Chronicle

Article history
Received: January 6th, 2024
Received in the **revised format:** January 9th ,2024
Accepted: January 10th 2024
Available online: January 3, 2024

Iqra Naseer is currently affiliated with National University of Sciences and Technology Nust, Pakistan.
Email: iqranaseer74@gmail.com

***Corresponding Author**

Abstract

Contemporary institutions are consistently confronted with fraudulent activities that exploit weaknesses in interconnected systems. Securing critical data against unauthorized access by hackers and other cybercriminals requires the application of robust cybersecurity protocols. As the number and complexity of cyber threats continue to grow, innovative prevention strategies are required. The objective of this study is to investigate the correlation between machine learning (ML) and cyber threat intelligence (CTI) to improve cybersecurity strategies. For the detection of anomalies, the analysis of malware, and the prediction of threats, ML techniques are indispensable in industries including retail, finance, healthcare, and cybersecurity. By employing critical threat information (CTI), security teams can gain a comprehensive understanding of adversary strategies and bolster defensive measures; thus, they play a pivotal role in proactive defense. Integration of ML and CTI facilitates exhaustive analysis by automating the acquisition, processing, and categorization of data. However, obstacles arise when confronted with issues such as risk assessment, the requirement for precise data, and the initial stages of machine learning implementation in business intelligence. In this paper, we present an extensive examination of the current literature concerning the visualization of Cyber Threat Intelligence (CTI) and the utilization of Machine Learning (ML. Therefore, the report concludes with an analysis of emergent threats, potential future applications of AI and ML in the field of cyber threat intelligence, and the critical contribution of machine learning to the improvement of cybersecurity.

Keywords Machine Learning, Cyber Threat Intelligence, Cyber Security, Defensive Measures.

INTRODUCTION

In the modern age of technology, every firm is related to several technological systems, operational methods, and procedural frameworks. This enables cyber assailants to move unrestrictedly inside the operational setting. Every company is vulnerable to facing a continuous stream of assailants and intruders. Their objective is to target both major firms and small enterprises throughout both the public and private domains. Therefore, the process of identifying cyber attackers and evaluating the severity of threats requires the use of cybersecurity defensive technologies, protocols, and algorithms. Cybersecurity refers to a range of techniques and tactics used to protect the reliability of devices, networks, and information against damage, attacks, and unwanted entry. Large corporations own large quantities of data, including vital and confidential information.

Hence, it is crucial to protect sensitive data against illegal access and possible threats. Cybersecurity defenses are put in place to counteract three categories of attacks: conventional attacks that make use of well-known weaknesses, sophisticated attacks that exploit complex flaws, and emerging attacks that exploit newly discovered weaknesses. The information system has infiltrated several aspects of the company, such as the production, operation, and management departments (Bastos, & Martini, 2015). These advancements entail the need for reliable information security policies and solutions. Cybersecurity defenses include educating people, quickly detecting intrusions, and analyzing new and previously unseen threat scenarios. Cybersecurity requires the reduction of both instances of false positive and false negative vulnerabilities (Rodriguez & Okamura, 2020). Notwithstanding the use of many preventative measures, such as antivirus software, encryption, and firewalls, enterprises persistently encounter a substantial amount of intrusions. Companies in the commercial banking, credit card, and communications sectors are particularly susceptible and need effective threat detection measures (Gupta & Chowdhary, 2017). The increase in cybersecurity threats in recent years has emphasized the need for automated threat analysis at every level of an organization or company (Sisiaridis & Markowitch, 2018). Cyber threat intelligence (CTI) has arisen as a viable remedy for firms to tackle the increasing quantity and intricacy of security breaches. Cyber Threat Intelligence (CTI) refers to the systematic process of identifying and assessing potential cyber risks across an entire organization in a proactive way. While subscribing to several threat intelligence sources might be detrimental, it may result in an abundance of data. A Threat Intelligence-Sharing Platform (TISP) facilitates the integration of intelligence into various technologies to assist in incident response by converting cyber threat information data into actionable intelligence. Currently, information security firms and the ecosystem provide TISP (Threat Intelligence and Sharing Platform) solutions that can be classified into two main categories: content aggregation, which involves gathering various updates on threat data, and threat intelligence management, which focuses on extracting economic value from the acquired data.

FOUNDATIONS OF MACHINE LEARNING IN CYBER THREAT INTELLIGENCE

Overview of Machine Learning

By using machine learning techniques, large datasets may be sifted for useful business insights. While these techniques are used by many businesses, they are especially common in the retail, banking, healthcare, and cybersecurity fields. In addition, new cyber dangers may be countered using machine learning. There is a wide variety of cyberattacks, including man-in-the-middle attacks, distributed denial of service attacks, SQL injection, phishing, malware, social engineering, and cross-site scripting assaults. Organizations utilize machine learning to continuously assess data, identify patterns that may indicate impending attacks, and implement countermeasures. Machine learning is frequently employed to detect previously unknown malware, identify anomalies by monitoring network behavior, and prevent access to malicious websites, among other tasks. Data in cloud settings may also be protected using these methods. According to Berghout et al. (2022), machine learning is mostly used in the security industry for network analysis, intrusion identification, and malware categorization. Data scarcity or poor quality data used for teaching the techniques is a major problem for anybody doing

security-related work. Due to the potentially severe repercussions of an error, this industry has stringent accuracy standards (Berghout et al., 2022).

TYPES OF MACHINE LEARNING

In the realm of cyber threat detection, machine learning techniques are implemented to fortify defense mechanisms and bolster resilience against ever-evolving threats.

- A fundamental approach in machine learning, **supervised learning** entails the utilization of labeled data to instruct models on the differentiation between benign and detrimental samples. By employing this methodology, cybersecurity systems can predict the malevolent attributes of newly encountered samples. Unsupervised learning is the process of discovering patterns and correlations in unlabeled data through analysis without external guidance.
- **Unsupervised learning** is advantageous within the domain of cybersecurity due to its capacity to reveal novel attack patterns or adversary behaviors. This enhances the capacity to detect anomalies within extensive datasets.
- **Reinforcement learning**, which functions similarly to how humans learn through the process of iteration and correction, is an indispensable component of cybersecurity. This method is particularly advantageous when it comes to tackling intricate challenges like safeguarding cyber-physical systems, detecting autonomous intrusions, and defending against distributed denial of service (DDoS) attacks; it consistently optimizes the overall advantages. By enhancing their ability to forecast, identify, and mitigate risks, these machine learning paradigms empower cybersecurity professionals to proactively safeguard against a wide range of cyber-attacks.

Security Applications of Machine Learning:

- **Real-time monitoring of email.** Anomaly detection and natural language processing are applied to email content by machine learning to identify fraud attempts.
- **Contrast with bots.** Some bots may take over an application and run harmful code; others produce 25% of all page traffic. Though several ML methods exist for this purpose, the most important ones for detecting bots are response rate, message variability, and temporal pattern recognition.
- **Anti-Malware Software**. Polymorphic malware can circumvent security safeguards. Countering this kind of malware is made easier with the use of several machine-learning approaches. These include decision trees, support vector machines, convolutional neural networks, and others.
- **Multi-Factor Authentication.** Machine learning facilitates the acceleration of security tool development, including multi-factor authentication, and subsequently permits their global scalability by IT resources. In summary, machine learning is an immensely promising technological advancement that has brought about a paradigm shift in the field of cybersecurity. The capacity of this technology to analyze extensive volumes of data, detect patterns, and generate precise forecasts has uncovered novel opportunities in a multitude of sectors, including healthcare, finance, transportation, and many others.

Proactive Defense : Cyber Threat Intelligence

CTI enables security teams to proactively defend, against cyber-attacks or detect malicious activities on enterprise networks before they intrude into those networks. For example, CTI can provide valuable insights to the teams regarding the adversary's strategy and methods of operation. Consequently, this empowers them to enhance their defenses against particular assault techniques that are recognized as having been employed by the adversary. Additionally, it facilitates the generation of actionable information that decision-makers can comprehend (Al-Fawa'reh et al., 2022).

There are primarily three qualities that CTI should have:

(1) Evidence-based: malware analysis may confirm the authenticity of cyber threats; (2) Utility: businesses need to figure out how to use the collected CTI to improve security incidents and (3) The collected CTI needs to motivate not only data or information but also measures to regulate security.

Threat intelligence is derived by gathering and analyzing threat streams; "security analytics" then uses this information to improve detection chances. Cyber threat intelligence (CTI) is compiled by continuously analyzing massive amounts of threat data to classify and put cyber threat actions, trends, and assaults into perspective. Research, inputs from external threats, internal networks, and study of previous attacks may all lead to its acquisition. For instance, it may be created by combining data from trustworthy, diverse, and fused sources. The open web, black web, deep web, web crawlers, security networks, spam traps, botnet monitoring services, social media, and historical data gathered regarding harmful things are all potential places to get such information. The next step is to apply several pre-processing algorithms to the aggregated data before processing it comprehensively, sometimes in real time. Statistical criteria, expert systems (including sandboxes, similarity tools, heuristic engines, and behavior profiling), validation of security analysts, and verification of whitelisting are all part of these strategies. (Carroll etal., 2021).

Machine Learning and CTI Modeling: Literature Survey

In this part, the literature on cyber threat intelligence (CTI) and how it relates to the visualization of cyber threat data is discussed. An analysis of the machine-learning techniques used to extract meaningful patterns from the data is also included in the package. (Strom et al., 2018) states that businesses are using threat data at an exponential rate, which highlights the need for CTI. Malware, the attacker, and countermeasures may all be better understood with the help of CTI. To prevent cyberattacks before they happen, some companies are investing in CTI (Cyber Threat Intelligence). However, there are major issues with CTI volume, interesting pattern discovery, and automated pattern use in security measures like honeypots, intrusion detection systems, and firewalls that the research community has to solve. CTI includes information regarding both internal and external dangers. Gaining a comprehensive understanding of external dangers can significantly enhance an organization's ability to safeguard itself against potential harm.

The current technologies for viewing CTI data have restricted functionalities. Security analysts cannot effectively visualize extensive CTI data from many angles and lack the

flexibility required to implement robust defensive measures. A cyber-security knowledge network that includes many degrees of abstraction in the subject is introduced by Bromiley (2016) as STUCCO. Security analysts and malware analysis tools will benefit from the standardized data collected from thirteen distinct sources, which will include both organized and unstructured information. Not only is the ontology compatible with the Graph JSON format but it has also been instantiated in JSON. This program's inability to include CTI visualization features is a major downside. In addition, the malware item in the ontology just provides details on the attack type; it does not reveal how it was carried out. In addition, the suggested ontology does not include any information regarding the behavior and tactics. Miles et al. (2014) have proposed a web service for the analysis of malware. This web service offers the ability to analyze malware by examining its code and semantics. This relationship visualization aids the security analyst in the categorization of malware.

For CTI visualization, there is a product named STIX that is available for purchase (Ejaz et al., 2022). To display STIX documents, this tool makes use of Javascript. For STIX documents, STIXViz offers three different kinds of visualizations: first, a timeline view; second, a graph view; and third, a tree view. Items in an STIX document, including incidents and the campaigns they were a part of, are shown in the timeline view according to the time and date they were recorded. To dynamically place the nodes, the graph view utilizes force-directed graph architecture. Because nodes may be moved, the graph's structure can rearrange itself without any human intervention. In the tree view, STIX entities are shown at their original or highest level. Among these things are things like events, exploit targets, campaigns, observables, indicators, and TTP. When the tree is first expanded, it displays a top-level node for each element category in the STIX file. The node with the black and white border may be expanded to display information at the second level by clicking on it.

Common Attack Pattern Enumeration and Classification (CAPEC) CAPEC is a database of known assault patterns. In his 2015 paper, MITRE Corporation's Steven Noel specifies many interactive visualization methods for this database (Noel, 2015). e. These graphics demonstrate how the CAPEC attack pattern taxonomy hierarchically links assaults. You may show the CAPEC taxonomy using three interactive visuals. Sunburst is the first example (Kaushal et al., 2017). It depicts all tree levels as diverging lines from a central point. Wetzel (2004) suggests the circular tree map. It has spacious exhibits for bigger assault patterns at higher tree levels. Thirdly, the Voronoi treemap (Hussain et al., 2018) iteratively partitions the screen according to the unity aspect ratio for a balanced interactive display. Interactive and maximizing screen space, these visualizations display huge hierarchical data. Treemaps were used to illustrate the ATT&CK knowledge base and malware data because they are practical.

Bronwyn, Perl, and Lindauer (2015) argue that machine-learning techniques can be employed to organize unstructured data. Their findings were unveiled through an analysis of real-life incident data. An important question was posed: "How can a community derive utility from it?" given the availability of a large dataset of shared event data. They came up with a plan to combine machine learning methods with low-quality unstructured event data to make the data more comprehensible and encourage community participation. They said that using machine learning methods improves the completeness of incident data for analysis and decision-making by revealing more about

shared data indications and occurrences. Nevertheless, they made use of clustering and other unsupervised machine-learning approaches. Malware detection and family identification using classification methods outperform clustering methods in static, dynamic, and hybrid investigations (Sisaat et al., 2017).

The evolution of malware over recent years is examined in a study on malware trends (Saeed, Selamat, & Abuagoub, 2013). The focus is on identifying the prevailing trends in the security business, enhancing organizational systems and networks to mitigate assaults, and anticipating future targets and advancements in the upcoming years. The report outlines the modifications in malware and the shifts in attacker techniques aimed at evading system defenses. This text provides a concise overview of prevalent malware, such as ransomware, that specifically targets cryptocurrencies like Bitcoin. These cryptocurrencies enable the transfer of funds without leaving a trace. Ransomware aims to infiltrate a user's system, seize their personal information, and encrypt all data. It thereafter issues a demand, typically requiring the transfer of funds in cryptocurrency within a defined timeframe. Failure to comply results in permanent irretrievable loss of the data. Efficiency and accuracy are measured for machine learning models. According to their assessment findings, KNN is more accurate than other approaches. Even with bias or inadequate training data, it is easy to deploy and produces better and reasonable results (Mohaisen, & Alrawi, 2013).

Role of Machine Learning in Cyber Threat Intelligence

Machine learning is an emerging area of study that has the potential to greatly enhance CS metrics.CS applications utilizing ML, for instance, can detect anomalies on a network more efficiently than those employing conventional techniques. As the rate of advancement quickens and there is a need for more efficient countermeasures, ML emerges as an obvious solution to the challenge of managing the escalating quantity of cyber-attacks. Computer scientists and network engineers have joined forces to create, model, and test new network penetration patterns and the characteristics that go along with them. Machine learning, genetic algorithms, intelligent agent multi-agent systems, neural networks, and machine learning are the centers of attention for several distinct methods. Fuzzy logic, pattern recognition algorithms, expert systems, inductive logic programming, Bayes classification, and associative techniques are all part of this category (Preuveneers & Joosen, 2021). Some examples of ML applications that can be used in CS solutions include spam filter applications, fraud detection, botnet detection, secure user authentication, cyber security ratings, and hacking incident forecasting (Kadoguchi et al., 2019). For instance, by teaching machine learning algorithms to look for certain traits, it is possible to find harmful software and good software.

Examples of such characteristics include the APIs that have been accessed, disc fields, environmental goods, processing power consumption, bandwidth use, and the amount of data transported over the internet. The system is built by using these distinct traits. After being installed on the system, a test program may examine these unique characteristics to identify malicious software. Within the framework of CTI, organizations may use ML approaches to automate data collecting and processing, combine unstructured data from disparate sources, and connect data from different places by giving context to how malicious actors operate and the state of the breach. Together with their current security systems, this can be accomplished. Incorporating data points from many sources, such

as the open web, deep web, black web, and technological sources, is necessary for this processing to provide the most thorough plan. Names, attributes, connections, and events are some more ways that ML algorithms may categorize data. Concepts are separated and then combined to accomplish this. To provide comprehensive category analysis, automate data categorization rather than rely on human sorting (Kure & Islam, 2019).

Furthermore, machine learning methods can be devised to classify text into distinct categories such as code, prose, or data records, and to eliminate ambiguity among entities sharing the same name by analyzing contextual cues in the adjacent text. Additionally, Machine Learning and statistical methodologies can be applied to classify events and entities based on their significance; for example, malevolent entities' risk scores can be evaluated. Risk scores can be computed using machine learning on a dataset that has been previously analyzed. Classifiers, such as risk scores, provide both an evaluation and contextual information regarding the score, given that multiple sources confirm that the IP address in question is malevolent. The process of automating risk classifications significantly reduces the time required to sift through false positives and prioritize tasks. In addition to its usage in constructing more accurate predictive analytical models than humans utilizing previously gathered and classified vast collections of data, ML may be used to forecast entity attributes and occurrences. Machine learning techniques might also serve as active sensors, collecting data and sending it to a centralized intelligence network that everyone could access. As stated previously, the process of implementing ML methodologies at various phases of CTI is in its infancy. For example, research endeavors in the field of operational intelligence are currently in the experimental and research phase, which requires significant resources (Montasari et al., 2021).

Threats and challenges

In actuality, cybercriminals use a variety of tactics to focus on a specific victim, with the goals of either (i) stealing sensitive information like financial records or (ii) taking over the victim's device and executing further malicious actions like botnet delivery of malware or ransomware encryption and locking process. While various cyberattacks may employ distinct methods of infection, their life cycles are essentially identical, commencing with victim reconnaissance and culminating in the execution of malicious activities on the targeted machine or network.

It is very important to know the exact places of entry and system weaknesses that hackers may use to protect against cyberattacks. Besides the usual tricks that have been used to trick people in the past, new methods have been used in recent years (e.g., phishing (Wardman et., 2018) to conduct the actions that the assailants desire, more sophisticated and inventive methods have been developed by others. The methods employed vary in nature, including the delivery of malicious software (e.g., PDF files or Word documents) in an unexpected format to the victim's device (Elingiusti et al., 2018), the exploitation of zero-day vulnerabilities (Ding et., 2018), and the breaching of anonymous communications to communicate with threat actors (Haughey et al., 2018). Some examples of these complex attacks are the new types of ransomware, which act like worms and have attacked hundreds of thousands of people, businesses, and important

systems. Because attack methods have gotten better, it is now very hard to figure out who did the attack and where it happened.

Cybercriminals using advanced escape and anti-forensics methods in their harmful code are another major worry when it comes to new cyberattacks. This makes common ways of checking for security flaws less useful, like the Common Vulnerability Scoring System (CVSS), static malware, and traffic analysis (Shalaginov et al., 2018). Also, companies are using more and more new networking technologies, like software-defined networking (SDN), the Internet of Things (IoT), and cloud computing. For example, they use cloud resources to store and process large amounts of data. This means that modern techniques need to be used to look into the data that was sent and stored (Conti, Dargahi & Dehghantanha, 2018).

To address the issues mentioned above, the field of cyber threat intelligence is expanding to include techniques of artificial intelligence and machine learning that can detect, understand, and respond intelligently to complex assaults. Security experts have needed ways to identify signs of cyber threats, thus researchers have been considering various artificial intelligence approaches in recent years. Machine learning (ML) and data mining methods in particular are seeing increased use as a result of their shown efficacy in malware analysis (static and dynamic) and network anomaly detection (Shalaginov et al., 2018). There are a variety of strategies that cyber-defenders might use, such as honeypots, to trick attackers into thinking they are not able to detect or prevent intrusions. These systems work by having security experts watch the actions of attackers and proactively prevent attacks, while also providing them with phony information or resources that look real (Papalitsas et al., 2018). To provide security practitioners and analysts with the most recent information, a mix of these strategies would be necessary.

CONCLUSION

Machine learning plays a vital role in the cybersecurity domain, particularly in the context of cyber threat intelligence (CTI). The exponential increase of CTI data is posing challenges for security analysts in effectively evaluating large amounts of data and promptly responding to emerging risks. Machine learning algorithms can automate threat detection by identifying patterns and irregularities that may suggest cyber threats. This enables rapid detection of these threats. This serves a dual purpose: it enhances the speed of detection while also allowing security specialists to visually study CTI data for valuable insights. Machine learning may assist security analysts in strengthening network defenses, mitigating vulnerabilities, and actively combating cyber attackers. It does this by standardizing the representation of threat information and facilitating comprehensive analysis from several perspectives. The cybersecurity sector may use open-source CTI repositories to enhance collaboration, facilitate information sharing, and fortify defenses against potential threats. Machine learning is a powerful tool that may enhance cybersecurity measures to adapt to the ever-changing threat landscape. It is a crucial element of CTI.

DECLARATIONS

Acknowledgement: We appreciate the generous support from all the supervisors and their different affiliations.

Funding: No funding body in the public, private, or nonprofit sectors provided a particular grant for this research.

Availability of data and material: In the approach, the data sources for the variables are stated.

Authors' contributions: Each author participated equally to the creation of this work.

Conflicts of Interests: The authors declare no conflict of interest.

Consent to Participate: Yes

Consent for publication and Ethical approval: Because this study does not include human or animal data, ethical approval is not required for publication. All authors have given their consent.

REFERENCES

Abu, M. S., Selamat, S. R., Ariffin, A., & Yusof, R. (2018). Cyber threat intelligence–issue and challenges. *Indonesian Journal of Electrical Engineering and Computer Science*, *10*(1), 371-379. DOI: 10.11591/ijeecs.v10.i1.pp371-379

Al-Fawa'reh, M., Al-Fayoumi, M., Nashwan, S., & Fraihat, S. (2022). Cyber threat intelligence using PCA-DNN model to detect abnormal network behavior. *Egyptian Informatics Journal*, *23*(2), 173-185. https://doi.org/10.1016/j.eij.2021.12.001

Bastos, M. R., & Martini, J. S. C. (2015, October). A model-free voltage stability security assessment method using artificial intelligence. In *2015 IEEE PES Innovative Smart Grid Technologies Latin America (ISGT LATAM)* (pp. 679-682). IEEE. doi: 10.1109/ISGT-LA.2015.7381238. https://doi.org/10.1109/ISGT-LA.2015.7381238

Berghout, T., Benbouzid, M., & Muyeen, S. M. (2022). Machine learning for cybersecurity in smart grids: A comprehensive review-based study on methods, solutions, and prospects. *International Journal of Critical Infrastructure Protection*, 100547. https://doi.org/10.1016/j.ijcip.2022.100547

Bromiley, M. (2016). Threat intelligence: What it is, and how to use it effectively. *SANS Institute InfoSec Reading Room*, *15*, 172. https://nsfocusglobal.com/wp-content/uploads/2017/01/SANS_Whitepaper_Threat_Intelligence_What_It_Is_and_How_to_Use_It_Effectively.pdf

Conti, M., Dargahi, T., & Dehghantanha, A. (2018). *Cyber threat intelligence: challenges and opportunities* (pp. 1-6). Springer International Publishing. https://doi.org/10.1007/978-3-319-73951-9_1

Ding, Q., Li, Z., Haeri, S., & Trajković, L. (2018). *Application of machine learning techniques to detecting anomalies in communication networks: Datasets and feature selection algorithms* (pp. 47-70). Springer International Publishing. https://doi.org/10.1007/978-3-319-73951-9_3

Ejaz, S., Noor, U., & Rashid, Z. (2022). Visualizing Interesting Patterns in Cyber Threat Intelligence Using Machine Learning Techniques. *Cybernetics and Information Technologies*, *22*(2), 96-113. DOI: https://doi.org/10.2478/cait-2022-0019

Elingiusti, M., Aniello, L., Querzoni, L., & Baldoni, R. (2018). Malware detection: A survey and taxonomy of current techniques. *Cyber threat intelligence*, 169-191. https://doi.org/10.1007/978-3-319-73951-9_9

Gupta, S., & Chowdhary, S. K. (2017, September). Authentication through electrocardiogram signals based on emotions-a step towards ATM security. In *2017 6th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO)* (pp. 440-442). IEEE. doi: 10.1109/ICRITO.2017.8342467.

Haughey, H., Epiphaniou, G., Al-Khateeb, H., & Dehghantanha, A. (2018). Adaptive traffic fingerprinting for darknet threat intelligence. *Cyber Threat Intelligence*, 193-217. https://doi.org/10.1007/978-3-319-73951-9_10

Hussain, H., Burgstaller, L., Grassberger, P., & Lvov, N. (2018). Data Visualisation on Mobile. https://courses.isds.tugraz.at/ivis/surveys/ss2018/ivis-ss2018-g3-datavis-mobile-survey.pdf

Kadoguchi, M., Hayashi, S., Hashimoto, M., & Otsuka, A. (2019, July). Exploring the dark web for cyber threat intelligence using machine leaning. In *2019 IEEE International Conference on Intelligence and Security Informatics (ISI)* (pp. 200-202). IEEE. 200-202, doi: 10.1109/ISI.2019.8823360.

Kaushal, K. K., Kaushik, S., Choudhury, A., Viswanathan, K., Chellappa, B., Natarajan, S., ... & Dutt, V. (2017, December). Patient journey visualizer: a tool for visualizing patient journeys. In *2017 International Conference on Machine Learning and Data Science (MLDS)* (pp. 106-113). IEEE. doi: 10.1109/MLDS.2017.19.

Kure, H., & Islam, S. (2019). Cyber threat intelligence for improving cybersecurity and risk management in critical infrastructure. *Journal of Universal Computer Science*, *25*(11), 1478-1502. http://www.jucs.org/jucs_25_11/cyber_threat_intelligence_for/jucs_25_11_1478_1502_kure.pdf

Miles, C., Lakhotia, A., LeDoux, C., Newsom, A., & Notani, V. (2014, August). VirusBattle: State-of-the-art malware analysis for better cyber threat intelligence. In *2014 7th International Symposium on Resilient Control Systems (ISRCS)* (pp. 1-6). IEEE. doi: 10.1109/ISRCS.2014.6900103.

Mohaisen, A., & Alrawi, O. (2013, May). Unveiling zeus: automated classification of malware samples. In *Proceedings of the 22nd International Conference on World Wide Web* (pp. 829-832). https://doi.org/10.1145/2487788.2488056

Montasari, R., Carroll, F., Macdonald, S., Jahankhani, H., Hosseinian-Far, A., & Daneshkhah, A. (2021). Application of artificial intelligence and machine learning in producing actionable cyber threat intelligence. *Digital Forensic Investigation of Internet of Things (IoT) Devices*, 47-64. https://doi.org/10.1007/978-3-030-60425-7_3

Noel, S. (2015). Interactive visualization and text mining for the CAPEC cyber attack catalog. In *Proceedings of the ACM Intelligent User Interfaces Workshop on Visual Text Analytics* (pp. 1-8). http://csis.gmu.edu/noel/pubs/2015_CAPEC_viz.pdf

Papalitsas, J., Rauti, S., Tammi, J., & Leppänen, V. (2018). A honeypot proxy framework for deceiving attackers with fabricated content. *Cyber Threat Intelligence*, 239-258. https://doi.org/10.1007/978-3-319-73951-9_12

Preuveneers, D., & Joosen, W. (2021). Sharing machine learning models as indicators of compromise for cyber threat intelligence. *Journal of Cybersecurity and Privacy*, *1*(1), 140-163. https://doi.org/10.3390/jcp1010008

Rodriguez, A., & Okamura, K. (2020). Enhancing data quality in real-time threat intelligence systems using machine learning. *Social Network Analysis and Mining*, *10*, 1-22. https://doi.org/10.1007/s13278-020-00707-x

Saeed, I. A., Selamat, A., & Abuagoub, A. M. (2013). A survey on malware and malware detection systems. *International Journal of Computer Applications*, *67*(16). https://www.researchgate.net/profile/Imtithal-Saeed/publication/272238656_A_Survey_on_Malwares_and_Malware_Detection_Systems/links/566284c608ae192bbf8cf1a5/A-Survey-on-Malwares-and-Malware-Detection-Systems.pdf

Shalaginov, A., Banin, S., Dehghantanha, A., & Franke, K. (2018). Machine learning aided static malware analysis: A survey and tutorial. *Cyber threat intelligence*, 7-45. https://doi.org/10.1007/978-3-319-73951-9_2

Shalaginov, A., Banin, S., Dehghantanha, A., & Franke, K. (2018). Machine learning aided static malware analysis: A survey and tutorial. *Cyber threat intelligence*, 7-45. https://doi.org/10.1007/978-3-319-73951-9_2

Sisaat, K., Kittitornkun, S., Kikuchi, H., Yukonhiatou, C., Terada, M., & Ishii, H. (2017). A Spatio-Temporal malware and country clustering algorithm: 2012 IIJ MITF case study. *International Journal of Information Security*, *16*, 459-473. https://doi.org/10.1007/s10207-016-0342-0

Sisiaridis, D., & Markowitch, O. (2018, April). Reducing data complexity in feature extraction and feature selection for big data security analytics. In *2018 1st International Conference on Data Intelligence and Security (ICDIS)* (pp. 43-48). IEEE. doi: 10.1109/ICDIS.2018.00014

Strom, B. E., Applebaum, A., Miller, D. P., Nickels, K. C., Pennington, A. G., & Thomas, C. B. (2018). Mitre att&ck: Design and philosophy. In *Technical report*. The MITRE Corporation. https://www.mitre.org/sites/default/files/2021-11/prs-19-01075-28-mitre-attack-design-and-philosophy.pdf

Wardman, B., Weideman, M., Burgis, J., Harris, N., Butler, B., & Pratt, N. (2018). A practical analysis of the rise in mobile phishing. *Cyber Threat Intelligence*, 155-168. https://doi.org/10.1007/978-3-319-73951-9_8.

Wetzel, K. (2004). Pebbles-using circular treemaps to visualize disk usage. *URL: http://lip. sourceforge. net/ctreemap. html*, *2*.

Woods, B., Perl, S. J., & Lindauer, B. (2015, October). Data mining for efficient collaborative information discovery. In *Proceedings of the 2nd ACM Workshop on Information Sharing and Collaborative Security* (pp. 3-12). https://doi.org/10.1145/2808128.2808130