IJITCE

ISSN 2347- 3657

International Journal of Information Technology & Computer Engineering

Email : ijitce.editor@gmail.com or editor@ijitce.com

International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 13, Issue 2, 2025
https://doi.org/10.62647/ijitce.2025.v13.i2.pp707-714

PROACTIVE CYBERSECURITY THROUGH THREAT INTELLIGENCE MINING: A SURVEY OF TECHNIQUES AND TRENDS

[1]*Pundukoora Ravi, MCA Student, Department of MCA*

[2] *M G K Priyanka, MCA, (Ph.D), Assistant Professor, Department of MCA*

[12]*Dr KV Subba Reddy Institute of Technology, Dupadu, Kurnool*

ABSTRACT

Cyberattacks have increased in frequency and severity in recent years, necessitating the development of new security measures to fend them off. Traditional security solutions that rely on heuristics and signatures are unable to keep up with the dynamic nature of new-generation threats, which are elusive, resilient, and complicated. In order to avoid attacks or, at the at least, to react swiftly and pro-actively, organisations seek to collect, disseminate, and transform real-time cyber threat information into threat intelligence. The field of cyber threat intelligence (CTI) mining, which finds, gathers, and evaluates important data regarding cyberthreats, is expanding rapidly. But instead of utilising the insights that such new intelligence can provide, the majority of organisations today primarily concentrate on simple use cases, like integrating threat data feeds with already-existing network and firewall systems, intrusion prevention systems, and Security Information and Event Management systems (SIEMs). In this paper, we provide a thorough analysis of current research efforts on CTI mining from various data sources in order to maximise CTI's potential to greatly improve security postures. To be more precise, we offer and develop a taxonomy to categorise the research on CTI mining according to the intended uses (i.e., entities and events related to cybersecurity, cyberattack tactics, techniques, and procedures, hacker profiles, indicators of compromise, vulnerability exploits and malware implementation, and threat hunting), as well as a thorough analysis of the state-of-the-art. Finally, we go over research issues and potential avenues for CTI mining research in the future.

I. INTRODUCTION

Cyber security enemies have improved their tactics to become even more advanced in the aftermath of the significant disruptions brought about by the COVID-driven social, economic, and technical developments of the 2020s. The Solar Winds supply chain assault [1], which shocked several organisations and signalled a sea change in cyber security, was one of several high-profile attacks that followed. Cyber Threat Intelligence (CTI), which is the process of gathering, processing, and evaluating data about the goals, targets, and attack patterns of threat actors, helps governments, businesses, and individual internet users make better, more informed security decisions more quickly and change their behaviour from a reactive to a proactive approach to combating threat actors.

CTI is defined in a number of ways. "Evidence-based knowledge, including context, mechanisms, indicators, implications, and actionable advice about an existing or emerging menace or hazard to assets that can be used to inform decisions regarding the subject's response to that menace or hazard" is one example of what CTI is characterised as. [2]. CTI is defined as "the collection, evaluation, and application of data pertaining to security threats, threat actors, exploits, malware, vulnerabilities, and compromise indicators" in [3]. "Data that has been refined, analysed, or processed such that it is relevant, actionable, and valuable" is how Dalziel et al. [4] define CTI. In general, the CTI pipeline receives raw data regarding cyber

International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 13, Issue 2, 2025
https://doi.org/10.62647/ijitce.2025.v13.i2.pp707-714

security as input, and its output is information that may be used to inform future decisions about proactive cyber security defence, such as tactics for preventing and restricting the scope of cyberattacks.

Organisations of all sizes may better understand their attackers, react to events more quickly, and proactively anticipate what threat actors will do in the near future by utilising CTI to monitor cyber threats. Small and medium-sized businesses profit greatly from CTI data as it gives them access to a degree of security that they were previously unable to get. In the meantime, companies with sizable security teams may use external CTI to save expenses and boost analyst efficacy.

Some research attempts have been undertaken to examine relevant publications, motivated by the growing awareness of the need of proactively aiming to develop cyber resilience. Table II provides an overview of the current CTI surveys. In particular, a research on the dark net as a useful method for keeping an eye on online activity and cyber security threats was provided in the seminar work [5]. In addition to offering a thorough examination of protocols, applications, and dangers using a substantial amount of data, this research [5] classified the components of dark net data as scanning, backscatter, and misconfiguration traffic. The dark net was defined and characterised using case studies including the greatest DRDOS assault, the Conficker worm, and the Sality SIP scan bot net. By examining data taken from the dark net, such as cyberthreats and events, and identifying dark net-related technologies, the article also examined the contributions of dark net measurement. In addition, Robertson et al. [6] suggested a framework based on game theory for simulating an attacker and defender during the CTI mining and analysis process as a security game involving previous attacks and security experts, as well as a system that consists of a crawler, parser, and classifier to find websites where security analysts can obtain information.

Tounsi et al. [7] also divided the threat intelligence kinds that are now in use into three categories: tactical, operational, and strategic. The paper [7] included a thorough analysis of the Tactical Threat Intelligence (TTI) problems, new research trends, and standards, with a primary emphasis on TTI that was primarily produced from the Indicators of Compromise (IOCs). Ibrahim et al. gave a succinct explanation of how to use AI and ML techniques to use CTI to prevent data breaches in light of the developments in AI. In addition, Rahman et al. [11] [12] gave a comprehensive analysis of several machine learning (ML) and natural language processing (NLP) methods for autonomously extracting CTI from textual descriptions. Wagner et al. [8] detailed their inquiry on the state-of-the-art techniques to sharing CTI and the accompanying issues of automating the sharing process with both technical and non-technical obstacles, as using CTI is one of the essential stages to maximising its efficacy. An overview of the concept, problems, and difficulties of CTI was provided by Abu et al. [9]. The present state of the languages and formats accessible for exchanging CTI was compiled by Ramsdale et al. [14]. A sample of CTI feeds was also examined, along with the information they carry and the difficulties in compiling and disseminating them.

In addition to CTI research, government agencies and businesses often employ and implement CTI, which reflects the rising understanding of the vital significance of cyber security. Dedicated teams from these two parties are in charge of gathering, evaluating, and sharing threat intelligence data, often using specialised CTI platforms and tools. For

International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 13, Issue 2, 2025
https://doi.org/10.62647/ijitce.2025.v13.i2.pp707-714

instance, centralised non-profit organisations called Information Sharing and Analysis Centres (ISACs) were created to help its members share CTI and other security-related information. A wide range of sectors and businesses are served by ISACs, including technology, healthcare, financial services, and critical infrastructure. In order to exchange threat information and best practices, as well as to work together on incident response and mitigation initiatives, they bring together organisations from within a certain industry or sector. Governmental organisations and other groups often fund ISACs, which usually adhere to stringent security and privacy guidelines to guarantee that private data is safeguarded and sent only to those who are permitted.

The 2022 Crowd Strike Threat Intelligence Report states that 72% of respondents want to increase their spending on CTI within the following three months in 2022, indicating that it is becoming more and more acknowledged as a useful tool [15]. Both businesses and government agencies are devoting substantial resources to improving their CTI capabilities because they understand that keeping ahead of the ever-changing threat picture requires ongoing adaptation and development. These initiatives include using state-of-the-art technology and processes, forming alliances with other organisations and industry leaders, and building up internal knowledge. The initiatives taken by businesses and government agencies to strengthen their CTI capabilities show a dedication to securing their vital resources and preventing the dangers of cyberattacks. CTI is a key element of a thorough cyber security plan and a vital instrument in the continuous endeavours to protect digital networks and systems for businesses and organisations. Additionally, 75 percent of respondents to the 2022 SANS CTI survey by Brown et al. [13] think that CTI enhances their company's ability to foresee security threats, identify them, and respond to them. Additionally, according to the poll, 52% of participants said that fast and comprehensive information was the most important feature for CTI's future.

Due to the current increase in cyberattacks, several attack artefacts have been actively gathered by various organisations and widely publicised by public web sources [16], [17]. Organisations may enhance their security posture by identifying early warning indicators of risks and continually enhancing their security measures by mining CTI to find evidence-based dangers. The source data for CTI mining may be obtained from both public and private sources, including publicly accessible cyber security reports and technical blogs, as well as internal network logs of the firm. Specifically, the bulk of the CTI is composed of natural language cyber security information. Data pertaining to cyber security may be obtained from a multitude of sources, which serves as a first stage in the process of mining CTI. It may be difficult to mine solid, useful, and authentic CTI while keeping up with the ever growing amount of data pertaining to cyber security. Twenty-one percent of respondents to the 2022 SANS CTI survey [13] do not believe that CTI has improved their organization's overall security posture, despite a favourable trend towards increased levels of context, analysis, and relevance. Nowadays, a lot of companies focus on basic use cases that include integrating threat data feeds with their existing intrusion prevention, network, and firewall systems, as well as Security Information and Event Management (SIEM) systems. They do not, however, take full use of the useful information that such fresh intelligence may provide. In order to create useful tools, it is crucial to investigate CTI mining usage at fine granularities. Specifically, to look at what sort

International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 13, Issue 2, 2025
https://doi.org/10.62647/ijitce.2025.v13.i2.pp707-714

of CTI may be gained by CTI mining, how to execute it, and how to utilise the artefacts that are produced as a proactive defence against cyberattacks. In order to proactively protect against cyber security threats, we provide a thorough literature study on how CTI may be obtained from a variety of data sources, particularly from information presented as natural language texts from many data sources. Although CTI has been thoroughly examined in the prior literature review, this viewpoint has not been examined in the survey studies that are now available.

This paper's main goal is to evaluate current research on CTI mining. Specifically, our study offers an overview of the CTI knowledge acquisition taxonomy and CTI mining methods. A taxonomy that groups CTI mining research according to their goals is presented in our paper. We also provide a thorough review of the most recent CTI mining research. In order to solve these problems, we also look at the difficulties that have been experienced in CTI mining research and recommend future lines of inquiry. An overview of the contributions this publication highlights is provided below:

• Our research outlines a six-step process that uses perception, understanding, and projection to turn cyber security-related data into knowledge based on evidence for proactive cyber security defence using CTI mining.

• In order to proactively protect against cyber attacks, we gather and evaluate the most recent solutions and provide a thorough analysis of the gathered work using the suggested taxonomies based on CTI consumption, paying special attention to attackers' perspectives. We explore new trends and future directions, debate open research concerns and difficulties, and work to broaden the viewpoints of other researchers and CTI communities.

II. LITERATURE SURVEY

Hacioglu Umit et al., 2019 Digital competitive factors and their capacity to change business possibilities in the context of digitisation are essential for effective company performance. This research assessed the changing function of global business automation systems in the context of Industry 4.0 from both a theoretical and a practical standpoint. It seeks to illustrate how cutting-edge technology contributes to a successful supply chain management system.

Cassidy, Daniel et al., 2020 Belfer's National Cyber Capability Index (NCPI), which uses 27 indicators of capability and 32 indications of intent gathered from publicly accessible data, compares the cyber capabilities of 30 nations to seven national objectives. In contrast to current cyber indices, we think there isn't a single way to gauge cyber strength.

Marchetti, Mirco et al., 2021 Though its impact and advancement will last for decades, the digital revolution has permeated every aspect of society. All organisations today depend on intricately linked information systems that gather, process, store, and send data about everything from social and recreational activities to intricate and safety-critical industrial facilities, financial systems, including business-to-business and business-to-consumer services, logistics, and stock markets.

Zahra, Syed Rameem et al., 2022 2020 has produced almost ideal circumstances for hackers due to the world's unrest and instability. The COVID-19 epidemic has altered our lifestyles and caused a widespread shift to digital platforms, while corporations have all but eradicated human encounters. People are now more susceptible to cybercrime as a result of this development. Attackers target victims in order to get money incentives, credentials, or both.

Marchetti, Mirco et al., 2021 Though its impact and advancement will last for decades, the digital revolution has permeated every aspect of society. All organisations nowadays depend on

International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 13, Issue 2, 2025
https://doi.org/10.62647/ijitce.2025.v13.i2.pp707-714

intricately linked information systems that gather, process, store, and send data about everything from social and recreational activities to intricate and safety-critical industrial facilities, financial systems, including business-to-business and business-to-consumer services, logistics, and stock markets [8]. Syed Rameem Zahra et al. (2022) claim that 2020 has produced almost ideal circumstances for hackers due to the world's unrest and instability. The COVID-19 epidemic has altered our lifestyles and caused a widespread shift to digital platforms, while corporations have all but eradicated human encounters. People are now more susceptible to cybercrime as a result of this development. Attackers target victims in order to get money incentives, credentials, or both.

III. SYSTEM ANALYSIS AND DESIGN

EXISTING SYSTEM

Sharing Cyber Threat Intelligence (CTI) has emerged as a new tool in cyber defenders' toolbox to proactively counteract the growing number of cyberattacks. For academics and practitioners, automating CTI sharing and even basic consumption has created new difficulties.

This comprehensive literature review addresses many topic areas of interest related to the broader subject of sharing cyber threat information and examines the state-of-the-art. The recent rise in cyber threat information sharing and the associated difficulties in automating its procedures serve as the driving forces behind this study.

This work focusses on both technical and non-technical difficulties and includes a significant number of pieces from academic and grey literature. Additionally, the results show which subjects were discussed extensively and, as a result, were deemed pertinent by the writers and communities that share cyber threat information.

Disadvantages

Cyber Threat Intelligence (CTI) was not implemented in the system to detect cyberattacks in the current development. Due to a lack of Tactical Threat Intelligence (TTI), this system performs worse.

PROPOSED SYSTEM

• Our research outlines a six-step process that uses perception, comprehension, and projection to turn cyber security-related information into knowledge based on evidence for proactive cyber security defence using CTI mining.

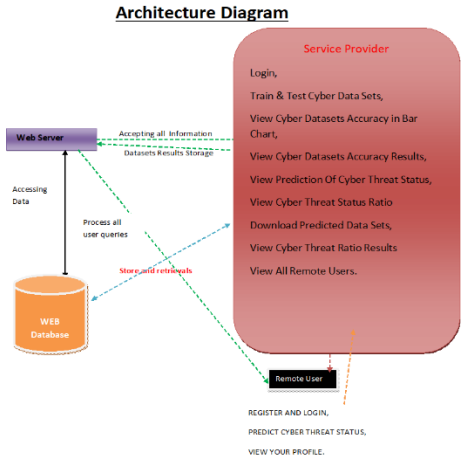
• In order to proactively protect against cyber attacks, we gather and evaluate the most recent solutions and provide a thorough analysis of the gathered work using the suggested taxonomies based on CTI consumption, paying special attention to attackers' perspectives. In an attempt to broaden the viewpoints of other researchers and CTI communities, we talk about difficulties and unresolved research questions in addition to seeing emerging patterns and potential paths forward.

ADVANTAGES

- Cybersecurity-related entities and events: In CTI mining, identifying cybersecurity-related entities and events is similar to diagnosing a certain condition or disease.
- Cyberattack strategies, methods, and practices: The objective of this work category is to analyse the tactics, techniques, and procedures (TTPs) of hackers and cyber threat actors in order to ascertain how they plan and carry out cyberattacks.
- The hacker profiles: Hacker profiles, which identify the source of cyberattacks, make up the third type in our taxonomy of CTI mining.
- The following are signs of compromise: Finding bits of forensic information that show potentially harmful activities on a company's system, such as malware names, signatures, and hashes, is the goal of IoC extraction.

SYSTEM ARCHITECTURE

International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 13, Issue 2, 2025
https://doi.org/10.62647/ijitce.2025.v13.i2.pp707-714



IV. IMPLEMENTATION

MODULES

Service Provider

The Service Provider must use a working user name and password to log in to this module. He may do many tasks after successfully logging in, including training and testing cyber data sets, View the Accuracy Results for Cyber Datasets, View the Accuracy Bar Chart, View the Cyber Threat Status Prediction and the Cyber Threat Status Ratio Get Predicted Data Sets here. View the Results of the Cyber Threat Ratio See Every Remote User.

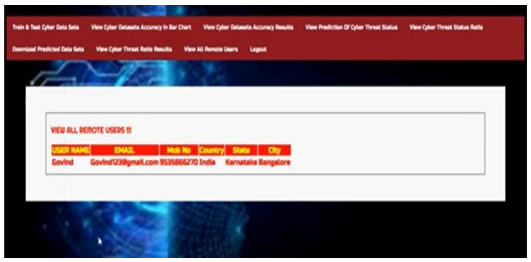
View and Authorize Users

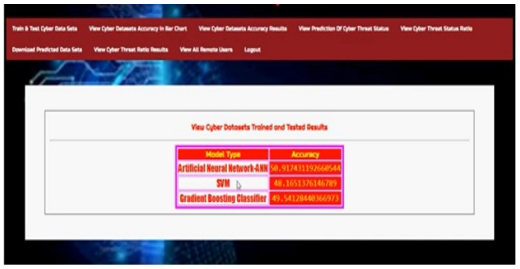
The administrator may see a list of all registered users in this module. Here, the administrator may see the user's information, like name, email, and address, and they can also grant the user permissions.

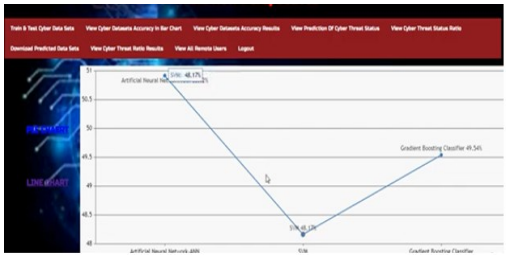
Remote User

A total of n users are present in this module. Before beginning any actions, the user needs register. Following registration, the user's information will be entered into the database. Following a successful registration, he must use his password and authorised user name to log in. Following a successful login, the user may do tasks including registering and logging in, predicting the status of cyber threats, and seeing their profile.

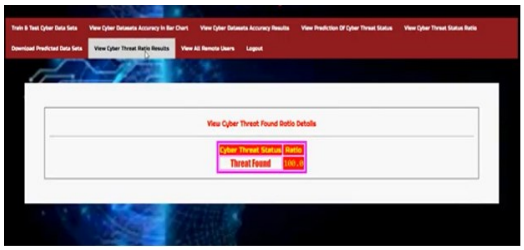
V. SCREEN SHOTS

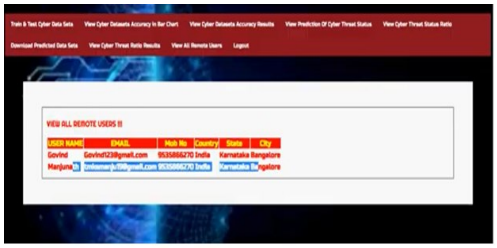


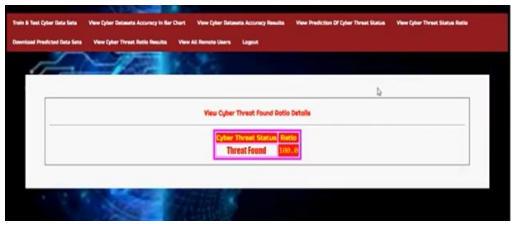




International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 13, Issue 2, 2025
https://doi.org/10.62647/ijitce.2025.v13.i2.pp707-714







VI. CONCLUSION

We provide a thorough analysis of the most important CTI mining publications to date in our study. In our study, we emphasised the technique used by the current studies and provided a categorisation strategy for grouping and classifying existing research works according to the goals of CTI knowledge acquisition. We thoroughly review and discuss current works in accordance with the suggested classification scheme, including entities and events related to cyber security, cyber attack tactics, techniques, and procedures, hacker profiles, indicators of compromise, vulnerability exploits and malware implementation, and threat hunting. We also spoke about prospective future research areas and present problems. CTI mining has attracted a lot of attention in recent decades, particularly for proactive cyber security defence. Numerous individuals are aware that a vast number of fresh models and strategies are created annually. This survey should clarify the most significant developments, help readers grasp the important facets of this topic, and provide insight into potential directions for future study.

REFERENCES

[1] "Solarwinds hackers linked to known russian spying tools, investigators say," https://cybernews.com/news/solarwinds-hackers-linked-to-known-russian-spying-tools-investigators-say/, 2022, accessed on 10/10/2022.

[2] R. McMillan, "Definition: threat intelligence," Gartner.com, accessed on 10/11/2022.

[3] D. Shackleford, "Who's using cyberthreat intelligence and how," SANS Institute, 2015.

[4] H. Dalziel, How to define and build an effective cyber threat intelligence capability. Syngress, 2014.

[5] C. Fachkha and M. Debbabi, "Darknet as a source of cyber intelligence: Survey, taxonomy, and characterization," IEEE Communications Surveys & Tutorials, vol. 18, no. 2, pp. 1197–1227, 2015.

[6] J. Robertson, A. Diab, E. Marin, E. Nunes, V. Paliath, J. Shakarian, and P. Shakarian, Darkweb cyber threat intelligence mining. Cambridge University Press, 2017

[7] W. Tounsi and H. Rais, "A survey on technical threat intelligence in the age of

International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 13, Issue 2, 2025
https://doi.org/10.62647/ijitce.2025.v13.i2.pp707-714

sophisticated cyber attacks," Computers & Security, vol. 72, pp. 212–233, 2018.
[8] T. D. Wagner, K. Mahbub, E. Palomar, and A. E. Abdallah, "Cyber threat intelligence sharing: Survey and research directions," Computers & Security, vol. 87, p. 101589, 2019.
[9] M. S. Abu, S. R. Selamat, A. Ariffin, and R. Yusof, "Cyber threat intelligence–issue and challenges," Indonesian Journal of Electrical Engineering and Computer Science, vol. 10, no. 1, pp. 371–379, 2018.
[10] A. Ibrahim, D. Thiruvady, J.-G. Schneider, and M. Abdelrazek, "Thechallenges of leveraging threat intelligence to stop data breaches,"Frontiers in Computer Science, vol. 2, p. 36, 2020.
[11] M. R. Rahman, R. Mahdavi-Hezaveh, and L. Williams, "What are the attackers doing now? automating cyber threat intelligence extraction from text on pace with the changing threat landscape: A survey," arXivpreprint arXiv:2109.06808, 2021.
[12] ——, "A literature review on mining cyberthreat intelligence from unstructured texts," in 2020 International Conference on Data Mining Workshops (ICDMW 2020). IEEE, 2020, pp. 516–525.
[13] R. Brown and P. Stirparo, "Sans 2022 cyber threat intelligence survey," SANS Institute, 2022.
[14] A. Ramsdale, S. Shiaeles, and N. Kolokotronis, "A comparative analysis of cyber-threat intelligence sources, formats and languages," Electronics, vol. 9, no. 5, p. 824, 2020.
[15] "What is cyber threat intelligence? 2022 threat intelligence report," http s://www.crowdstrike.com/cybersecurity-101/threat-intelligence/, 2022, accessed on 2/13/2023.
[16] N. Sun, C.-T. Li, H. Chan, M. Z. Islam, M. R. Islam, andW. Armstrong, "How do organizations seek cyber assurance?–investigations on the adoption of the common criteria and beyond," IEEE Access, 2022.
[17] N. Sun, J. Zhang, S. Gao, L. Y. Zhang, S. Camtepe, and Y. Xiang, "Data analytics of crowdsourced resources for cybersecurity intelligence,"in Network and System Security: 14th International Conference, NSS 2020, Melbourne, VIC, Australia, November 25–27, 2020, Proceedings 14. Springer, 2020, pp. 3–21.
[18] "Alienvault open threat intelligence," https://otx.alienvault.com/, 2022, accessed on 10/10/2022.
[19] "A community openioc resource," https://openiocdb.com/, accessed on 10/10/2022.
[20] "Iocbucket," https://www.iocbucket.com/, accessed on 10/10/2022.
[21] "Facebook threatexchange," https://developers.facebook.com/products/threat-exchange, 2022, accessed on 10/10/2022.
[22] "National Vulnerability Database," https://nvd.nist.gov/vuln, accessed on 10/10/2022.
[23] "2018 Verizon annual Data Breach Investigations Report," https://www.verizonenterprise.com/verizon-insights-lab/dbir/, accessed on 10/11/2022.
[24] N. Sun, J. Zhang, P. Rimba, S. Gao, L. Y. Zhang, and Y. Xiang, "Datadriven cybersecurity incident prediction: A survey," IEEE communications surveys & tutorials, vol. 21, no. 2, pp. 1744–1772, 2018.
[25] "Defense industrial base cybersecurity information sharing program," https://dibnet.dod.mil/portal/intranet/, 2022, accessed on 10/10/2022