ORISHA: Improving Threat Detection through Orchestrated Information Sharing (Discussion Paper)

Luca Caviglione[1], Carmela Comito[2], Massimo Guarascio[2,*], Giuseppe Manco[2], Francesco Sergio Pisani[2] and Marco Zuppelli[1]

[1]Institute for Applied Mathematics and Information Technologies, Via de Marini 6, Genova, 16149, Italy
[2]Institute for High Performance Computing and Networking, Via P. Bucci 8-9/C, Rende, 87036, Italy

Abstract

The exponential growth in the number of cyber threats requires sharing in a timely and efficient manner a wide range of Indicators of Compromise (IoCs), i.e., fragments of forensics data that can be used to recognize malicious network or system activities. To this aim, a suitable architecture is required, especially to distribute and process the various IoCs. Unfortunately, the continuous creation of offensive techniques, along with the diffusion of advanced persistent threats, imposes the ability to update and extend the platform used to manage the multitude of IoCs collected in the wild. In this paper, we present the ORISHA architecture, which takes advantage of a distributed threat detection system to match performance and scalability requirements. The paper also discusses how the platform can be extended to handle the most recent "stealthy" malware as well as campaigns aimed at spreading fake news.

Keywords

Threat Intelligence, Risk Mitigation, Active Learning, Collaborative Approach

1. Introduction

In recent years, we observed exponential growth in the number of attacks targeting organizations and users. Successful attacks performed by Blackhats were able to provoke a wide variety of damages and proved the weakness (in terms of security) of both government computer systems as well as user devices. As reported in [1], DDoS, information leakage, phishing, identity theft, and botnet were among the most frequent attacks performed in 2020, and the outbreak of the pandemic emergency has done nothing but further exacerbate this complex scenario. The vulnerabilities of popular platforms, applications, and systems discovered during this critical period have fed the interest in employing information-sharing technologies to increase attack detection and risk mitigation capabilities of enterprises and organizations [2, 3].

Quick decisions and adequate countermeasures can be set up if information concerning

SEBD 2023: 31st Symposium on Advanced Database System, July 02–05, 2023, Galzignano Terme, Padua, Italy

*Corresponding author.

luca.caviglione@ge.imati.cnr.it (L. Caviglione); carmela.comito@icar.cnr.it (C. Comito); massimo.guarascio@icar.cnr.it (M. Guarascio); giuseppe.manco@icar.cnr.it (G. Manco); francescosergio.pisani@icar.cnr.it (F. S. Pisani); marco.zuppelli@ge.imati.cnr.it (M. Zuppelli)
0000-0001-6466-3354 (L. Caviglione); 0000-0001-9116-4323 (C. Comito); 0000-0001-7711-9833 (M. Guarascio); 0000-0001-9672-3833 (G. Manco); 0000-0003-2922-0835 (F. S. Pisani); 0000-0001-6932-3199 (M. Zuppelli)

CEUR Workshop Proceedings (CEUR-WS.org)

threat events and Indicators of Compromise (IoCs) is shared promptly. Specifically, proactive threat information sharing and defensive mitigation strategies can be exploited to boost the resilience of the entities belonging to trusted communities generating herd immunity against new (possibly unknown) threats. Therefore, an emerging research line focuses on devising new platforms, approaches, and methodologies to deliver and share threat events to prevent further damage by quickly arranging countermeasures.

Recently, *Cyber Threat Intelligence* (CTI) platforms have proved their effectiveness in managing threat information [4]. These tools are currently adopted to gather, preprocess, enrich, correlate, analyze, and share threat events [5]. As highlighted in [6], *Threat Intelligence Platforms* (TIP) have to satisfy some main requirements, i.e., providing (*i*) services for information sharing, (*ii*) facilities for automatizing the process, and (*iii*) functionalities for collaborative threat data analysis. Alas, devising a complete solution for handling information from different sources is a challenging objective [7]. Indeed, different open issues (e.g., standardization, privacy, and reliability of the shared information, just to cite a few) have to be addressed to realize a fully operational platform. Although some recent works propose advanced solutions for easing threat data sharing, many only focused on some of the issues mentioned above [8]. A comprehensive review of the current state of the art and open challenges can be found in [9].

In this work, we provide an overview of ORISHA [10], a platform for ORchestrated Information SHaring and Awareness that combines TIPs with AI-based Threat Intelligence solutions in a single comprehensive framework. ORISHA allows for improving the accuracy of Threat Detection Systems (TDS) in recognizing incoming attacks and also enables the sharing of reliable and relevant threat information among organizations and threat detection algorithms. The main idea is that TDSs can benefit each other mutually by sharing knowledge since a threat feed produced by a TDS can be exploited to improve the threat modeling strategies of another one. The platform allows for publishing threat information on a distributed TIP and making them accessible to other actors. Although the current implementation is fully general, ORISHA has been mainly used for Network Intrusion Detection Systems. In this work, we discuss how ORISHA can also be employed to mitigate the risk of new emerging threats, such as information-hiding-based attacks and spreading fake news.

The rest of the paper is structured as follows. Section 2 surveys state-of-the-art solutions for threat information sharing and awareness. Section 3 describes ORISHA, our platform for threat event sharing. Section 4 introduces the new threats to be managed and discusses how ORISHA can be extended. Finally, Section 5 concludes the paper and outlines future research directions.

2. Background

Threat Intelligence refers to the task of gathering data concerning attacks or breaches (e.g., context, methods, indicators, or devices) with the aim to help organizations to set up effective countermeasures by leveraging a wide range of information [11]. Specifically, organizations can cooperate to improve the detection and prevention of new threats by sharing information about recently identified attacks. In this respect, *Indicators of Compromise* (IoCs) are the mean typically used to share this information. An IoC is a piece of forensic data identifying potentially malicious activities on a system or network. The IP address of a DoS attack, a hash of a malicious executable file or the URL of a phishing website are examples of IoCs.

Threat Intelligence is a relatively new research line in the field of cybersecurity and, as reported in [12], both academic and industrial entities have shown a growing interest in this topic. Cooperation and data sharing can boost the security of computer networks and mitigate the risk of compromising. However, the research in this field mainly aimed at developing tools for threat information sharing; hence in recent years, there has been a proliferation of threat intelligence platforms [13]. The lack of standards and solid approaches yielded several combinations of solutions and methods incorrectly tagged as threat intelligence.

Tentative guidelines have been proposed in [4], where the authors define information-sharing goals for organizations by also specifying threat information sources and rules for handling the publication and distribution of the data. Nevertheless, there is no consensus among researchers and practitioners on adopting a methodology or technology, as no complete solution exists for handling the standardization, privacy, and reliability issues related to the sharing process.

Although channels such as mail messages, phone calls, ticket systems, or face-to-face meetings have been widely used as a primary way to share threat information quickly, the growing number of cyberattacks made these tools inadequate to handle the volume of data produced, hence the necessity to replace them with semi-automatic tools. Recently, several standards, such as Structured Threat Information CybereXpression (STIX) [14], Cyber Observable eXpression (CybOX) [15], Incident Object Description Exchange Format (IODEF) [16] and Trusted Automated eXchange of Indicator Information (TAXII) [17], have been proposed to facilitate the sharing of IoCs.

In more detail, in [18], the authors describe the main platforms for threat information sharing based on the standards introduced above [12, 5]. One of the most adopted solutions is MISP (Malware Information Sharing Platform), an open-source software solution for collecting, storing, distributing, and sharing cyber security indicators and threat information [19].

MITRE CRITs (Collaborative Research Threats) is another widely used open-source malware and threat repository that leverages different open-source software to create a unified tool for analysts and security experts engaged in threat defense [20]. CIF (Collective Intelligence Framework) is an open-source cyber threat intelligence platform that allows for gathering data from different sources and exploiting them for threat identification, detection, and mitigation. Finally, EclecticIQ Platform is a commercial platform based on STIX and TAXII standards that gathers and interprets intelligence data from open sources.

3. The ORISHA Platform

In this section, we illustrate the main components composing ORISHA. Figure 1 depicts the core actors cooperating within the system: *Distributed TIP*, *TDS Layer*, and *Honeynet*. The TIP is devoted to orchestrate the interactions among the components and represents the core of ORISHA. Basically, it performs two main tasks: *(i)* it allows for storing and encrypting the information (gathered from heterogeneous sources) in a distributed fashion, and *(ii)* it permits to share the collected data to the other components. The distributed TIP is realized by connecting several MISP instances. Among the tools introduced in Section 2, the MISP exhibits different benefits as highlighted in [21]: (*i*) integration with SIEMs and Intrusion Detection Systems

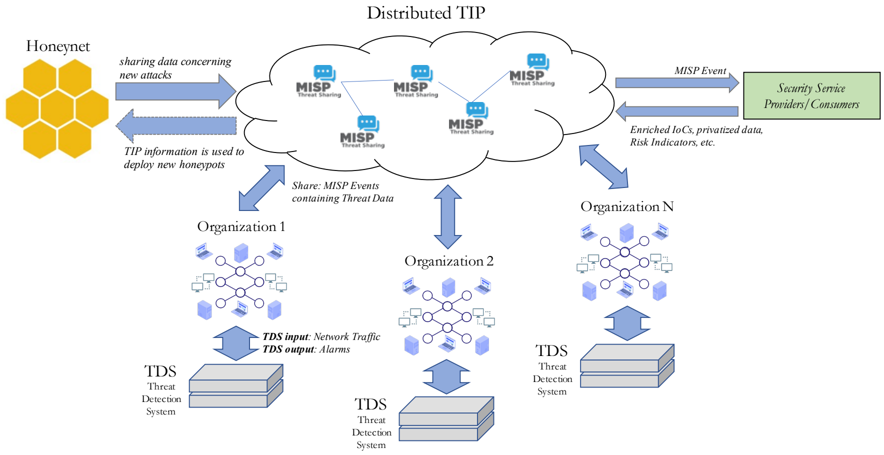
Distributed TIP
Honeynet
sharing data concerning new attacks
TIP information is used to deploy new honeypots
MISP Threat Sharing
MISP Threat Sharing
MISP Threat Sharing
MISP Threat Sharing
MISP Threat Sharing
MISP Event
Security Service Providers/Consumers
Enriched IoCs, privatized data, Risk Indicators, etc.
Share: MISP Events containing Threat Data
Organization 1
Organization 2
Organization N
TDS input: Network Traffic
TDS output: Alarms
TDS Threat Detection System
TDS Threat Detection System
TDS Threat Detection System

Figure 1: ORISHA Platform.

(IDSs) functionalities, (*ii*) extensible and flexible architecture, (*iii*) support for different standards (e.g., STIX, TAXI), (*iv*) detailed documentation, and (*v*) several active communities. Basically, the different MISP instances cooperate by sharing data about upcoming threat events gathered by the other actors.

To this aim, ORISHA mainly leverages two elements: the data exchange format defined in [10] and the layers handling the communication between TDSs and TIP. As an example, in the following, we consider the case in which ORISHA is used to share data about anomalous flow connections discovered by ML-Based IDSs. Specifically, we focused on describing how ORISHA can be used to realize an Active Learning scheme [22]. It is important to note that the platform can be extended to integrate other TDSs by customizing the data exchange format.

3.1. Leveraging ORISHA for Active Learning

In this section, we show how the cooperation among different TDSs is realized using ORISHA and how the decision-making process is improved. Figure 2 depicts the overall information flow. The process begins by monitoring the system. The computer network periodically yields traffic flow that the underlying TDS layer will analyze. In this scenario, a specific anomaly detector (TDS_1 in the figure) processes the `.pcap` files containing the traffic traces and detects an anomaly. A MISP security event is generated and shared with the TIP, which plays the role of "security event hub". Then, a different IDS (TDS_2 in the figure) reads the event, analyzes the embedded `.pcap` files, and labels the event with additional information. The updated MISP object, now with two consensus labels, is examined by an expert who can accept or reject the threat classification. Once validated, the event can be used by other IDSs (e.g., TDS_n in the figure) for the training stage in order to improve their predictive performances.

Now let us consider a different case where the event produced by TDS_1 is classified differently (non-anomalous) by TDS_2. Again, the domain expert examines the event with dissimilar scores and realizes that the event is a false alarm. The event is then returned to TDS_1, which can include the validated event in its training set and refine the underlying model for better accuracy

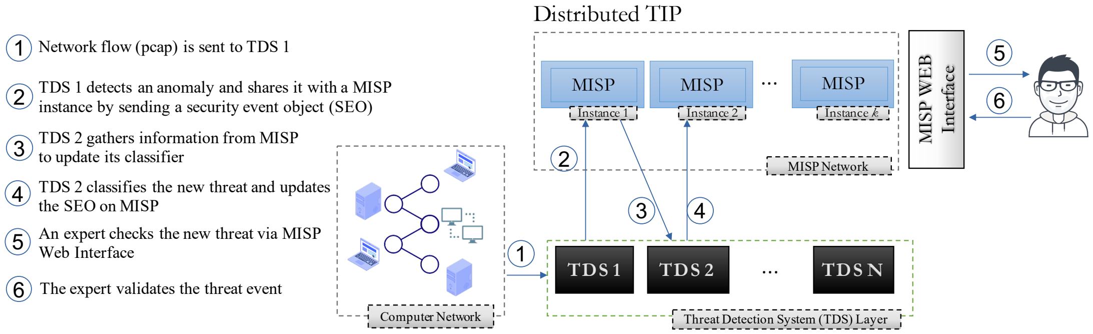
1 Network flow (pcap) is sent to TDS 1
2 TDS 1 detects an anomaly and shares it with a MISP instance by sending a security event object (SEO)
3 TDS 2 gathers information from MISP to update its classifier
4 TDS 2 classifies the new threat and updates the SEO on MISP
5 An expert checks the new threat via MISP Web Interface
6 The expert validates the threat event
Distributed TIP
MISP
Instance 1
MISP
Instance 2
...
MISP
Instance k
MISP Network
MISP WEB Interface
5
6
2
3
4
1
Computer Network
TDS 1
TDS 2
...
TDS N
Threat Detection System (TDS) Layer

Figure 2: An example of execution flow.

and improve its false positive rate.

The solution described above corresponds to the well-known Query-By-Committee strategy [23], with the difference that, here, we foresee that the expert validates both the agreement (in the first situation) and the disagreement (in the second one). More sophisticated validation criteria can be adopted to implement different optimization objectives. For example, to reduce human intervention, automatic validation can be used to confirm the agreements based on confidence values and reduce human analysis to the most uncertain cases based, e.g., on label entropy.

4. Extending ORISHA for Emerging Threats

In this section, we present two classes of emerging threats, i.e., multi-vector attacks leveraging information hiding and fake news. We then discuss how to extend the ORISHA platform with IoCs able to capture the challenging and fast-paced modern security scenario.

4.1. Multi-Vector and Information Hiding Attack Campaigns

In recent years, threat actors are increasingly taking advantage of complex attack chains, especially to elude detection or target large-scale organizations and critical infrastructures. For instance, many modern malware deploy multi-stage loading architectures, e.g., offensive routines are retrieved only when needed to reduce the footprint of the malicious software [24]. Moreover, advanced persistent threats are now able to exploit different portions of the attack surface, making them intrinsically multi-vector (see, e.g., [25], for the case of smart manufacturing systems). Notable recent examples of sophisticated offensive campaigns are the ransomware attack against the Italian vaccination booking system in August 2021 and against the US Colonial Pipeline facility in May 2021. In both cases, threat actors used social-engineering-like techniques, e.g., malicious mail attachments dropped in the home computer of a remote worker, jointly with multi-vector approaches, e.g., flawed password policies or known exploits in commercial software suites. Besides societal and economic losses, such attack campaigns highlighted several limits of CTI and TIP frameworks. First, the sharing of data needs to overcome resistance to disclosing information that can reveal insights on how the security of a complex organization is enforced. Second, collaborative analysis demands a precise methodology for collecting data about a security incident and making it compliant with several (often incompatible) regulations. Third, critical infrastructures are often characterized by sensitive information, which could be actively exploited by state-level threat actors to infer details, such as work shifts or energy requirements.

Unfortunately, the surge of offensive techniques leveraging information hiding is expected to challenge the process of creating IoCs to be shared across various organizations. For the sake of clarity, we present two recent use cases observed in real attacks [26].

Steganographic Malware and Covert Channels

To prevent detection, many recent threats deploy information-hiding mechanisms. For instance, malicious payloads are hidden in digital images by means of steganography. The most used technique encodes bits of secret data by altering the least significant bits of the red, green, and blue color components of pixels belonging to the target image. Altered files can then be sent via mail attachments, embedded within Word/PDF documents, or bundled within an application. Despite the chosen vector, the typical use of steganography is to conceal additional information, such as remote URLs, configuration files, or IP addresses, without leading to a visible signature. From the perspective of developing IoCs or supporting collaborative threat analysis, steganographic malware represents a challenging scenario. In fact, images can be sent or embedded in different manners, thus requiring specialized procedures for the creation of the IoC. Moreover, the steganographic process could introduce overheads in the TIP. As an example, a URL used to retrieve a remote payload hidden within an image could not be a complete IoC. Specifically, also the "carrier" concealing the secret data and the used steganographic mechanism (e.g., the Invoke-PSImage techniques observed in Ursnif) should be part of the IoC itself. In other words, steganographic malware may "inflate" the IoC space to explicitly consider both the malicious hidden content and the container.

Another challenge deals with the abused carriers, which could be very mixed or hard to collect (in principle, any digital content could be used to conceal information). In more detail, attackers could hide malicious information by manipulating icons or images bundled with applications [27] as well as in concurrent code or HTML files [28]. The creation of IoCs should then consider a multitude of heterogeneous assets (e.g., HTML pages, icons, and additional files), which could not be retrieved in a simple manner. Specifically, the original IoC could require to interact with an application store/repository or to crawl/scrape contents through the Web.

Advancements in IDSs, firewalls, and traffic analyzers partially ignited the diffusion among threat actors of covert channels hidden within network traffic. In essence, network covert channels are parasitic communications cloaked within legitimate traffic flows [29], which are created with the ultimate goal of bypassing security tools or blockages. As an example, the attacker could hide sensitive data in unused protocol fields or botnet commands in HTTP headers. Similarly to the case of steganographic malware, network covert channels require the preparation of multiple IoCs. Even if one may consider to share `.pcap` traces containing covert communications, this could lead to several hazards. First, recognizing in which part of the protocol (or flow) the data has been hidden may require to store a non-negligible volume of traffic. Second, an attacker targeting the payload could be identified only via complete traffic traces, which usually conflicts with standard anonymization procedures. Third, covert communications are usually long-lasting, thus collecting data for preparing the IoC could need to gather a huge amount of information at a wire speed, thus lacking of proper scalability [30].

Lastly, a possible realistic example considering the mitigation of the aforementioned threats by using the ORISHA platform could be as follows. A TDS (TDS_1) detects the presence of an image containing malicious content concealed via steganography, e.g., it deploys well-known heuristics or an AI-based countermeasure. For instance, a Web server has been instrumented to spot the presence of skimmers or additional payloads hidden in favicons [31]. The tampered favicon is then "quarantined" and the hidden content is retrieved if possible, e.g., the script or URL is stored in a textual form. A suitable IoC composed of the original favicon, the script, and companion metadata is then prepared and sent via the MISP interface. The IoC could also be enriched with information such as the name of the threat (e.g., Magecart/Magento), the size of the favicon, and the type of the cloaked data (e.g., JavaScript or PowerShell). If needed, additional details on the obfuscation technique used by the attacker (e.g., zipx or Base64) can be put in metadata as well. In a similar manner, a detector (TDS_2) in charge of revealing the presence of network covert channels could bundle fragments of traffic in one or more `.pcap` files, along with information on which part of the protocol has been exploited (e.g., the TTL or the Flow Label). The IoC could also contain data on the detection accuracy to avoid propagating false positives/negatives or help in setting up suitable labels to train an AI-based framework.

4.2. Fake News

Recent years have also seen an increased concern for the threats that fake news and online misinformation present to the democratic debate. Online Web sources and social media are the main means of news information dissemination and spreading. In particular, an exponential increase in the use of social media has accelerated information diffusion. The speed at which misinformation spreads, alongside social media's open access content production and dissemination, increases the potential damage, making online platforms primary targets for fake news propagation.

Therefore, it is necessary to mitigate the impact of misinformation as well as develop specific tools and services to allow citizens and the professional community to access reliable and trustworthy information on the Web and social media. The automatic detection of fake news is a relevant problem attracting great interest from the research community. Most previous research studied the problem of fake news detection, by typically using feature extraction from news content. Text content-based approaches mainly explore lexical and syntactic features like word usage and linguistic styles to identify fake news or to detect the differences in the writing style of real and fake news, such as deception [32, 33, 34, 35, 36, 37, 38]. Other methods, such as the one reported in [39], capture and exploit sensational emotions for learning emotion-enhanced representations. Moreover, some works analyze the images in the news along with the text content for fake news detection [36, 40]. Exploiting user-based features as auxiliary information for improving the identification of fake news was explored in [32, 41].

With the advent of social media, the nature of misinformation has evolved from text-to-visual based modalities, such as images, audio, and video. Therefore, the identification of media-rich fake news requires an approach that exploits and effectively combines the information acquired from different multi-modal data.

Multi-modality is a key approach to improve fake news detection, but successful solutions supporting different data modalities, with their different structure and dimension, is still poorly explored. Multi-Modal Deep Learning based approaches demonstrated to be effective in providing accurate predictions but require feeding with different types of labeled data. In this respect, the integration with ORISHA could represent an effective solution to obtain sufficient data for their learning. In particular, the multi-modality can be implemented through the cooperation of the different TDS within the ORISHA architecture, which can exhibit peculiar classification abilities according to the specific data modality. We can envisage a deep learning based cooperative model that uses the feedbacks of the different organizations within the ORISHA frameworks to estimate news trust levels and ranks the news accordingly.

Let us consider the following scenario. A fake news detector (TDS_1) identifies a fake news by analyzing the textual content of a social media post. At this point, the TDS creates a proper IoC that specifies the news text, the url of an image posted together with the textual content, and a set of metadata reporting the source of the news and its social context (e.g., engaged users, retweets, replies). Further metadata include information such as the probability with which the news has been detected as fake (e.g., the accuracy of the fake news classifier) and that the news has not been validated yet as fake. A MISP security event is then produced and delivered. The event is distributed in the TIP, where a different TDS (TDS_2) analyzes the IoC and classifies the news as real, differently than TDS_1. The updated MISP object, now with two opposite labels, is delivered in the TIP. At this point a third TDS (TDS_3) handles the IoC and analyzes the image (e.g., exploiting a CNN network), classifying it as malicious. Again the MISP security event is updated and delivered. The domain expert inspects the event with dissimilar scores and realizes that the event represents a fake news. The event is then returned to TDS_2, which will adapt its training set with the validated news, improving the classification accuracy of the model by adapting its false negative rate.

5. Conclusions and Future Works

In this paper, we provided an overview of the ORISHA platform, which allows for sharing different pieces of forensics information as well as specific IoCs. As shown, our approach can be used to both improve the accuracy of the detection or foster cooperative threat mitigation campaigns among different organizations. However, the recent surge of advanced attack schemes using information hiding and the diffusion of fake news requires extending the platform and addressing specific challenges. For instance, the heterogeneity of IoCs, privacy constraints, and scalability properties should be considered to effectively deploy ORISHA in realistic deployments. Moreover, we want also to investigate new emerging types of threats aiming at compromising ML models through specific attacks against the learning or deployment stages [42].

Acknowledgments

This work was partially supported by project SERICS (PE00000014) under the NRRP MUR program funded by the EU - NGEU.

References

[1] ENISA, Enisa threat landscape 2020 - list of top 15 threats, 2020. https://www.enisa.europa.eu/publications/enisa-threat-landscape-2020-list-of-top-15-threats.

[2] Interpol, Covid-19 cybercrime analysis report., 2020. https://tinyurl.com/6wek2rk.

[3] Microsoft 365 Defender Threat Intelligence Team, Exploiting a crisis: How cybercriminals behaved during the outbreak, 2020. https://tinyurl.com/cybercrime-during-outbreak.

[4] C. S. Johnson, M. L. Badger, D. Waltermire, J. Snyder, C. Skorupka, Guide to cyber threat information sharing, NIST Special Publication 800-150 (2016).

[5] S. Brown, J. Gommers, O. Serrano, From cyber security information sharing to threat management, in: Proceedings of the 2nd ACM Workshop on Information Sharing and Collaborative Security, 2015, pp. 43–49. doi:10.1145/2808128.2808133.

[6] L. Dandurand, O. S. Serrano, Towards improved cyber security information sharing, in: 2013 5th International Conference on Cyber Conflict (CYCON 2013), 2013, pp. 1–16.

[7] A. Zibak, A. Simpson, Cyber threat information sharing: Perceived benefits and barriers, in: Proceedings of the 14th International Conference on Availability, Reliability and Security, ARES '19, Association for Computing Machinery, 2019, pp. 1–9. doi:10.1145/3339252.3340528.

[8] S. Qamar, Z. Anwar, M. A. Rahman, E. Al-Shaer, B.-T. Chu, Data-driven analytics for cyber-threat intelligence and information sharing, Computers & Security 67 (2017) 35–58. doi:https://doi.org/10.1016/j.cose.2017.02.005.

[9] T. D. Wagner, K. Mahbub, E. Palomar, A. E. Abdallah, Cyber threat intelligence sharing: Survey and research directions, Computers & Security 87 (2019) 101589. doi:https://doi.org/10.1016/j.cose.2019.101589.

[10] M. Guarascio, N. Cassavia, F. S. Pisani, G. Manco, Boosting cyber-threat intelligence via collaborative intrusion detection, Future Generation Computer Systems 135 (2022) 30–43. URL: https://www.sciencedirect.com/science/article/pii/S0167739X22001571. doi:https://doi.org/10.1016/j.future.2022.04.028.

[11] V. Mavroeidis, S. Bromander, Cyber threat intelligence model: An evaluation of taxonomies, sharing standards, and ontologies within cyber threat intelligence, in: 2017 European Intelligence and Security Informatics Conference (EISIC), 2017, pp. 91–98. doi:10.1109/EISIC.2017.20.

[12] C. Sauerwein, C. Sillaber, A. Mussmann, R. Breu, Threat intelligence sharing platforms: An exploratory study of software vendors and research perspectives, Wirtschaftsinformatik und Angewandte Informatik (2017).

[13] M. C. Libicki, Sharing information about threats is not a cybersecurity panacea, Santa Monica, CA: RAND Corporation, 2015, pp. 1–9.

[14] B. Jordan, R. Piazza, T. Darley, Stix™ version 2.1 committee specification 01 (2020).

[15] T. Darley, I. Kirillov, R. Piazza, D. Beck, Cybox™ version 2.1.1. part 01: Overview - committee specification draft 01 / public review draft 01 (2016).

[16] R. Danyliw, J. Meijer, Y. Demchenko, The incident object description exchange format, in: RFC 5070 (Proposed Standard), 2007.

[17] T. Darley, I. Kirillov, R. Piazza, D. Beck, Taxii™ version 2.1 committee specification 01 (2020).

[18] ENISA, Exploring the opportunities and limitations of current threat intelligence platforms, December 2017.

[19] C. Wagner, A. Dulaunoy, G. Wagener, A. Iklody, Misp: The design and implementation of a collaborative threat intelligence sharing platform, in: Proceedings of the 2016 ACM on Workshop on Information Sharing and Collaborative Security, WISCS '16, Association for Computing Machinery, 2016, p. 49–56. doi:10.1145/2994539.2994542.

[20] M. Goffin, Crits: Collaborative research into threats, https://crits.github.io/, 2014. [Online].

[21] G. González-Granadillo, M. Faiella, I. Medeiros, R. Azevedo, S. González-Zarzosa, Etip: An enriched threat intelligence platform for improving osint correlation, analysis, visualization and sharing capabilities, Journal of Information Security and Applications 58 (2021) 102715. doi:https://doi.org/10.1016/j.jisa.2020.102715.

[22] P. Ren, Y. Xiao, X. Chang, P. Huang, Z. Li, B. Gupta, X. Chen, X. Wang, A survey of deep active learning, ACM Comput. Surv. 54 (2021). doi:10.1145/3472291.

[23] D. A. Cohn, L. E. Atlas, R. E. Ladner, Improving generalization with active learning, Machine Learning 15 (1994) 201–221. doi:10.1007/BF00993277.

[24] A. Afianian, S. Niksefat, B. Sadeghiyan, D. Baptiste, Malware dynamic analysis evasion techniques: A survey, ACM Computing Surveys (CSUR) 52 (2019) 1–28.

[25] F. Zahid, G. Funchal, V. Melo, M. M. Kuo, P. Leitao, R. Sinha, Ddos attacks on smart manufacturing systems: A cross-domain taxonomy and attack vectors, in: 2022 IEEE 20th International Conference on Industrial Informatics (INDIN), IEEE, 2022, pp. 214–219.

[26] L. Caviglione, W. Mazurczyk, Never mind the malware, here's the stegomalware, IEEE Security & Privacy 20 (2022) 101–106.

[27] N. Cassavia, L. Caviglione, M. Guarascio, G. Manco, M. Zuppelli, Detection of steganographic threats targeting digital images in heterogeneous ecosystems through machine learning, Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications 13 (2022) 50–67.

[28] Y. Liu, Z. Xu, M. Fan, Y. Hao, K. Chen, H. Chen, Y. Cai, Z. Yang, T. Liu, Concspectre: Be aware of forthcoming malware hidden in concurrent programs, IEEE Transactions on Reliability 71 (2022) 1174–1188.

[29] S. Zander, G. Armitage, P. Branch, A survey of covert channels and countermeasures in computer network protocols, IEEE Communications Surveys & Tutorials 9 (2007) 44–57.

[30] W. Mazurczyk, K. Powójski, L. Caviglione, IPv6 covert channels in the wild, in: Proceedings of the third central european cybersecurity conference, 2019, pp. 1–6.

[31] M. Guarascio, M. Zuppelli, N. Cassavia, L. Caviglione, G. Manco, Revealing MageCart-like threats in favicons via artificial intelligence, in: Proceedings of the 17th International Conference on Availability, Reliability and Security, 2022, pp. 1–7.

[32] K. Shu, L. Cui, S. Wang, D. Lee, H. Liu, Defend: Explainable fake news detection, in: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '19, 2019, p. 395–405.

[33] C. Raj, P. Meel, Arcnn framework for multimodal infodemic detection, Neural Networks 146 (2022) 36–68.

[34] T. Sachan, N. Pinnaparaju, M. Gupta, V. Varma, Scate: Shared cross attention transformer encoders for multimodal fake news detection, in: Proceedings of the 2021 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM

’21, 2021, p. 399–406.

[35] R. Kumari, A. Ekbal, Amfb: Attention based multimodal factorized bilinear pooling for multimodal fake news detection, Expert Systems with Applications 184 (2021) 115412.

[36] Z. Jin, J. Cao, H. Guo, Y. Zhang, J. Luo, Multimodal fusion with recurrent neural networks for rumor detection on microblogs, in: Proceedings of the 25th ACM International Conference on Multimedia, Association for Computing Machinery, New York, NY, USA, 2017, p. MM ’17.

[37] Q. Jing, D. Yao, X. Fan, B. Wang, H. Tan, X. Bu, J. Bi, Transfake: Multi-task transformer for multimodal enhanced fake news detection, in: IJCNN, 2021, pp. 1–8.

[38] J. Wang, H. Mao, H. Li, Fmfn: Fine-grained multimodal fusion networks for fake news detection, Applied Sciences 12 (2022).

[39] X. Zhang, J. Cao, X. Li, Q. Sheng, L. Zhong, K. Shu, Mining dual emotion for fake news detection, in: Proceedings of the Web Conference 2021, WWW ’21, Association for Computing Machinery, 2021, p. 3465–3476.

[40] Y. Wang, F. Ma, Z. Jin, Y. Yuan, G. Xun, K. Jha, L. Su, J. Gao, Eann: Event adversarial neural networks for multi-modal fake news detection, in: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD ’18, Association for Computing Machinery, 2018, p. 849–857. URL: https://doi.org/10.1145/3219819.3219903. doi:10.1145/3219819.3219903.

[41] K. Shu, D. Mahudeswaran, S. Wang, D. Lee, H. Liu, Fakenewsnet: A data repository with news content, social context and dynamic information for studying fake news on social media, arXiv preprint arXiv:1809.01286 (2018).

[42] L. Caviglione, C. Comito, M. Guarascio, G. Manco, Emerging challenges and perspectives in deep learning model security: A brief survey, Systems and Soft Computing 5 (2023) 200050. doi:https://doi.org/10.1016/j.sasc.2023.200050.