www.ijbar.org
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86
https://doi.org/10.5281/zenodo.14066336

MINING CYBER THREAT INTELLIGENCE FOR PROACTIVE DEFENSE: A SURVEY OF CURRENT TRENDS AND FUTURE DIRECTIONS

[1]K. Kalyani,[2]Bathini Pavan kalyan
[1]Assistant Professor, [2]MCA Student
Department Of MCA Student
Sree Chaitanya College of Engineering, Karimnagar

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 recent 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.

1. INTRODUCTION

IN the wake of the massive disruptions that have been caused by the COVID-driven social, economic, and technological changes of the 2020s, cyber security adversaries have refined their tradecraft to become even more sophisticated. A series of high-profile attacks followed, such as the Solar Winds supply chain attack [1], which rocked many organizations and marked a turning point in cyber security. As the process of collecting, processing, and analyzing information about threat actors' motives, targets, and attack

www.ijbar.org
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86
https://doi.org/10.5281/zenodo.14066336

behaviors, Cyber Threat Intelligence (CTI) assists organizations, governments, and individual internet users in making faster, more informed, data-backed security decisions and changing their behavior in order to fight threat actors from a reactive to a proactive one.

Several definitions exist for CTI. An example of what CTI is defined as is "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" [2]. In [3], CTI refers to "the set of data collected, assessed and applied regarding security threats, threat actors, exploits, malware, vulnerabilities and compromise indicators". Dalziel et al. [4] describe CTI as "data that has been refined, analyzed, or processed such that it is relevant, actionable, and valuable". Generally speaking, the input of the CTI pipeline is the raw data about cyber security, while the output is the knowledge that can help in future decision-making for proactive cyber security defense, including strategies for limiting the extent and prevention of cyber attacks.

By using CTI to observe cyber risks, organizations of all shapes and sizes can better understand their attackers, respond quicker to incidents, and proactively get ahead of what threat actors will do in the near future. For small and medium-sized enterprises, CTI data is of great benefit to them because it allows them to access a level of protection they were previously unable to achieve. Meanwhile, enterprises with large security teams can reduce costs and increase the effectiveness of their analysts by leveraging external CTI.

Driven by the increasing awareness of proactively striving to achieve cyber resilience, some research efforts have been made to review related works. The existing surveys on CTI are summarized in Table II. Specifically, the seminar work [5] presented a study on the dark net as a practical approach to monitoring cyber activities and cyber security attacks. This study [5] defined dark net data components as scanning, backscatter, and misconfiguration traffic, and provided a detailed analysis of protocols, applications, and threats using a large volume of data. Case studies such as Conficker worm, Sality SIP scan bot net, and the largest DRDOS attack were used to characterize and define the dark net. The paper also reviewed the contributions of dark net measurement by analyzing data extracted from it, including cyber threats and events and identified technologies related to the dark net. Additionally, Robertson et al. [6] proposed a system consisting of a crawler, parser, and classifier to locate sites where security analysts can gather information, as well as a game theory-based framework for simulating an attacker and defender in the process of CTI mining and analyzing as a security game involving past attacks and security experts.

www.ijbar.org
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86
https://doi.org/10.5281/zenodo.14066336

Further, Tounsi et al. [7] classified the existing threat intelligence types into strategic threat intelligence, operational threat intelligence, and tactical threat intelligence. With the focus mainly on the Tactical Threat Intelligence (TTI) that was mainly generated from the Indicators of Compromise (IOCs), the work [7] provided a comprehensive study on the TTI issues, emerging research trends, and standards. With the advancements in Artificial Intelligence (AI), Ibrahim et al. provided a brief discussion on how to apply AI and Machine Learning (ML) approaches to leverage CTI to stop data breaches. Rahman et al. [11] [12] further provided a holistic discussion of various technologies in the area of ML and Natural Language Processing (NLP) for automatically extracting CTI from the textual descriptions. As the usage of CTI is one of the key steps to maximizing its effectiveness, Wagner et al. [8] reported the investigation on the state-of-the art approaches to sharing CTI and the associated challenges of automating the sharing process with both the technical and non-technical challenges. Abu et al. [9] gave an overall survey on CTI definition, issues and challenges. Ramsdale et al. [14] summarized the current landscape of available formats and languages for sharing CTI. They also analyzed a sample of CTI feeds, including the data they contain and the challenges associated with aggregating and sharing that data.

Beyond the research works on CTI, the use and implementation of CTI is a common practice in government organizations and enterprises, reflecting the growing recognition of the critical importance of cyber security. These two parties have dedicated teams responsible for collecting, analyzing, and disseminating threat intelligence information, often through specialized CTI platforms and tools. For example, the Information Sharing and Analysis Center (ISACs) are centralized non-profit organizations that are established to facilitate the sharing of CTI and other security-related information among their members. ISACs serve a variety of industries and sectors, including critical infrastructure, financial services, healthcare, technology, and others. They bring together organizations from within a specific industry or sector to share threat intelligence and best practices, as well as collaborate on incident response and mitigation efforts. ISACs are often supported by government agencies and other organizations, and they typically follow strict security and privacy protocols to ensure that sensitive information is protected and shared only among authorized individuals.

According to the 2022 Crowd strike threat intelligence report, CTI is increasingly being recognized as a valuable asset, with 72 percent planning to spend more on it over the next three months in 2022 [15]. Government organizations and enterprises alike are investing significant resources into enhancing their CTI capabilities, recognizing that staying ahead of the constantly evolving threat landscape requires continuous improvement and adaptation. Such efforts include the

www.ijbar.org
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86
https://doi.org/10.5281/zenodo.14066336

development of in-house expertise, the establishment of partnerships with other organizations and industry leaders, and the use of cutting-edge technologies and methodologies. The efforts made by government organizations and enterprises to improve their CTI capabilities demonstrate the commitment to protecting their critical assets and safeguarding against the risks posed by cyber threats. CTI is a crucial component of a comprehensive cyber security strategy and an essential tool in the ongoing efforts to secure digital systems and networks for organizations and enterprises. Furthermore, according to the 2022 SANS CTI survey conducted by Brown et al. [13], 75 percent of the participants believe that CTI improves their organization's security prediction, threat detection, and response. The survey also revealed that 52 percent of the respondents considered detailed and timely information as the most crucial characteristic for the future of CTI.

As a result of the surge in cyber attacks in recent years, a large number of attack artifacts have been reported extensively by public online sources and actively collected by different organizations [16], [17]. By mining CTI, organizations can discover evidence-based threats and improve their security posture by detecting early signs of threats and continuously improving their security controls. The source data for mining CTI can be retrieved from private channels, such as company internal network logs, as well as public channels, such as technical blogs or publicly available cyber security reports. In particular, cyber security information written in natural language comprises the majority of the CTI. Cyber security related data can be gathered from a wide variety of sources, and this provides a stepping stone on the path towards mining CTI. However, mining robust, actionable, and genuine CTI while keeping pace with the rapidly increasing cyber security related information is challenging. Although there is a positive trend towards higher levels of context, analysis, and relevance of CTI, 21 percent of the participants in the 2022 SANS CTI survey [13] do not perceive any improvement in their organization's overall security situation due to CTI. Currently, many organizations concentrate on fundamental usage scenarios that involve merging threat data feeds with their current network and firewall systems, intrusion prevention systems, and Security Information and Event Management systems (SIEMs). However, they do not make the most of the valuable knowledge that such new intelligence can provide. Consequently, it is important to study CTI mining consumption at fine granularities to develop effective tools. To be specific, to investigate what kind of CTI can be obtained through CTI mining, the methodology to achieve it, and how to use the acquired artifacts as proactive cyber security defense . Based on the above motivation, we conduct a comprehensive literature review of how CTI can be acquired from diverse data sources, especially from information written in the form of natural language texts from various data sources, to defend against cyber security attacks proactively. This

www.ijbar.org
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86
https://doi.org/10.5281/zenodo.14066336

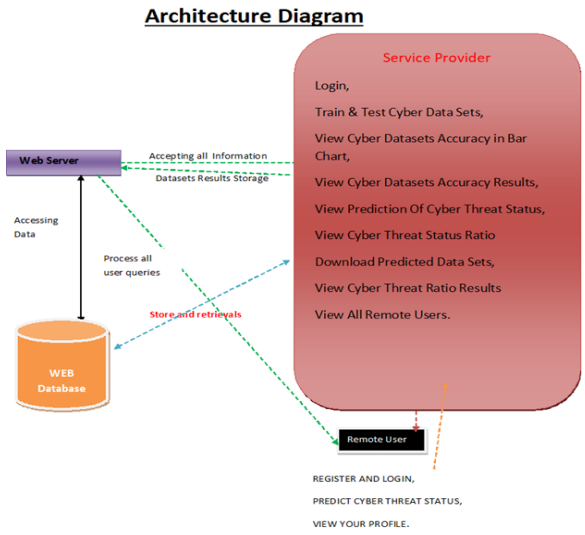
perspective has not been explored in the existing survey works despite the fact that CTI has been extensively studied in the previous literature review.

The primary focus of this paper is to review recent studies on CTI mining. In particular, our work provides a summary of the CTI mining techniques and the CTI knowledge acquisition taxonomy. Our article presents a taxonomy that classifies CTI mining studies based on their objectives. Additionally, we offer a comprehensive analysis of the latest research on CTI mining. We also examine the challenges encountered in CTI mining research and suggest future research directions to address these issues. Below is a summary of the contributions highlighted in this paper:

- Our review summarizes a six-step methodology that transforms cyber security-related information into evidence based knowledge through perception, comprehension, and projection for proactive cyber security defense using CTI mining.
- We collect and review the state-of-the-art solutions and provide an in-depth analysis of collected work with the proposed taxonomies based on CTI consumption, particularly seeing through the eyes of attackers for proactively defending against cyber threats.
- As part of our efforts to expand the perspectives of other researchers and CTI communities, we discuss challenges and open research issues as well as identify new trends and future directions.

As follows is an overview of this survey. Firstly, Section II provides an overview of CTI mining, including its methodology of CTI mining and taxonomy. Section III presents a comprehensive review of existing work in the field of CTI mining according to our proposed taxonomy. Section IV discusses the challenges and future direction in this area. Finally, Section V concludes the paper. Table I lists and describes the acronyms used throughout this paper.

2. SYSTEM ARCHITECTURE



3. EXISTING SYSTEM

- Cyber Threat Intelligence (CTI) sharing has become a novel weapon in the arsenal of cyber defenders to proactively mitigate increasing cyber attacks. Automating the process of CTI sharing, and even the basic consumption, has raised new

www.ijbar.org
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86
https://doi.org/10.5281/zenodo.14066336

challenges for researchers and practitioners.

- This extensive literature survey explores the current state-of-the-art and approaches different problem areas of interest pertaining to the larger field of sharing cyber threat intelligence. The motivation for this research stems from the recent emergence of sharing cyber threat intelligence and the involved challenges of automating its processes.
- This work comprises a considerable amount of articles from academic and gray literature, and focuses on technical and non-technical challenges. Moreover, the findings reveal which topics were widely discussed, and hence considered relevant by the authors and cyber threat intelligence sharing communities.

Disadvantages

- In the existing work, the system did not implement Cyber Threat Intelligence (CTI) for finding cyber attacks.
- This system is less performance due to lack of Tactical Threat Intelligence (TTI).

4. PROPOSED SYSTEM

• Our review summarizes a six-step methodology that transforms Cyber security-related information into evidence based knowledge through perception, comprehension, and projection for proactive cyber security defense using CTI mining.

• We collect and review the state-of-the-art solutions and provide an in-depth analysis of collected work with the proposed taxonomies based on CTI consumption, particularly seeing through the eyes of attackers for proactively defending against cyber threats.

• As part of our efforts to expand the perspectives of other researchers and CTI communities, we discuss challenges and open research issues as well as identify new trends and future directions.

Advantages

- Cybersecurity related entities and events: The identification of cybersecurity-related entities and events in CTI mining is like a diagnosis step that identifies the nature of a particular illness or disease.
- Cyber attack tactics, techniques, and procedures: In this task category, the goal is to determine how cyber threat actors and hackers prepare and execute cyber attacks by analyzing their Tactics, Techniques, and Procedures (TTPs).
- The profiles of hackers: The third category in our taxonomy of CTI mining is called profiles of hackers which trace the origin of cyber attacks.
- Indicators of compromise: The extraction of IoCs aims to find pieces of forensic data that provide evidence of potentially malicious

www.ijbar.org
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86
https://doi.org/10.5281/zenodo.14066336

activity on an organization's system, for example, the names, signatures, and hashes of malware.

5. IMPLEMENTATION

Modules description

Service Provider

In this module, the Service Provider has to login by using valid user name and password. After login successful he can do some operations such as Train & Test Cyber Data Sets, View Cyber Datasets Accuracy in Bar Chart, View Cyber Datasets Accuracy Results, View Prediction Of Cyber Threat Status, View Cyber Threat Status Ratio Download Predicted Data Sets, View Cyber Threat Ratio Results View All Remote Users.

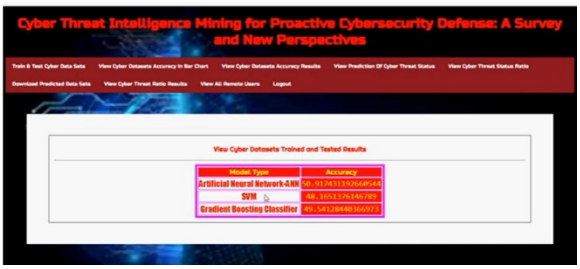
View and Authorize Users

In this module, the admin can view the list of users who all registered. In this, the admin can view the user's details such as, user name, email, address and admin authorizes the users.

Remote User

In this module, there are n numbers of users are present. User should register before doing any operations. Once user registers, their details will be stored to the database. After registration successful, he has to login by using authorized user name and password. Once Login is successful user will do some operations like REGISTER AND LOGIN, PREDICT CYBER THREAT STATUS, VIEW YOUR PROFILE.

6. RESULTS



Index in Cosmos
Nov 2024, Volume 14, ISSUE 4
UGC Approved Journal

www.ijbar.org
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86
https://doi.org/10.5281/zenodo.14066336

7. CONCLUSION

We offered 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 directions and present problems. CTI mining has attracted a lot of attention in recent decades, particularly for proactive cyber security defence. Numerous people 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-t o-known-russian-spying-tools-investigators-say/, 2022, accessed on 10/10/2022.

[2] R. McMillan, “Definition: threat intelligence,” Gartner.com, accessed

https://doi.org/10.5281/zenodo.14066336
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 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, "The challenges of leveraging threat intelligence to stop data breaches," Frontiers in Computer Science, vol. 2, p. 36, 2020.