Practical Cyber Threat and OSINT Analysis based on Implementation of CTI Sharing Platform

Ibrahim Alzahrani , Seokhee Lee , Kyounggon Kim *

Posted Date: 6 May 2024

doi: 10.20944/preprints202405.0277.v1

Keywords: Cybercrime; Cyber investigation; Cyber Threat Intelligence; Indicator of compromised; Malware Information Sharing Platform; Intelligence Sharing Platform

Article

Practical Cyber Threat and OSINT Analysis Based on Implementation of CTI Sharing Platform

Ibrahim Alzahrani †, Seokhee Lee † and Kyounggon Kim *

Center of Excellence in Cybercrime and Digital Forensics, College of Criminal Justice, Naif Arab University for Security Sciences, Riyadh, 14812 Saudi Arabia; ialzahrani@nauss.edu.sa (I.A.); slee@nauss.edu.sa (S.L.)

* Correspondence: kkim@nauss.edu.sa

† These authors contributed equally to this work.

Abstract: Cybercrime Threat Intelligence (CTI) allows us to change our actions from reactive to proactive in the fight against threat actors, and to make more informed, data-driven security decisions. Thus this study proposes the practical implementation of CTI in the Arab world. This study focuses on strengthening CTI by integrating Indicators of Compromise (IoCs) and collecting realistic security alerts from honeypot systems and open source intelligence. The collected data is stored in the Malware Information Sharing Platform (MISP), an open source platform that allows users to create and share IoCs with other organizations while staying informed about new threats. It features an intuitive interface for data analysis and threat identification, facilitating sharing, storage, and correlation of IoCs. Therefore, we leveraged MISP to generate IoCs based on the collected data and analyzed the results to identify potential cyber threats. The implemented platform aims to provide organizations with actionable information to prevent, detect and respond to cybercrime. This study presents a practical approach to strengthening CTI in the Arab world and provides an opportunity to strengthen the region's security posture.

Keywords: cybercrime; cyber investigation; cyber threat intelligence; indicator of compromised; Malware Information Sharing Platform; intelligence sharing platform

1. Introduction

In today's digital age, where our world is highly connected, cyber threats are everywhere, presenting a constant challenge. The move towards using digital technology has brought remarkable advancements, but it also brings a complex network of weaknesses that puts our digital systems at risk [1,2]. Moreover, the rising number and sophistication of cyber-attacks pose significant risks, including data theft, financial losses, and reputational damage [3]. The escalating sophistication and frequency of cyber-attacks demand a strategic and collaborative response, and at the center of this response lies the critical discipline of Cybercrime Threat Intelligence (CTI) [4,5].

Lowenthal argues that intelligence should be viewed as a 'working concept' encompassing three perspectives: process (the intelligence cycle – how information is needed, requested, collected, analyzed and shared), product (the outcome of the process), and organizations (actors participating in information processes) [6,7]. Viewed from this perspective, successful understanding of cyber threats and responding it requires a knowledge base of threat information and effective ways to represent this knowledge [8]. One of the method is the creating own knowledge wherehouse which has Indicators of Compromise (IoC). Additionally, organizations must have a reliable, robust and flexible CTI system to implement these kind of intelligence. Thus we have approached to create our own IoCs database. These system operate without constraints of time, place, or motivation, emphasizing the need for proactive preparedness to prevent or minimize the severity of potential threats.

In this study, we explore the details of creating a Cyber Threat Intelligence sharing platform. The aim is not just to tackle the challenges of today's digital world but to reshape how we work together for cyber security. As we navigate through this research, the core framework supporting our effort is the Malware Information Sharing Platform & Threat Sharing (MISP) [9]. MISP stands not just as a technological foundation but as a figure of collaborative intelligence-sharing, and, this research goes beyond just technical pursuit; it stands as a testament to our dedication to shaping a

 doi:10.20944/preprints202405.0277.v1

cyber landscape that cross boundaries, fosters collaboration, and safeguards the very essence of our interconnected world.

Likewise, recognizing the increasing for the development of CTI platforms, researchers are motivated to create an openly accessible platform that shares details of global cyber attacks [4]. This proposal aims to specifically address cyber threats affecting organizations in Arab countries and the wider region, thereby mitigating their impact and strengthen collaborative efforts in proactive defense measures.

The spread of cyber threats respects no borders. Thus, our research takes on a global perspective with a dedicated focus on addressing the unique cyber security needs of the Arab world. The following describe the objectives and contributions of this research, towards a CTI sharing platform tailored to the cyber security needs of the Arab world.

The primary objective of this research is to conceptualize, design, and implement a CTI sharing platform using the Open Source Platform. This Platform is designed based on MISP Platform to foster collaboration among cyber security entities by establishing a secure ecosystem for efficiently sharing actionable threat intelligence. It aims to enhance organizational situational awareness, enabling proactive defense against evolving cyber threats. CTI Sharing Platform is developed as a cost-free cyber intelligence platform. It actively facilitates the exchange of intelligence indicators for proactive threat mitigation in the region, creating a repository for storing unique cyber threat intelligence data. With a vision to become the leading provider of intelligence indicators to the Arab World, CTI Sharing Platform strives to reinforce regional cyber security and foster collaboration against cybercrime.

This research makes a distinctive contribution by addressing the unique cyber security needs of the Arab world. Currently, there is a conspicuous absence of a dedicated CTI sharing platform tailored to the nuances of the Arab digital landscape. The MISP-based CTI Sharing Platform aspires to fill this void by introducing a solution specifically crafted to meet the challenges and requirements of Arab organizations. This contribution is envisioned to significantly enhance the collective cyber resilience of Arab entities, fostering a collaborative defense against cyber threats. In essence, this research endeavors to bridge a critical gap in the cyber security infrastructure of the Arab world, establishing a foundation for effective threat intelligence sharing and collaboration. By leveraging the MISP platform, the proposed CTI sharing platform is poised to not only enhance the security posture of individual organizations but also contribute to the collective defense against cyber threats on a regional scale. The contributions of this paper include the following:

- We proposed a scheme for the organization to collect information and build a system using open source tools without using expensive commercial CTI systems provided by cyber security companies.
- Utilizing advanced tools and techniques, this research analyzes 1,013,033 data collected from honeypot and 6,877 data from Open Source Intelligence (OSINT) sources to identify patterns and trends indicating potential threats.
- The platform monitors multiple threat intelligence sources to enable security system integration to detect and analyze potential cyber threats specific to the Arab region.

And when our platform is activated and actively used, the following effects can be expected:

- The CTI platform provide the Arab world up-to-date information on potential security threats, aiding in the detection and prevention of cyber-attacks for an overall fortified security posture.
- The platform facilitates collaboration between the Arab world's Cybercrime team and external stakeholders, including other organizations, government agencies, and cyber security experts.
- The CTI platform enables organizations in the Arab world to access timely and precise information regarding their cyber threat landscape. This empowers them to respond promptly and effectively, thus minimizing damage, reducing downtime of critical systems, and formulating robust security posture and response strategies. With this capability, organizations can make informed decisions.

 doi:10.20944/preprints202405.0277.v1

The remainder of this paper is organized as follows: Section 2, we conduct a review of CTI-related research, specifically exploring projects utilizing the MISP platform similar to our study and examining the outcomes achieved. In Section 3, we explain our platform architecture which consists of realtime threat collection from honeypot and OSINT Data Collection method which are two types (Manual Feed and Automatic Module Feed) and method for checking accuracy. We research into the CTI analysis and statistics, assessing the nature of research and the meaningful results generated. Section 4, we describe the implementation of platform developed in our research. Section 5, we focus extended to understanding the perspectives on CTI sharing, and we thoroughly examined the relevant research and outcomes in this regard. We analysed the result and provide brief summary on statistic of collected IoCs from OSINT and honeypot. Section 6, we discuss our limitation and future work for further research. Finally, in Section 7, we present our conclusions.

2. Literature Review

Cynthia Wagner et al. addressed that the IT community is faced with all kinds of incidents, with new threats appearing every day, and that it is almost impossible to respond to these security incidents individually [10]. Thus sharing information about threats across communities has become a key element of incident response to identify attackers, and trusted intelligence resources that provide reliable information can be found within the IT community or the broader intelligence community or fraud detection group. It is essential and in this regard, the author presented the MISP and the Threat Sharing Project.

According to the M. Mutemwa et al. research, in developing countries like South Africa, security and defense role players often lack the necessary capabilities to effectively defend their national cyberspace against fast-moving and persistent threats. The authors state that addressing this challenge requires improved security solutions and increased collaboration within the cyber domain. They emphasized the importance of information sharing as a crucial element in detecting, defending against, and responding to constantly evolving cyber threats and attacks. To address this need, the authors proposed a conceptual CTI sharing model and platform. This model aims to stimulate and enable various stakeholders to seamlessly and collaboratively aggregate, analyze, and share contextually actionable cyber threat information in a timely manner [11].

Abdullahi et al. conducted a systematic literature review of AI methods for detecting cybersecurity attacks in IoT environments, analyzing 80 studies from 2016 to 2021. From the literature review, they found that deep learning and machine learning techniques, especially SVM and RF, are very effective in solving security problems. They also proposed to explore advanced methods such as XGBoost, NN, and RNN to improve detection accuracy [12]. Kattamuri et al. used 51,409 samples including the SOMLAP dataset for static malware detection for cyber threat intelligence. They used colony optimization algorithms wrapped in machine learning, Ant Colony Optimization (ACO), Cuckoo Search Optimization (CSO), and Gray Wolf Optimization (GWO), and achieved an accuracy of 99.37% with 12 features optimized using ACO [13].

Sakellariou et al. defined the core concepts of the CTI framework and presented an eight-layer CTI reference model for advanced system design. The authors validated the proposed model through three case studies and created a CTI reference architecture based on them [14]. The authors examined standardized shared environments for cyber threat intelligence, such as STIX, TAXII, and CybOX, and evaluated their implementations. The authors highlighted various challenges that arise when analyzing threat feeds, identifying data types, and aggregating and sharing data. The study concluded that although standardized shared environments are widely known, real-world adoption is low, with many providers often preferring customized or simple formats [15].

Melo e Silva et al. showed that the cyber security landscape has fundamentally changed over the past few years, with organizations being encouraged to develop the ability to respond to incidents in real time using sophisticated threat intelligence platforms. However, as the field grows rapidly, the concept of CTI today lacks a consistent definition, and a heterogeneous market has emerged that

 doi:10.20944/preprints202405.0277.v1

includes a variety of systems and tools with different capabilities and goals, creating a need for threat intelligence standards. Therefore, the author presented a comprehensive evaluation methodology for intelligence platforms [16].

Borce Stojkovski et al. studied mixed-methods user experience investigation of MISP. Effective incident response in the realm of security relies on standardized CTI as vital threat information. They studied takes a comprehensive approach to security incident response, presenting eighteen key concepts that support the evaluation and establishment of standardized approaches. By analysing six incident response formats, the author identifies their structural elements, highlights characteristics, exposes format deficiencies, and showcases how key concepts aid in selecting the appropriate format for specific use cases. Additionally, the author highlighted an ongoing research task aimed at fully harnessing the potential of incident response. CTI sharing platforms are becoming essential tools for collaborative and cooperative cybersecurity, the focus is often on the technical aspects, incentives or implications associated with CTI sharing instead of examining the challenges experienced by platform users. MISP is an open source CTI sharing platform used by over 6,000 organizations worldwide, as a technologically advanced CTI sharing platform, aims to accommodate a diverse range of security information workers with distinct needs and goals [17].

As mentioned above, many researchers have emphasized the need for a CTI sharing platform, including the MISP platform, and have developed systems that can be used by independent communities. In this context, this study aimed to develop a CTI sharing platform for the Arab world, since organizations in the Arab world do not have a CTI sharing platform. In addition to research on building a CTI sharing platform, much research has been conducted on CTI collection and analysis sharing methods, and many notable research results have been published.

Meanwhile, Abu, M. S. reviewed existing research related to CTI, addressing the most basic question of what CTI is by comparing existing definitions to find commonalities or inconsistencies. They argued that more research is needed to define CTI, as both organizations and vendors lack a complete understanding of what information is considered CTI. It was explained that research institutes such as Financial Services Information Sharing and Analysis Center (FS-ISAC) and MITRE Corporation are developing standard formats for intelligence sharing such as Structured Threat Information eXpression (STIX) and Trusted Automated eXchange of Intelligence Information (TAXII) [18].

Daniel Schlette el al. studied CTI as threat information used for security purposes, requiring standardization in incident response. Authors reviewed a broader security incident response perspective, presenting 18 key concepts and analysing six incident response formats. They identified format defects and explained how to choose the appropriate format using key concepts. Survey results consistently focused on incident response measures in all formats, with playbooks indicating procedures. Various use cases allow organizations to combine formats. The authors also discussed ongoing research to maximize incident response potential [19].

In addition to various studies on the CTI sharing platform, a lot of research is being conducted on intelligence analysis techniques and results. Therefore, we will not only install the CTI sharing platform together with the MISP platform for the Arab world, but also establish procedures and an environment that enable us to collect and analyze actual attack data and provide unique intelligence information.

3. Methodology

Abu, M. S. outlined the process of generating CTI information, systematically categorizing it into five well-recognized steps: planning and direction, data collection, processing, analysis and production, and dissemination [7,20,21]. In alignment with this, our study also follows a five-step methodology; however, we propose the incorporation of an additional classification step specifically tailored for collecting IoCs that are more suited to our unique research model. The CTI data flow lifecycle for fetched IoCs typically involves five stages:

 doi:10.20944/preprints202405.0277.v1

- Research and Planning: We aim to establish a repository of free IoCs for cybercrime threat intelligence in the Arab World. Conduct initial research to comprehend the project scope and objectives, identifying potential data sources and determining specific project requirements.
- Data Collection: Develop mechanisms to collect data, leveraging international OSINT IoCs and security alert from honeypot. Utilize advanced methods to ensure a comprehensive and diverse collection of relevant cyber threat information.
- Analysis: Process and analyze the collected data using appropriate techniques, such as data mining. Identify patterns, anomalies, and trends within the data to gain deeper insights into emerging cyber threats in the Arab world.
- Classification: Establish a database of OSINT IoCs that require classification to derive the most effective IoCs for public sharing. Implement a accurate classification process to enhance the quality and relevance of the shared IoCs.
- Dissemination: Publish the acquired IoCs for free, aiming to become the leading free IoCs provider in the Arab World. Present IoCs with attractive graphics and in a user-friendly format to empower users to easily and effectively utilize IoCs for preventing exposure to the risks of cyber attacks.

Our whole design of system architecture is described in the Figure 1. We use two main methods to collect CTI information. First, we set up honeypot systems to gather data on real and upcoming cyber attacks. Second, we use OSINT, which involves white papers and security alert reports created by trusted organizations. We also gather IoCs shared by verified organizations through social media platforms. Additionally, MISP's Feed Provider function is used to collect and store enhanced CTI information. In this research, we provide a brief explanation of how to generate and share CTI data by obtaining direct threat information using the honeypot system. We also learn about resources that can collect OSINT data and explain our two methods (Manual Collection and Module Collection) for collecting CTI information, including IoCs. In addition, we analyze and explain statistics on data collected in a month recently, as well as statistical information such as attack type, threat actor, and threat actor's tactics.

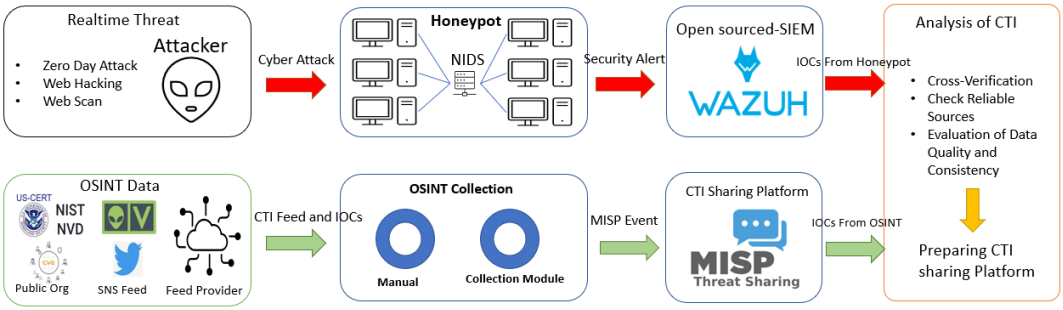

Figure 1. Cybercrime Threat Intelligence Platform.

3.1. Collecting Security Alert from Honeypot

We have deployed 6 honeypot systems as shown in Figure 2. These system are on the cloud and running a virtual web server, which could attract potential attackers. The cloud server is physically located in the MENA region, however presently, the data we used may have some bias as the attacks target the United States and Saudi Arabia due to the relocation of honeypot server from United States to MENA region. Cyber attacks can occur simultaneously across different regions, and similar attack patterns are widespread globally. Despite this, our future plan involves installing physical honeypots in seven locations worldwide to ensure even data collection. By gathering information from diverse global sources, we aim to identify unbiased and accurate trends, providing more precise CTI insights.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2024 doi:10.20944/preprints202405.0277.v1

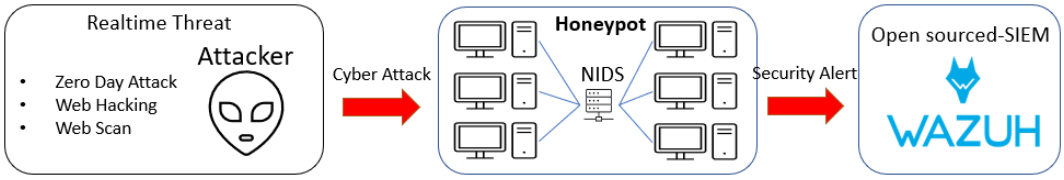

Figure 2. Honeypot System (Collection and Analysis Platform).

To ensure accurate Security Alerts, it is crucial to maintain a robust security policy. While most Security Information Event Monitoring (SIEM) solutions generate a multitude of Security Alerts, ranging from mild to severe attacks, accurately identifying and tracking meaningful alerts pose a distinct challenge compared to the sheer volume of generated alerts. In this study, we formulated a rule policy for the Wazuh system, which gathers security alerts from the honeypot system, aiming to filter data that holds significance as CTI. This initiative helped reduce extraneous data and enabled the monitoring of attack trends.

3.2. Collecting OSINT IoCs

We have deployed MISP systems and collection module as shown in Figure 3. The combat against cybercrime is becoming more tangled every day, demanding collaborative efforts. In the pursuit of shared goals, several OSINT providers have surfaced to combat cybercrime. There are diverse potential sources for gathering IoCs data, they can be categorized in Table 1.

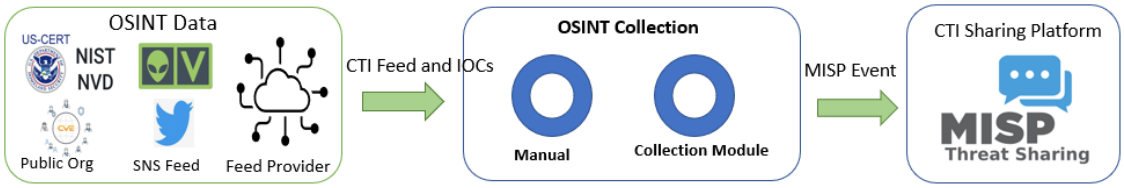

Figure 3. OSINT Data (Collection and Analysis).

Table 1. OSINT Sources.

Type	Description
Threat Intelligence Feeds	These are commercial or open-source feeds that provide IoCs data for various types of threats. Some examples include AlienVault [22], VirusTotal [23], and OpenPhish [24].
Cybersecurity Reports	These reports provide IoCs data on the latest threats and vulnerabilities. Some examples include the Verizon Data Breach Investigations Report, Symantec's Internet Security Threat Report, and the McAfee Threats Report.
Publicly Available Data	IoCs data can also be found in publicly available data, such as security advisories, blog posts, and research papers.
Dark Web Monitoring	Dark web monitoring services can help organizations track IoCs data related to their digital assets that have been compromised and are being sold on the dark web.

To gather IoCs data, we employed two distinct methodologies, classifying data collection into module collection and manual collection. The Table 2 addresses a comparison of the key steps involved in incorporating IoCs into a MISP system, along with the level of automation associated with each feed type. It is evident that manual feeds demand the highest degree of manual effort, whereas module collection feeds necessitate the least manual intervention.

 doi:10.20944/preprints202405.0277.v1

Table 2. OSINT Feed Type.

Type of Feed	Key Steps
Manual Feed	1. Search for an event manually 2. Gather information using OSINT techniques 3. Identify IOCs from the gathered information 4. Add the identified IOCs to the MISP system 5. Classify whether there are more events to investigate.
Collection Module Feed	1. Integrate with sources using an API key 2. Fetch IOCs automatically using a script 3. Add the identified IOCs to the MISP system 4. The system it will determine if there are more IoCs to adding automated to feed the MISP.

4. Implementation

4.1. Honeypot System Deployment

We have deployed honeypot server as shown in the Table 3.

Table 3. Honeypot List.

ID#	IP address	OS	Group
gcc-Bag-server-1	154.xx.xx.128	Ubuntu 22.04.4 LTS	GCC
gcc-Bah-server-1	38.xx.xx.27	Debian GNU/Linux 12	GCC
gcc-Mus-server-1	38.xx.xx.204	AlmaLinux 8.6	GCC
gcc-Dub-server-1	38.xx.xx.166	Rocky Linux 8.9	GCC
gcc-Riy-server-1	38.xx.xx.109	Ubuntu 22.04.4 LTS	GCC
gcc-Kuw-server-1	38.xx.xx.45	Ubuntu 22.04.4 LTS	GCC
Africa-Cai-server-1	38.xx.xx.53	Ubuntu 22.04.4 LTS	Africa

The honeypot server is primarily located in the MENA region, encompassing the GCC area and Africa. By focusing on this specific region, we gather targeted information relevant to our geographical location. This approach enables us to obtain more precise intelligence on cyber threats affecting the MENA region, which are distinct from those in other regions.

4.1.1. Wazuh Security Rule Setting

The detection rules in the Wazuh system are assigned security levels ranging from 0 to 15, predefined based on the threat's risk. Levels 0 to 5 represent a risk from zero to relatively low, while levels 6 to 10 indicate an elevated risk, requiring a response. Levels 11 to 15 signify a high risk of attack, necessitating active response from the security analysis team through additional analysis. For this study, data collection commenced from level 6 to discern genuine trends in security attacks.

4.2. OSINT IOCs Collection

4.2.1. Manual Collection

Gathering information from open sources using various collection methods can serve multiple purposes, including competitive intelligence, research, and security assessments. Here are some common methods for collecting information from open sources:

 doi:10.20944/preprints202405.0277.v1

- Web Crawling: Utilizing software for automated browsing and information collection from websites. This method is efficient for quickly gathering data from numerous websites, although it may not capture all relevant information.
- Search Engines: Platforms like Google and Bing are valuable for finding information on specific topics or entities, providing a quick and easy way to access publicly available information.
- Social Media Monitoring: Platforms such as Twitter, Facebook, and LinkedIn offer valuable insights into individuals or organizations. Social media monitoring tools can track mentions, keywords, and hashtags related to specific topics or entities.
- Public Records Requests: Making requests to government agencies for information related to specific topics or individuals. While time-consuming, this method can provide access to information not available through other sources.
- Online Forums: Platforms like Reddit and Quora offer insights into specific topics or industries, helping identify emerging trends and issues.
- News Aggregators: Services like Google News and Feeds collect news articles related to specific topics or entities, aiding in tracking news and updates over time.
- Data Scraping: Extracting data from web pages using software. This method is efficient for quickly collecting large amounts of structured data, though it may not be legal or ethical in all cases.

Our methodologies used to collect OSINT data include search engines, public records, online forums, and news aggregators, among the methods previously mentioned. This comprehensive approach ensures not only the systematic collection and analysis of CTI information but also emphasizes the importance of providing valuable, freely accessible IOCs to enhance cyber threats across the Arab World. The data was mainly collected in this project can be seen in the Tables 4 and 5.

Table 4. Security Alert Sites.

No	Name	Description
1	US-CERT	The United States Computer Emergency Readiness Team provides security alerts, tips, and resources to protect against cyber threats.
2	CVE	The Common Vulnerabilities and Exposures database provides information on known vulnerabilities in software and hardware products.
3	NIST National Vulnerability Database	Comprehensive database of vulnerabilities maintained by the National Institute of Standards and Technology.
4	The Hacker News	A popular online news outlet for cybersecurity-related news and alerts
5	Threatpost	Another popular online news outlet for cybersecurity-related news and alerts
6	KrebsOnSecurity	Blog maintained by cybersecurity expert Brian Krebs that focuses on cybercrime news and alerts.
7	Dark Reading	Cybersecurity news and analysis site that covers a wide range of topics and trends.

For collecting data, we have joined and researched open-source intelligence communities. These communities provide various sites or tools where security experts share events and indicators they've encountered. These platforms are highly useful for search threat information and security alert reports.

- AlienVault : Alien Vault Open-source Threat Exchange is a group source cybersecurity platform. It has more than 180,000 participants in 140 countries who share more than 19 million potential threats daily. Also, after integration with this platform we have Alerts directly if there is any new attacks or IOCs.

 doi:10.20944/preprints202405.0277.v1

- Google Dorks : Google Dorks OSINT data gathering method using clever Google search queries with advanced arguments [25].

Table 5. OSINT Sites.

No	Name	Description
1	Shodan	Search engine for internet-connected devices
2	ZoomEye	Search engine for internet-connected devices and web applications.
3	Censys	Search engine for internet-connected devices and web applications
4	Whois	Domain registration lookup tool.
5	Google Dorks	Advanced search operators for finding sensitive information online
6	Social Media Platforms	Meta, X, LinkedIn, etc. for gathering information about individuals or organizations.
7	Wayback Machine	Digital archive of the internet that allows you to view historical versions of websites.

4.2.2. Collection Module

Peter Amthor et al. addressed that effective cyber security management necessitates a prompt and cost-efficient response to all threat alerts [26]. Utilizing automated cyber threat detection and incident response proves to be an efficient approach to promptly address real threats. Thus, there is a need for automated tools for threat detection, such as threat intelligence sharing platforms and security policy control systems. These tools encompass various technologies, methods, and instruments aimed at responding to threat occurrences and events. In our study, we implemented the automatic collection of OSINT data and are currently exploring methods to automatically incorporate the gathered IoCs into security policies and equipment.

The process performed involved setting up a MISP instance and generating a MISP API key. Then, we created a Twitter developer account and obtained API credentials. Based on this, we developed methods which is the python script to integration by API key that uses the sntwitter library to get feeds from twitter (X.com). In other words, data was transmitted to MISP through API and the confirmation process for the retrieved data was implemented using a python script. Overall, the dataflow process in this script can be represented as follows:

1. Input node (imported modules)
2. Processing node (classify_iocs function)
3. Output node (MISP instance)
4. Transformation node (get_query_date_range function)
5. Processing node (Iterate through tweets and classify IOCs)
6. Output node (MISP instance)

The detailed Algorithm 1 show the pseudocode that we used to collect the data from Twitter (X.com).

The results of conversion of data collected by the Collection Module are as follows. Figure 4 shows the example of twitter post which provide IoCs to followers. Figure 5 shows the IoCs on the MISP platform after we collect data from twitter (X.com).

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2024 doi:10.20944/preprints202405.0277.v1

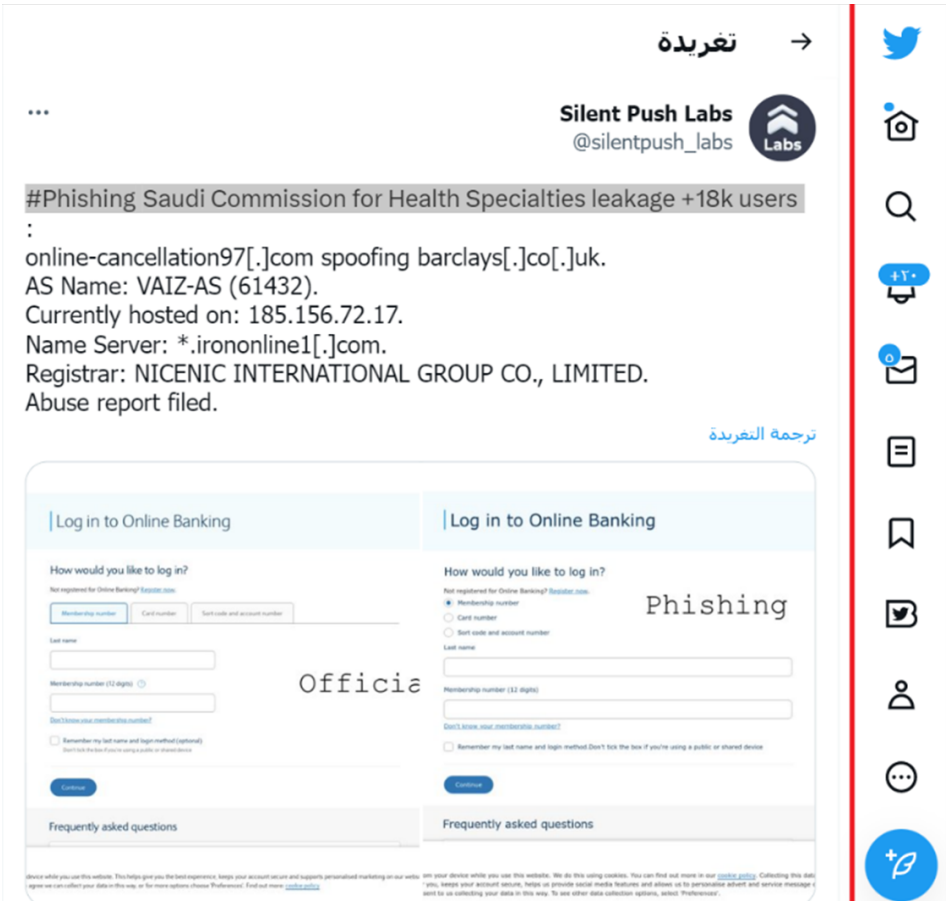

Figure 4. IoCs Source from twitter.

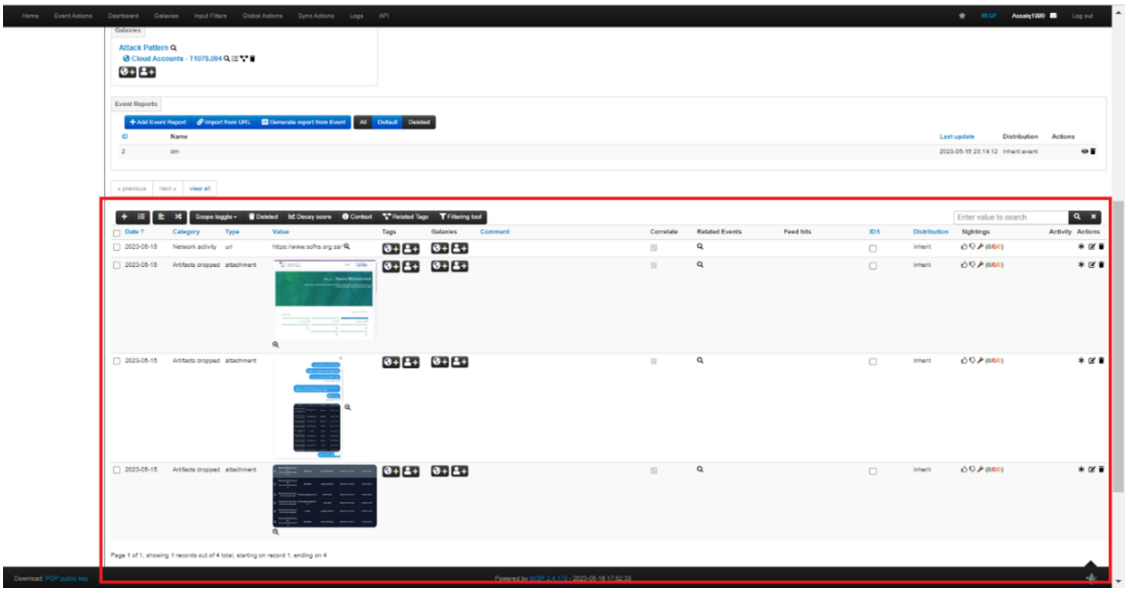

Figure 5. IoCs examples screen.

 doi:10.20944/preprints202405.0277.v1

Algorithm 1 Class Cyber Threat Monitor

Define class CYBERTHREATMONITOR
Define method __INIT__
Set *twitter_api* to method call SETUP_TWITTER_API
Set *misp_api* to new PYMISP instance with *URL*, *key*, and *verify* parameters
Define *tags* list with "#phishing", "#malware", "#infosec", "#cybersecurity", "#ransomware", "#APT", "#zeroDay", "#dataBreach", "#hacking", "#cybercrime"
Define method SETUP_TWITTER_API
Create *auth* with Twitter *consumer key* and *secret*
Set *access token* and *secret* on *auth*
return new TWEEPY API instance with *auth*, set to wait on rate limit
Define method FETCH_TWEETS with parameter *query*
Create a cursor with TWEEPY.CURSOR to search tweets using API with *query*, *tweet mode* extended, and *language* English
return cursor *items* up to 100
Define method EXTRACT_IOCS with parameter *tweet*
if tweet has 'retweeted_status' **then**
    Use that text
else
    Use tweet's full text
end if
Extract *URLs* from text using IOCEXTRACT with *refang* True
Extract *IPs* from text using IOCEXTRACT
Extract *SHA256 hashes* from text using IOCEXTRACT
Extract *MD5 hashes* from text using IOCEXTRACT
return dictionary with lists of *URLs*, *IPs*, *SHA256s*, and *MD5s*
Define method REPORT_TO_MISP with parameters *iocs* and *tweet_info*
for each *ioc_type* and *iocs_list* in *iocs* **do**
    for each *ioc* in *iocs_list* **do**
        Create a new MISP event with *info* "Twitter-based IOC alert"
        Add attribute to event with *ioc_type*, *ioc* value, and a *comment* with tweet's user
        Submit event to MISP API
    end for
end for
Define method MONITOR_TWEETS
Join *tags* with " OR " and append "-filter:retweets -filter:replies" to form *query*
Set *tweets* to result of method call FETCH_TWEETS with *query*
for each *tweet* in *tweets* **do**
    Set *iocs* to result of method call EXTRACT_IOCS with *tweet*
    Set *tweet_info* to dictionary with *user* and *date* from *tweet*
    Call method REPORT_TO_MISP with *iocs* and *tweet_info*
end for

5. Analysis

The platform focuses on optimizing threat response mechanisms to minimize the impact of cyber incidents and promotes interoperability by ensuring compatibility with existing cyber security infrastructures, encouraging widespread adoption across diverse organizational settings.

The gathered IoCs are subsequently archived in a dedicated section within the our database, established explicitly for this research initiative. The overarching objective of these five steps is to methodically classify and share IoCs, thereby augmenting our capacity to respond promptly and efficiently to the most pressing threats confronting the Arab World. This strategic approach enhances the overall cyber security posture and resilience of the region.

Gong, S' conducted research that CTI information can also be applied to security systems in internal IT and OT infrastructure, such as IoT (Internet of Things) and Supervisory Control and Data Acquisition (SCADA) networks. And the performance of a security system depends on the accuracy of the data, and they provided data accuracy results for four CTI feeds using approximately 40,000 data sets [27]. Likewise, there is a need to confirm the accuracy of the IoCs collected. In general, there are three methods to check data accuracy: Cross-Verification, Check Reliable Sources, and Evaluation of Data Quality and Consistency.

- Cross-Verification: We meticulously compared the gathered OSINT information with data acquired from various independent sources. Utilizing the MISP platform's functionality, we

 doi:10.20944/preprints202405.0277.v1

establish connections when identical IOCs are found across events in the standard data format. When identical IOCs were identified, the information was considered more accurate, having been corroborated by multiple sources.

- Check Reliable Sources: We conducted thorough checks to ascertain the credibility of the information, verifying its origin from reputable and reliable sources. Information sourced from certified organizations, government agencies, or trusted experts is deemed more likely to be accurate.
- Evaluation of Data Quality and Consistency: After the distribution of CTI data to members of the MISP community, the members assess the quality and consistency of the collected data. In cases of inconsistencies, contradictions, or inappropriate data, concerns about the reliability of the information are raised, prompting suggestions for modification and revision through the functionalities provided by the MISP platform.

Inside of Collected Data

The data collected during the month of January 2024 totaled 1,013,033 threats which was from honeypot and OSINT recorded during the month. And as seen in the Table 6, the data collected from the honeypot shows in monthly total 1,006,156 attacks and shows a concerning increase every week, and 323,895 attacks in the final week.

Table 6. Threat from Honeypot.

Period (JAN)	Week1	Week2	Week3	Week4	Total
No of Attack	218,571	232,425	231,265	323,895	1,006,156

The data provided includes counts for a total of 140 countries, each associated with the corresponding sum of counts representing cyber-attack metrics. However we describe the top 10 countries in the chart, due to space constraints. China emerges as the leading contributor with 365,949 attacks, emphasizing the global nature of cyber security threats. The data further delineates the significant role played by the United States, Japan, and other nations in the cyber threat landscape.

Table 7. Top 10 Countries of Attack Sources.

Top 10 Countries	Count of Attack
China	83624302
United States	53412711
Japan	46020758
Singapore	16261275
South Korea	11387451
India	11367809
Russia	10556219
Germany	7122717
Brazil	6257157
Hong Kong	5922749

The attacker's tatics and techniques were as shown in the Table 8. Credential Access stands out as a prevalent method, emphasizing the importance of securing user credentials. The data reveals the multifaceted nature of attacks, covering Defense Evasion, Initial Access, Lateral Movement, Persistence, Privilege Escalation, and Reconnaissance. Understanding these tactics is crucial for developing effective defense strategies. Regarding the techniques, it shows that password Guessing, with a staggering count of 3,229,147, signals a substantial threat to password security. Brute Force attacks, Exploits on Public-Facing Applications, and SSH-based intrusions highlight the diversity of techniques employed.

 doi:10.20944/preprints202405.0277.v1

This information is pivotal for cybersecurity professionals to implement targeted countermeasures against specific attack vectors.

Table 8. Types of Tatics and Techniques

Tactics	Counts	Techniques	Counts
Credential Access	2054265	Password Guessing	3229147
Lateral Movement	1681643	SSH	2927690
Defense Evasion	1246204	Valid Accounts	2492094
Privilege Escalation	1246204	Brute Force	71165
Initial Access	1246162	Vulnerability Scanning	252
Persistence	1246047	Process Injection	238
Reconnaissance	252	File and Directory Discovery	205
Discovery	167	Exploit Public-Facing Application	153

As seen in the Table 9, there are a total of 6,877 OSINT information through Twitter, and most of the information is mainly about phishing and malware. Malware entries include Qakbot, Njrat, GootLoader, RedLine, Remcos, Dcrat, AsyncRAT, AgentTesla, IcedID, SocGholish, BazarLoader, and Lazarus in Table 10. This intelligence can help security team to identify active malware. Based on this intelligence, security experts can develop customized response strategies and respond immediately to malware. The following is a rough description of the threat group and malware we identified in this study. Since it represents a threat group or malware currently active, the intelligence offers security organizations insights that can aid in the effective allocation of limited resources for preemptive action.

Table 9. Data from X (Twitter).

Type of threat	Malware	Phishing	Ransomware	Scam	Total
Count	1416	5308	5	148	6877

Table 10. Top Malwares collected our CTI platform.

Name	**Description**
Njrat	Remote Access Tool (RAT), has been demonstrating enhanced techniques and more sophisticated attacks recently. New variants are also emerging.
GootLoader	Focusing on spreading Trojan horses, has been incorporating new propagation and evasion techniques to make detection more challenging.
RedLine	Malware emphasizing information theft and malicious activities, tends to enhance evasion capabilities with advanced concealment technologies.
Remcos	Remote access tool, has been strengthening its capabilities with various encryption technologies and features to bypass detection.
Dcrat	Multipurpose malware, has seen an increase in campaigns using new phishing and social engineering techniques to deceive users.
AsyncRAT	Lightweight remote access tool, provides higher flexibility by adding new command and control functionalities recently
AgentTesla	Spyware primarily focusing on keylogging, is introducing more sophisticated theft and evasion techniques.
IcedID	Trojan horse targeting financial institutions, is adopting advanced campaigns and new infection techniques.
SocGholish	Phishing kit, is expanding its impact by utilizing a variety of attack vectors recently.
BazarLoader	Loader propagating various payloads, is prioritizing evasion through diverse propagation and advanced hiding features
Lazarus	Nation-state actor engaged in advanced persistent cyber attacks, has been gaining international attention with rapid and continuous cyber operations.

 doi:10.20944/preprints202405.0277.v1

6. Limitation and Future work

The real-data collection and analysis methodology for practical Cybercrime Threat Intelligence focuses on putting a threat intelligence project into action. The methodology encourages always keeping an eye on things and making things better to deal with changing cybercrime threats. However, in this context, currently honeypot system is located in the MENA region but future research will focus on gathering real-time information on cyber threats by establishing honeypot system physically located in the all over the world, expanding out platform to gather information about direct cyber attacks against the world. Our plan involves developing an evaluation methodology to offer more precise and accurate intelligence based on the currently collected data, considering criteria such as attacker identity, attacker goals, execution plans and methods, and indicators for tracing execution. Through this approach, we aim to provide not only refined intelligence information but also visibility into attack patterns and trends targeting the Arab world.

Kris Oosthoek and Christian Doerr's research [28] highlighted the need for a CTI framework. In their research, they investigated the application and impact of these frameworks on reporting analysis results, particularly in the context of reproducing CTI reports for APT malware, and aimed to ensure accurate CTI sharing and distribution in the context of the rapid increasing of new malware samples every day. The importance of behavioral labeling was emphasized. Likewise, our future research will also analyze the data collected from Honeypot and aim to share information through accurate labeling on the CTI sharing platform. We plan to apply machine learning algorithms to classify and cluster security alerts by attacker to generate more valuable intelligence from security alert data.

In trying out this approach for a cybercrime threat intelligence project, we experiment with different ways to gather data, analyze it, and show it visually to get useful insights from real-world information. We keeps refining and adjusting the methodology based on feedback and how the threat landscape is changing. This way of doing things lets everyone involved test ideas, check results, and make things better using real evidence and practical observations. Finally, our goal is to generate and share direct CTI information related to cybercrime, allowing members of the Arab World to promptly integrate it into their security systems and policies.

We are also considering the promotion of the implemented CTI platform, currently in temporary operation, in the Arab world. The intention is to use it as a platform to launch the tentatively named Arab CTI Community (ACC). To achieve this, it is crucial to encourage the participation of major organizations in Arab countries and establish a relationship and atmosphere conducive to freely sharing CTI information obtained by each organization. Above all, for the effective activation of the platform, someone needs to consistently share CTI data that can serve as priming material. Thus we are willing to take on that role and subsequently, when a consensus is reached among major organizations in Arab countries, the plan is to officially launch the community.

7. Conclusions

In this study, we examined the life cycle of CTI and the CTI sharing platform practically implemented based on MISP platform. Additionally, we outlined the scheme of platform based on open sourced security system without commercial CTI platform and the methods for collecting data to enhance CTI platform and highlighted the acquisition of meaningful intelligence through a concise statistical analysis of the gathered data. Especially, analysis from real threat alerts and automatic collection module implemented to gather OSINT data is distinguish point to other research.

As conclusion, this research has played a crucial role in furnishing intelligence information aimed at safeguarding against potential future cyber attacks or mitigating their risks. It is meaningful that our research is the start to share the CTI information freely in the Arab world. By anticipating potential threat and motivating to proactively implement appropriate measures, the initiative would significantly contribute to decision-making at the forefront. Notably, it will be expected to stand as the first free intelligence information provider within the MENA region, and want to extend its services to member states and the international community.

 doi:10.20944/preprints202405.0277.v1

Author Contributions: Conceptualization, S.L; methodology, K.K.; formal analysis, I.A; investigation, K.K, S.L, and I.A.; resources, I.A.; writing—original draft preparation, S.L, and I.A; writing—review and editing, K.K, S.L., and I.A.; supervision, S.L, K.K.; project administration, K.K.; funding acquisition, K.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research work has received funding from the Security Research Center of Naif Arab University for Security Sciences, under grant agreement no NAUSS-23-R12.

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

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

References

1. Kim, K.; Alshenaifi, I.M.; Ramachandran, S.; Kim, J.; Zia, T.; Almorjan, A. Cybersecurity and cyber forensics for smart cities: a comprehensive literature review and survey. *Sensors* **2023**, *23*, 3681.
2. Kim, K.; Alfouzan, F.A.; Kim, H. Cyber-attack scoring model based on the offensive cybersecurity framework. *Applied Sciences* **2021**, *11*, 7738.
3. Jajodia, S.; Samarati, P.; Yung, M. Encyclopedia of Cryptography, Security and Privacy, 2019.
4. Kotsias, J.; Ahmad, A.; Scheepers, R. Adopting and integrating cyber-threat intelligence in a commercial organisation. *European Journal of Information Systems* **2023**, *32*, 35–51.
5. Van Haastrecht, M.; Golpur, G.; Tzismadia, G.; Kab, R.; Priboi, C.; David, D.; Răcătăian, A.; Baumgartner, L.; Fricker, S.; Ruiz, J.F.; others. A shared cyber threat intelligence solution for SMEs. *Electronics* **2021**, *10*, 2913.
6. Lowenthal, M.M. *Intelligence: From secrets to policy*; CQ press, 2022.
7. Ainslie, S.; Thompson, D.; Maynard, S.; Ahmad, A. Cyber-Threat Intelligence for Security Decision-Making: A Review and Research Agenda for Practice. *Computers & Security* **2023**, p. 103352.
8. Mavroeidis, V.; Bromander, S. Cyber threat intelligence model: an evaluation of taxonomies, sharing standards, and ontologies within cyber threat intelligence. 2017 European Intelligence and Security Informatics Conference (EISIC). IEEE, 2017, pp. 91–98.
9. MISP. MISP Open Source Threat Intelligence Platform & Open Standards for Threat Information Sharing. https://www.misp-project.org/, 2024. Accessed on February 4, 2024.
10. Wagner, C.; Dulaunoy, A.; Wagener, G.; Iklody, A. Misp: The design and implementation of a collaborative threat intelligence sharing platform. Proceedings of the 2016 ACM on Workshop on Information Sharing and Collaborative Security, 2016, pp. 49–56.
11. Mutemwa, M.; Mtsweni, J.; Mkhonto, N. Developing a cyber threat intelligence sharing platform for South African organisations. 2017 Conference on Information Communication Technology and Society (ICTAS). IEEE, 2017, pp. 1–6.
12. Abdullahi, M.; Baashar, Y.; Alhussian, H.; Alwadain, A.; Aziz, N.; Capretz, L.F.; Abdulkadir, S.J. Detecting cybersecurity attacks in internet of things using artificial intelligence methods: A systematic literature review. *Electronics* **2022**, *11*, 198.
13. Kattamuri, S.J.; Penmatsa, R.K.V.; Chakravarty, S.; Madabathula, V.S.P. Swarm optimization and machine learning applied to pe malware detection towards cyber threat intelligence. *Electronics* **2023**, *12*, 342.
14. Sakellariou, G.; Fouliras, P.; Mavridis, I.; Sarigiannidis, P. A reference model for cyber threat intelligence (CTI) systems. *Electronics* **2022**, *11*, 1401.
15. Ramsdale, A.; Shiaeles, S.; Kolokotronis, N. A comparative analysis of cyber-threat intelligence sources, formats and languages. *Electronics* **2020**, *9*, 824.
16. de Melo e Silva, A.; Costa Gondim, J.J.; de Oliveira Albuquerque, R.; García Villalba, L.J. A methodology to evaluate standards and platforms within cyber threat intelligence. *Future Internet* **2020**, *12*, 108.
17. Stojkovski, B.; Lenzini, G.; Koenig, V.; Rivas, S. What's in a Cyber Threat Intelligence sharing platform? A mixed-methods user experience investigation of MISP. Annual Computer Security Applications Conference, 2021, pp. 385–398.
18. Abu, M.S.; Selamat, S.R.; Ariffin, A.; Yusof, R. Cyber threat intelligence–issue and challenges. *Indonesian Journal of Electrical Engineering and Computer Science* **2018**, *10*, 371–379.

 doi:10.20944/preprints202405.0277.v1

19. Schlette, D.; Caselli, M.; Pernul, G. A comparative study on cyber threat intelligence: The security incident response perspective. *IEEE Communications Surveys & Tutorials* **2021**, *23*, 2525–2556.
20. Abu, M.S.; Selamat, S.R.; Yusof, R.; Ariffin, A. Comparative Study of Cyber Threat Intelligence Framework. 2nd Global Conference on Computing and Media Technology, 2018.
21. Kime, B. Cyber Threat Intelligence Support to Incident Handling, 2017.
22. AlienVault Open Threat Exchange. https://otx.alienvault.com/dashboard/new. Accessed on February 29, 2024.
23. VirusTotal. https://www.virustotal.com/. Accessed on February 29, 2024.
24. OpenPhish. https://openphish.com/. Accessed on February 29, 2024.
25. GoogleDorks. https://www.exploit-db.com/google-hacking-database. Accessed on February 29, 2024.
26. Amthor, P.; Fischer, D.; Kühnhauser, W.E.; Stelzer, D. Automated cyber threat sensing and responding: integrating threat intelligence into security-policy-controlled systems. Proceedings of the 14th International Conference on Availability, Reliability and Security, 2019, pp. 1–10.
27. Gong, S.; Cho, J.; Lee, C. A reliability comparison method for OSINT validity analysis. *IEEE Transactions on Industrial Informatics* **2018**, *14*, 5428–5435.
28. Oosthoek, K.; Doerr, C. Inside the matrix: CTI frameworks as partial abstractions of complex threats. 2021 IEEE International Conference on Big Data (Big Data). IEEE, 2021, pp. 2136–2143.