2023 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech) | 979-8-3503-0460-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/DASC/PICOM/CBDCOM/CY59711.2023.10361416

Analyzing Edge IoT Digital Forensics Tools: Cyber Attacks Reconstruction and Anti-Forensics Enhancements

Elena Becker*, Maanak Gupta†, and ‡Feras M. Awaysheh
*†Department of Computer Science, Tennessee Tech University, Cookeville, TN, USA
‡Institute of Computer Science, University of Tartu, Tartu, Estonia
*ebecker42@tntech.edu, †mgupta@tntech.edu, ‡feras.awaysheh@ut.ee

***Abstract*—Digital forensics is a rapidly growing sub-domain of forensic science, primarily due to technological advancements and their integration into everyday life. In this paper, our study delves deep into the potential of digital forensics tools, particularly emphasizing their capability to reconstruct the timelines of cyber attacks on edge Internet of Things (IoT) systems. Our research compares ten distinct and widely used digital forensics tools to discern the most detailed, efficient, and user-friendly instrument for IoT attack reconstruction. Our analysis over fourteen different cyber attacks found that Splunk Enterprise is the most optimal static PCAP and CSV file analysis solution, while Microsoft Azure stands out for live analysis in the context of attack reconstruction. Furthermore, we outline potential enhancements for these tools with consideration to the evolving dynamics of anti-forensics.**

***Index Terms*—Digital Forensics, Forensics Tools, Attack Reconstruction, Anti-Forensics, Edge Computing, IoT**

I. Introduction

Edge computing and the Internet of Things (IoT) are rapidly transforming our world, with interconnected devices becoming increasingly ubiquitous in our homes, workplaces, and communities. This transformation has created new challenges for digital forensics investigators, who must now contend with the unique characteristics of IoT devices and networks [32, 35, 36]. One of the biggest challenges in IoT digital forensics is the diversity of devices. IoT devices come in various shapes and sizes, with varying operating systems, hardware configurations, and communication protocols. This diversity makes developing and using forensic tools compatible with all IoT devices challenging. Another challenge is the need for more procedural standardization in the IoT and edge computing industry [33, 37]. There is currently a lack of standard procedures for collecting, preserving, and analyzing digital evidence from IoT devices. This can make it difficult for investigators to ensure their findings are admissible in court for trials. IoT devices often maintain poor security and privacy protections. This exposure can make them vulnerable to cyber attacks and make it difficult for investigators to collect and preserve digital evidence without compromising users' privacy.

Traditional digital forensics methods face added challenges as processing capabilities shift closer to data sources in edge devices. Edge digital forensics extends the investigation to these decentralized edge devices, ensuring that potential evidence distributed across networks, especially in edge nodes, is not overlooked. The security of these edge devices is paramount, as vulnerabilities in edge nodes can compromise vast networks, emphasizing the need for robust security measures and subsequent forensic tools tailored to the edge environment.

Digital forensics tools come into play to aid investigators in navigating these challenges. These applications offer capabilities for efficient, secure, and comprehensive analysis and reporting of digital evidence [16]. They also assist with attack reconstruction, which is finding the "What", "Where", "When", "How", and sometimes "Why" of a cyber attack. Attack reconstruction can be made easier in digital forensics tools when they include certain features, such as the date and time, a list of running processes, hash filtering, and ability to integrate machine learning (ML) models. However, as digital forensic methodologies evolve, so do tactics to hinder them. *Anti-forensics* aims to obstruct investigations by tampering with or reducing the quality of digital evidence. Investigators must anticipate and recognize such tactics, ensuring robust investigations in the face of deliberate obfuscation [30].

This study aids in tackling this gap by comprehensively investigating the capabilities of ten distinct and widely used digital forensics tools to discern the most detailed, efficient, and user-friendly instrument for edge IoT attack reconstruction. Our study discusses the current literature's shortcomings in addressing this research gap. Next, we compare and contrast these ten digital forensics tools, and provide a taxonomy that classifies the current state-of-the-art solutions based on the tool's features and capabilities in providing the required digital forensics artifacts. Finally, we discuss the possible improvements for each tool to strengthen its ability to reconstruct the timelines of IoT attacks using Anti-forensics.

The key contributions of this paper are as follows:

1) We demonstrate that digital forensics tools can be utilized to reconstruct the timelines of fourteen different types of edge IoT cyber attacks.
2) We compare ten different digital forensics tools to determine the most detailed, efficient, and easy-to-use tool for cyber attacks reconstruction.
3) We discuss the strengths and weaknesses of each tool, with best practice recommendations.

979-8-3503-0460-2/23/$31.00 ©2023 IEEE

4) We conjecture that Splunk Enterprise is the best choice for static PCAP and CSV file analysis, while Microsoft Azure is the top tool for live analysis, for attack reconstruction.
5) We propose improvements to the digital forensics process and tools examined with counters to anti-forensics techniques.

The rest of the paper is organized as follows: Section II covers an overview of related work and highlights the current limitations of existing approaches. Section III discusses the problem definition and solution approach. Section IV presents the results, discussion, and proposed improvements. Lastly, Section V covers the conclusion, a summary of findings, and future work.

II. Related Work

Authors in [7] proposed a methodology for the digital forensics technique of attack reconstruction and showcased several existing techniques for reconstruction. This work noted that none of the identified pre-existing methods are able to adequately handle the complexity that is inherent when trying to reconstruct a timeline of events and artifacts from the massive amount of data that can be extracted from a disk image. In their proposed methodology, the authors addressed this issue by focusing on an abstracted approach, which helps provide forensics investigators with key information in a compact form. The timeline is abstracted out into four different levels, two that focus on the events at different levels of granularity and two that focus on the artifacts in the same way. The two levels of event abstractions provide information on what kind of activity was performed, while the two levels of artifact abstraction provide more complete information about a particular activity. The authors also conducted a case study in which they show how their methodology, once implemented, assists with the task of reconstruction. For this, two types of experiments were conducted, with one group focusing on online operations and the other group focusing on offline activities. Both groups of experiments included running various applications both sequentially and concurrently. The results of this case study showed that the authors' proposed methodology was able to reconstruct a timeline of events and artifacts once it was implemented.

Work in [8] provides an introduction to digital forensics, identifies various evaluation metrics that can be used to evaluate digital forensics tools, and provides a comparison of digital forensics tools that are used for various areas of digital forensics. The evaluation metrics identified for digital forensics tools are: absolute speed, completeness, relative speed, reliability, preciseness, auditability, and repeatability. In addition to these metrics, the authors also identify various challenges that are faced by digital forensics tools. For the comparison, the various tools are divided up into groups based on the type of digital forensics they are used for. The authors provide a taxonomy that consists of these various digital forensics groups. Included in the taxonomy are the following groups: computer forensics, cloud forensics, database forensics, email forensics, IoT forensics, memory forensics, mobile forensics, and network forensics.

Authors in [10], compare and contrast various digital forensics tools. Specifically, they look at tools relating to the areas of desktop forensics - where evidence is extracted from the secondary memory, live forensics - where evidence is extracted from the primary memory, and live network forensics - where evidence is extracted from packets traveling on a live network. The authors identify two problems that currently exist in the digital forensic tool space from their literature review: one, it can be hard to select which tool is the best to use, and two, several forensics tools require massive amounts of time to run before they provide any results. For each tool being compared, a short description of its purpose and capabilities was provided. The parameters used for comparison were created by examining the different features supported by the various tools investigated. Results of the comparison were presented in the form of a table in the paper. Since this paper was a comparative study, the results of the comparison were not discussed in any significant detail.

The paper, [11], examines how applicable the ML technique of neural networks is with assisting in digital forensics. More specifically, the task examined is "identifying the exact sequence of actions affecting a file system," [11]. Before the use of neural networks could be investigated, a dataset had to be created first. The dataset collected by the authors has attributes relating to file system activities, as well as system event audit log entries, and a class that is the application that was used. Four different scenarios are considered in the dataset: applications executed sequentially with no file interaction, applications being executed sequentially and loading a file, applications being executed sequentially and performing a variety of file interactions, and lastly, applications being run concurrently while performing various file interactions. The authors describe how they went about cleaning and pre-processing the dataset to ready it for a neural network, and then detail the architecture of the network used. The chosen network was a feed-forward artificial neural network classification model. After the network was trained, it was then used with the dataset created to classify which application was used to produce a set of attributes. The results showed that the neural network produced good results in most cases.

A. Limitations of Literature

Table I, particularly the emphasized row in red, delineates the distinctions in focus and the suite of digital forensics tools adopted in our methodology. Notably, our work stands apart as it encompasses digital forensics and attack reconstruction capabilities — a combination not thoroughly addressed in extant literature. The salient contributions of this paper, which seek to bridge identified gaps, are as follows:

1) Contrasting the focal areas illuminated in [2]- [4], [6] and [11]- [15], our investigation integrates digital forensics and the intricacies of attack reconstruction (TTP).
2) While digital forensics analyses in prior studies such as [7, 8, 10, 11, 13], and [15] are acknowledged, our

	Focus		Tools Used												
Paper	Digital Forensics	Attack Reconstruction	Autopsy & TSK	Volatility	Wireshark	SIFT	CAINE	Xplico	LastActivityView	Tcpdump	Log2timeline	Splunk Enterprise & MLTK	NetworkMiner	Elastic	Microsoft Azure
Sachdeva et al. 2022 [2]	✓														
Shakeele et al. 2021 [3]	✓														
Elhoseny et al. 2021 [4]	✓														
Qadir et al. 2020 [6]	✓														
Bhandari et al. 2020 [7]	✓	✓									✓				
Narwal et al. 2020 [8]	✓	✓	✓	✓	✓		✓	✓							
Soltani et al. 2019 [9]	✓	✓													
Lovanshi et al. 2019 [10]	✓	✓	✓	✓	✓										
Mohammad 2018 [11]	✓		✓			✓			✓						
Babun et al. 2018 [12]	✓														
Koroniotis et al. 2017 [13]	✓									✓					
Tallón-Ballesteros et al. 2014 [14]	✓														
Marturana et al. 2012 [15]	✓								✓						
Our Approach	✓	✓	✓	✓	✓	✓	✓	✓				✓	✓	✓	✓

TABLE I
DIFFERENCES AMONG RELATED WORKS WITH RESPECT TO FOCUS COVERED AND TOOLS USED

methodology entails a comprehensive comparison of ten disparate digital forensics tools, elaborated further in subsequent sections.

3) Our approach brings a nuanced lens to *Attack Reconstruction*, adding more depth to this critical aspect than prior works like [7]- [10].
4) Our work includes potential enhancements to the digital forensics tools examined with anti-forensics in mind.

III. PROPOSED METHODOLOGY

A. Problem Definition

Digital forensics tools are necessary for reconstructing a timeline of cyber attacks, which are often used for security remediation efforts or even as evidence in criminal cases. There are many digital forensics tools to choose from, and it is time-consuming for analysts, institutions, and businesses to choose which one is best. In this paper, we aim to remediate this issue by highlighting the best tools to utilize for both static PCAP or CSV file and live attack analysis to reconstruct cyber attacks.

B. Our Approach

We analyzed ten different and widely used digital forensics tools to determine the most detailed, efficient, and easy-to-use one for cyber attack reconstruction. Figure 1 displays an overview of the IoT digital forensics process as data feeds through a digital forensics tool and outputs many different features to be used for attack reconstruction. The more features a tool supports, the easier it is for a more thorough attack reconstruction.

The edge IoT attack dataset we utilized for attack reconstruction tool comparison consisted of fourteen different attacks, as shown in Figure 2 [1]. For comparing, we downloaded each tool and fed in the cleaned edge IoT attack data used in our prior work titled, "Using Machine Learning for Detection and Classification of Cyber Attacks in Edge IoT" [1, 29]. The preprocessed data was converted into either a PCAP or CSV file format, depending on what configuration the tool accepted. We then evaluated the tools with their different functions to see which one had the most features - which we found to be the most important for attack reconstruction - as listed in Table II. We then reconstructed at least ten different attacks on each by taking a random sample of the attacks from the dataset in Figure 2.

C. Digital Forensics Tools Used

We used the following tools for our analysis.

1) *Autopsy and The Sleuth Kit (TSK):* A digital forensics platform and graphical user interface (GUI) for several digital forensics tools, including TSK, which is an open-source command line tool that supports the forensic inspection of file systems and disk volumes [17].

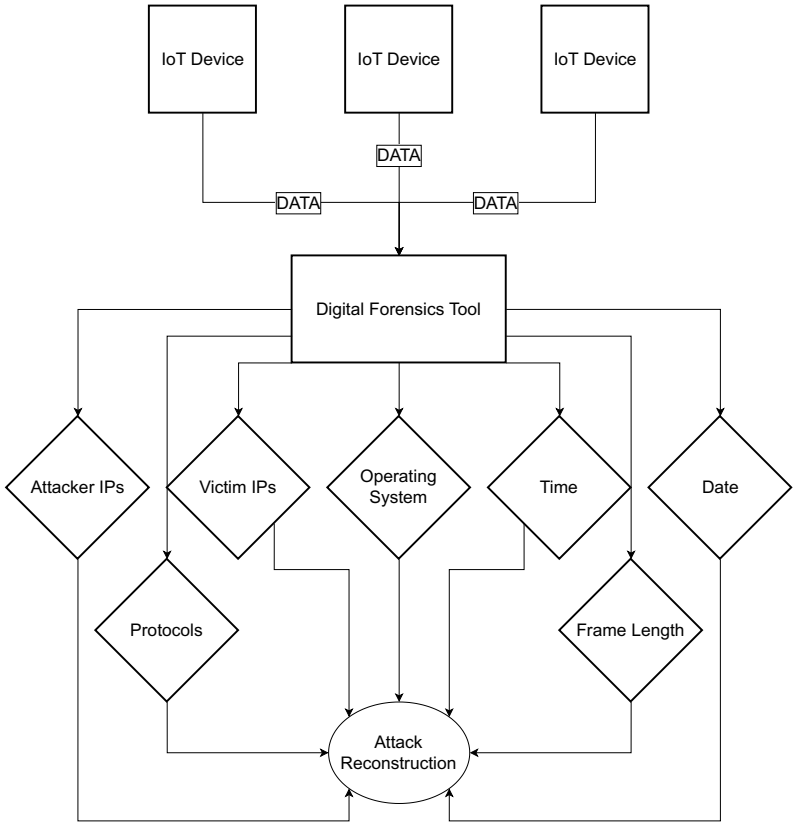

Fig. 1. IoT Digital Forensics Process and Attack Reconstruction

```
Normal                      24101
DDoS_UDP                    14498
DDoS_ICMP                   13096
DDoS_HTTP                   10495
SQL_injection               10282
DDoS_TCP                    10247
Uploading                   10214
Vulnerability_scanner       10062
Password                     9972
Backdoor                     9865
Ransomware                   9689
XSS                          9543
Port_Scanning                8921
Fingerprinting                853
MITM                          358
Name: Attack_type, dtype: int64
```

Fig. 2. Preprocessed Class Numbers

2) *Volatility:* A command line tool that lets digital forensics and incident response (DFIR) analysts acquire and analyze the volatile data that is temporarily stored in random access memory (RAM) [25].
3) *Wireshark:* An open-source packet analyzer tool for network analysis, troubleshooting, and protocol development [18].
4) *Sans Investigative Forensic Toolkit (SIFT) Workstation:* A collection of open-source digital forensics tools for performing examinations in many settings [26].
5) *Computer Aided Investigative Environment (CAINE):* An investigative environment and GUI with built-in digital forensics tools [21].
6) *Xplico:* An open-source Network Forensic Analysis Tool (NFAT) used for capturing and analyzing network traffic [23].
7) *Splunk Enterprise and Splunk Machine Learning Toolkit (MLTK):* A data platform for capturing, managing, and analyzing data from any source. It also has the MLTK plug-in, which allows for integrating ML models [18].
8) *NetworkMiner:* An open-source network forensics tool for examining data captured in PCAP files. It can also be used to capture live network traffic [24].
9) *Elastic, Elasticsearch, Kibana, and Integrations:* A platform designed for capturing, analyzing, and viewing data with additional plug-ins for more features, like Elasticsearch, Kibana, and Integrations [27].
10) *Microsoft Azure, Sentinel, and 365 Defender:* A cloud platform with products and cloud services for collecting data, as well as building, managing, and running applications on-premises, on multiple clouds, and at the edge. Sentinel and 365 Defender are products used for

analyzing data and detecting and mitigating attacks [28].

IV. Analysis and Enhancements

A. Results

Our analysis found that Splunk Enterprise is the best digital forensics tool for static PCAP and CSV file analysis and attack reconstruction. Furthermore, our analysis conjectures that Microsoft Azure, with the addition of Sentinel and 365 Defender, is the best for live attack analysis and attack reconstruction, though this type of analysis was not done in this work and is left for future work.

B. Discussion

Table II and Figure 3 show the comparison of the ten tools for all 33 features they potentially possess. Our analysis suggests that the more features the tools support, the better they are for attack reconstruction.

Autopsy has a GUI similar to Windows Device Manager, so it is simple and organized, but also detailed enough for thorough investigations. TSK allows for the forensic analysis of files, hash filtering, analysis of e-mail and web artifacts, and keyword search [17]. Autopsy and TSK are more structured towards disk reconstruction, however, so they were not able to do as in-depth attack reconstruction as the other tools. Autopsy and TSK could be improved by including the ability to examine PCAP files without the addition of Wireshark, as it could be useful to report evidence for cases all on one platform. Autopsy and TSK also need to take in account their security, as attackers can craft attacks to specifically evade Autopsy's methods of investigation through anti-forensics.

Volatility is very detailed and provides the date and time of captured images, running processes, open network sockets and connections, loaded libraries and open file names for each process, memory addresses, operating system kernel modules, and a mapping of physical offsets to virtual addresses [19]. The downside is that the user must work through the command line, as there is no GUI. A simple GUI could be added to improve the tool as a whole, as it makes attack reconstruction much simpler for analysts.

Wireshark has many strengths, but it also has impactful weaknesses. It has many options for looking at data, it's easier to set up than its highest competitor - Splunk, and it has bandwidth monitoring, Internet Protocol (IP) address monitoring, internet usage monitoring, network analysis, performance metrics, real-time monitoring, reporting and statistics, server monitoring, Service Level Agreement (SLA) management, threshold alerts, uptime reporting, support for the Simple Network Management Protocol (SNMP), and web traffic reporting [18]. The weaknesses are that its GUI could be improved - as it looks outdated, and it can be overwhelmed by the amount of packets that are captured. Users also can't integrate ML models in Wireshark itself, so they need to use a ML model to label attacks and then open the marked file in Wireshark to analyze it.

SIFT Workstation generates detailed timelines for system logs, which are especially useful for attack reconstruction, and it allows for thorough examination of data files. Its weaknesses are that it has poor usability, as well as a substandard GUI and user documentation [26].

CAINE's strengths are that it provides a complete investigative environment for all stages of the digital forensics investigation process - including data preservation, collection, examination, and analysis, a user-friendly GUI, and a large collection of third-party digital forensics tools [21]. The weakness is that it must be run on a Linux system or virtual environment. CAINE could be improved by adding Windows OS functionality.

Xplico is a very in-depth NFAT with many strengths. It incorporates Port Independent Protocol Identification (PIPI) for each application protocol, multi-threading, outputting data and information in SQLite or MySQL databases and/or files, Transmission Control Protocol (TCP) reassembly with acknowledgment (ACK) or soft ACK verification for any packet, reverse Domain Name System (DNS) lookup from DNS packages contained in the input files that are not from an external DNS server, IPv4 and IPv6 support, modularity for each component, real-time elaboration, and no size limit on data entries or the number of files [23]. Additionally, each data file reassembled by Xplico is associated with a Extensible Markup Language (XML) file that uniquely identifies the flows and the PCAP containing the data reassembled [23]. The downside is that it is not specifically a network protocol analyzer, which could be helpful when reconstructing attacks if it was added.

Splunk Enterprise and Splunk Machine Learning Toolkit (MLTK) have many strengths and only one noticeable weakness. Its strengths include being able to integrate ML models with the MLTK, a good GUI and visualizations, activity tracking, alerts and notifications, a baseline manager, a user-friendly dashboard overall platform, event logs, IP address monitoring, internet usage monitoring, network analysis, network resource management, patch management, performance metrics, policy management, real-time monitoring and notifications, reporting and statistics, server monitoring, SLA management, support for the SNMP, threshold alerts, uptime reporting, and web traffic reporting [18]. Splunk's main weakness is that it is more complex to set up at the beginning for PCAP file analysis than its top competitor, Wireshark. It could be improved by adding the functionality of PCAP analysis by default.

NetworkMiner's strengths are that it can analyze files, images, and passwords from captured network traffic in static PCAP files as well as capture live network traffic by sniffing network interfaces [24]. It also provides detailed information about each IP address in the analyzed network traffic, which is aggregated to a network host inventory and can be used for passive asset discovery, as well as get an overview of which devices that are communicating [24]. Its weakness is that it has a poor GUI that looks like an outdated version of Windows Device Manager. If the GUI was improved, this would be a solid choice for network data analysis.

Elastic is a very well-rounded platform overall for data aggregation, live analysis, and attack reconstruction with the

Features	Tools Compared									
	Autopsy & TSK	Volatility	Wireshark	SIFT	CAINE	Xplico	Splunk Enterprise & MLTK	NetworkMiner	Elastic	Microsoft Azure
File analysis	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Hash filtering	✓		✓	✓	✓		✓	✓	✓	✓
E-mail & web artifact/traffic analysis	✓		✓	✓	✓		✓	✓	✓	✓
Keyword search	✓		✓	✓	✓		✓		✓	✓
Disk reconstruction	✓									
Date & time	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Running processes	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Open network sockets & connections		✓	✓	✓	✓	✓	✓	✓	✓	✓
Loaded libraries & file names	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Memory addresses		✓	✓	✓	✓		✓	✓	✓	✓
OS kernel modules	✓	✓	✓	✓	✓		✓	✓	✓	✓
Mapping of physical offsets to virutal addresses		✓	✓	✓			✓		✓	✓
Graphical user interface	✓		✓	✓	✓		✓	✓	✓	✓
Bandwidth monitoring			✓		✓	✓	✓		✓	✓
IP address monitoring			✓		✓	✓	✓	✓	✓	✓
Internet usage monitoring			✓		✓	✓	✓		✓	✓
Network analysis			✓		✓	✓	✓	✓	✓	✓
Performance metrics			✓				✓		✓	✓
Reporting & statistics	✓		✓	✓	✓		✓		✓	✓
Server monitoring			✓		✓		✓		✓	✓
Service Level Agreement management			✓				✓		✓	✓
Threshold alerts			✓				✓		✓	✓
Uptime reporting			✓				✓		✓	✓
Support for Simple Network Management Protocol			✓				✓		✓	✓
Able to integrate machine learning models							✓		✓	✓
System log timelines			✓	✓	✓		✓		✓	✓
Includes additional third-party digital forensic tools	✓			✓	✓	✓			✓	✓
User-friendly					✓		✓			✓
Visualizations and/or insights	✓		✓				✓		✓	✓
Patch management							✓		✓	✓
Network resource management							✓		✓	✓
Static analysis	✓	✓	✓	✓	✓	✓	✓	✓		
Live analysis	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

TABLE II
TOOLS FEATURES COMPARISON

addition of tools like Elasticsearch, Kibana, and Integrations. Elasticsearch lets the user store, search, and analyze data quickly at any scale, Kibana displays informative data visualizations for enhanced analysis, and Integrations provides valuable insights for process improvement [27]. Elastic also has the ability to directly connect ML models to Elastic to assist in attack detection and classification. Its downsides mostly involve user misconfiguration, as the default settings for Elastic should be changed to better fit each user's specific needs and environment for it to work best. Its GUI could also be improved a little, and in regards to IoT, tags should be added to better identify each device even when they have the same IPs. Additionally, Elastic is popular enough that it is easier for attackers to actively use anti-forensics methods to go against digital forensic ones used by analysts. This can be mitigated by keeping up-to-date with the latest vulnerabilities and updating tools with patches frequently.

Microsoft Azure is the best choice out of the ones compared for live network traffic data analysis and attack reconstruction with the addition of Sentinel, and 365 Defender. Azure has a very detailed and user-friendly GUI, methodical data aggregation and analysis functions, helpful visualizations, and beneficial insights and cybersecurity plug-ins. Microsoft Sentinel provides a comprehensive Security Information and Event Management (SIEM) solution for detecting threats, completing in-depth investigations, responding to incidents, and hunting

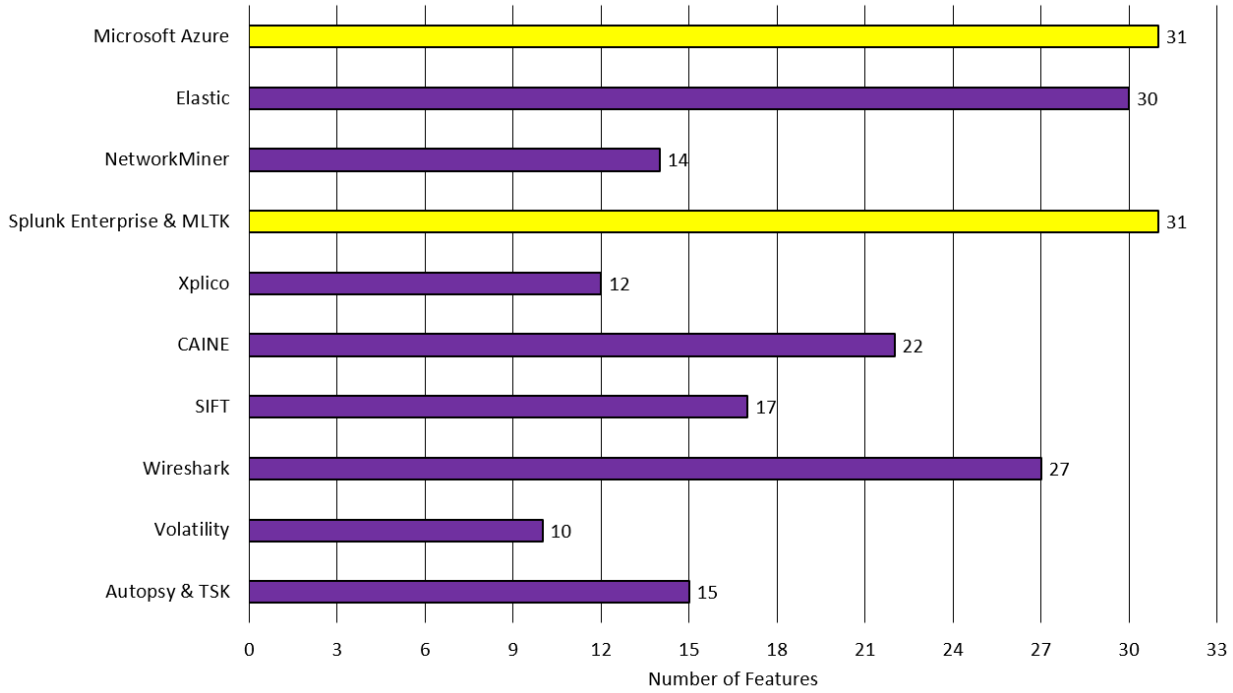

Fig. 3. Tool Features Comparison Chart

proactively for threats [28]. Microsoft 365 Defender offers a very informative dashboard detailing threat analytics, users and devices at risk, device health, active incidents, discovered devices, and more [28]. Azure can also directly integrate ML models for assistance in attack detection and classification. Its downsides are the same as Elastic, as user misconfiguration can result in the suite of products not being used to their full potential. Microsoft is also well-known, so many attackers and cybercrime organizations try to create ways around their security defenses with anti-forensics, so they must constantly update their tools with patches. Similar to Elastic, tags can also be added to IoT devices to assist in identifying those with the same IP addresses.

C. Proposed Anti-Forensics Enhancements

In all of the tools analyzed, anti-forensics can be used at the data collection stage to disrupt the analysis process. Data wiping, a tactic of securely erasing data, is one of the most well-known anti-forensic strategies and can be done directly on the device or remotely [31]. There are multiple ways to combat this method, a lot of which can be done in the digital forensics tools with the capability to collect data themselves. For instance, analysts should always make multiple copies of collected data, and fragments or traces of data can also recovered and used to restore the entirety of the data if need be. The tools that have data collection abilities should make sure to include an automatic data copying and fragment or trace scavenger mechanisms.

Another common anti-forensics approach is encrypting data [31]. Live digital forensics tools can improve by including the means to extract decryption keys from live random access memory (RAM) dumps. Modern Windows versions also automatically use BitLocker encryption on the system volume, which creates escrow keys (Recovery Keys) that can decrypt it [31]. Analysts can find an uploaded Recovery Key on the user's OneDrive account, so digital forensics tools could be enhanced by having the automatic retrieval of such Recovery Key. Both of these methods would assist analysts in decrypting the data collected all in one digital forensics tool.

V. Conclusion and Future Work

In an era where the technological landscape is rapidly evolving and the nexus between edge computing and IoT devices intensifies, the importance of digital forensics cannot be understated. With the rise in cyber crimes, especially those targeting edge nodes and IoT devices, the need for proficient and adaptive digital forensic tools becomes paramount. These tools must ensure comprehensive analysis and anticipate and counteract anti-forensic tactics that sophisticated adversaries employ. This research undertook a meticulous evaluation of ten prominent digital forensics tools, emphasizing their potential in the realm of edge and IoT environments. Our findings illuminated each tool's particular strengths and areas of improvement, especially against the backdrop of anti-forensics. Notably, Splunk Enterprise emerged as the most adept tool for reconstructing attacks based on static PCAP and CSV files, while Microsoft Azure displayed superior capabilities for real-time attack analysis.

In our future work, we will consider deeper exploration into real-time attack analyses, especially in edge environments,

which could enhance our understanding. Controlled attacks under varied scenarios will further validate and expand upon our preliminary findings and are also labeled as future work.

ACKNOWLEDGEMENTS

This work is partially supported by the DoD Cybersecurity Scholarship H98230-22-1-0248 and the NSF grants 2025682 and 2230609.

REFERENCES

[1] M. A. Ferrag, O. Friha, D. Hamouda, L. Maglaras and H. Janicke, "Edge-IIoTset: A New Comprehensive Realistic Cyber Security Dataset of IoT and IIoT Applications for Centralized and Federated Learning," in IEEE Access, vol. 10, pp. 40281-40306, 2022, doi: 10.1109/ACCESS.2022.3165809.

[2] S. Sachdeva and A. Ali, "Machine learning with digital forensics for attack classification in cloud network environment," International Journal of System Assurance Engineering and Management, vol. 13, no. 1, pp. 156–165, 2022.

[3] P. M. Shakeel, S. Baskar, H. Fouad, G. Manogaran, V. Saravanan, and C. E. Montenegro-Marin, "Internet of things forensic data analysis using machine learning to identify roots of data scavenging," Future Generation Computer Systems, vol. 115, pp. 756–768, 2021, doi: https://doi.org/10.1016/j.future.2020.10.001.

[4] M. Elhoseny, M. M. Selim, and K. Shankar, "Optimal deep learning based convolution neural network for digital forensics face sketch synthesis in internet of things (IoT)," International Journal of Machine Learning and Cybernetics, vol. 12, no. 11, pp. 3249–3260, 2021.

[5] I. Selim Gamal Eldin et al, "Anomaly events classification and detection system in critical industrial internet of things infrastructure using machine learning algorithms," Multimedia Tools Appl, vol. 80, (8), pp. 12619-12640, 2021.

[6] A. M. Qadir and A. Varol, "The Role of Machine Learning in Digital Forensics," 2020 8th International Symposium on Digital Forensics and Security (ISDFS), 2020, pp. 1-5, doi: 10.1109/ISDFS49300.2020.9116298.

[7] S. Bhandari and V. Jusas, "An Abstraction Based Approach for Reconstruction of TimeLine in Digital Forensics," Symmetry, vol. 12, no. 1, 2020, doi: 10.3390/sym12010104.

[8] B. Narwal and Nimisha Goel, "A Walkthrough of Digital Forensics and its Tools," 2020. [Online].

[9] S. Soltani and S. A. H. Seno, "A formal model for event reconstruction in digital forensic investigation," Digital Investigation, vol. 30, pp. 148–160, 2019, doi: https://doi.org/10.1016/j.diin.2019.07.006.

[10] Lovanshi, M., Bansal, P. (2019). Comparative Study of Digital Forensic Tools. In: Shukla, R.K., Agrawal, J., Sharma, S., Singh Tomer, G. (eds) Data, Engineering and Applications. Springer, Singapore. https://doi.org/10.1007/978-981-13-6351-1_15

[11] R. M. Mohammad, "A Neural Network based Digital Forensics Classification," 2018 IEEE/ACS 15th International Conference on Computer Systems and Applications (AICCSA), 2018, pp. 1-7, doi: 10.1109/AICCSA.2018.8612868.

[12] L. Babun, A. K. Sikder, A. Acar, and A. S. Uluagac, "Iotdots: A digital forensics framework for smart environments," arXiv preprint arXiv:1809.00745, 2018.

[13] N. Koroniotis, N. Moustafa, E. Sitnikova, and J. Slay, "Towards developing network forensic mechanism for botnet activities in the IoT based on machine learning techniques," in International Conference on Mobile Networks and Management, 2017, pp. 30–44.

[14] A. J. Tallón-Ballesteros and J. C. Riquelme, "Data mining methods applied to a digital forensics task for supervised machine learning," in Computational Intelligence in Digital Forensics: Forensic Investigation and Applications, Springer, 2014, pp. 413–428.

[15] F. Marturana, G. Me and S. Tacconi, "A Case Study on Digital Forensics in the Cloud," 2012 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery, 2012.

[16] "What is Digital Forensics?," EC-Council.org, 2023, https://www.eccouncil.org/cybersecurity/what-is-digital-forensics/

[17] B. Carrier, "Autopsy," SleuthKit.org, 2023, https://www.sleuthkit.org/autopsy/desc.php#:~:text=Autopsy%20is%20a%20graphical%20interface,2%2C%20Ext2%2F3).

[18] "Compare splunk enterprise vs wireshark 2023," Capterra, 2023, https://www.capterra.com/network-monitoring-software/compare/94317-209737/Splunk-vs-Wireshark.

[19] Pavel, "Using the volatility framework for analyzing physical memory dumps," Apriorit, 2023, https://www.apriorit.com/qa-blog/662-cybersecurity-using-volatility-framework-for-analyzing-physical-memory-dumps.

[20] "Free & open source computer forensics tools," Infosec, 2023, https://resources.infosecinstitute.com/topics/digital-forensics/free-open-source-computer-forensics-tools/.

[21] R. Garg, "Caine Forensic environment," GeeksforGeeks, 2023, https://www.geeksforgeeks.org/caine-forensic-environment/.

[22] D. Taylor, "Digital Forensics Review and tutorial project (Caine)," Daniel Taylor's Projects Blog, 2023, https://danoataylor.wordpress.com/digital-forensics-review-and-tutorial-project-caine/.

[23] G. Costa and A. D. Franceschi, "Open Source Network Forensic Analysis Tool (NFAT) ," Xplico, 2023, https://www.xplico.org/about.

[24] "Networkminer - the NSM and network forensics analysis tool," Netresec, 2023, https://www.netresec.com/?page=NetworkMiner.

[25] R. Olatona, "Volatility Is an Essential DFIR Tool-Here's Why," Booz Allen, 2023, https://www.boozallen.com/insights/cyber/tech/volatility-is-an-essential-dfir-tool-here-s-why.html.

[26] R. Lee, "SIFT Workstation," SANS, 2023, https://www.sans.org/tools/sift-workstation/.

[27] "Elastic Stack: Elasticsearch, Kibana, Beats & Logstash," Elastic, 2023, https://www.elastic.co/elastic-stack/.

[28] "Microsoft Sentinel - Cloud SIEM Solution: Microsoft Security," Cloud SIEM Solution — Microsoft Security, 2023, https://www.microsoft.com/en-us/security/business/siem-and-xdr/microsoft-sentinel?rtc=1#x3d7179d29ae5426595f7e3d948ad89bb.

[29] E. Becker, M. Gupta and K. Aryal, "Using Machine Learning for Detection and Classification of Cyber Attacks in Edge IoT," 2023 IEEE International Conference on Edge Computing and Communications (EDGE), Chicago, IL, USA, 2023, pp. 400-410, doi: 10.1109/EDGE60047.2023.00063.

[30] K. Singhan, "Anti Forensics," GeeksforGeeks, 2023, https://www.geeksforgeeks.org/anti-forensics/.

[31] O. Afonin, D. Nikolaev, and Y. Gubanov, "Countering Anti-Forensic Efforts - Part 2," Belkasoft, 2023, https://belkasoft.com/countering-anti-forensic-efforts-part-2.

[32] Awaysheh FM, Aladwan MN, Alazab M, Alawadi S, Cabaleiro JC, Pena TF. Security by design for big data frameworks over cloud computing. IEEE Trans. on Engineering Management. 2021 Feb 8;69(6):3676-93.

[33] Awaysheh, F.M., Alawadi, S. and AlZubi, S., 2022, December. FLIoDT: A Federated Learning Architecture from Privacy by Design to Privacy by Default over IoT. In 2022 Seventh International Conference on Fog and Mobile Edge Computing (FMEC) (pp. 1-6). IEEE.

[34] Awaysheh, F.M., 2022. From the Cloud to the Edge Towards a Distributed and Light Weight Secure Big Data Pipelines for IoT Applications. In Trust, Security and Privacy for Big Data (pp. 50-68). CRC Press.

[35] Gupta, Maanak, and Ravi Sandhu. "Towards activity-centric access control for smart collaborative ecosystems." In Proceedings of the ACM Symposium on Access Control Models and Technologies. 2021.

[36] Cathey, Glen, James Benson, Maanak Gupta, and Ravi Sandhu. "Edge centric secure data sharing with digital twins in smart ecosystems." In IEEE International Conference on Trust, Privacy and Security in Intelligent Systems and Applications (TPS-ISA), pp. 70-79. IEEE, 2021.

[37] Gupta, Maanak, and Ravi Sandhu. "Authorization framework for secure cloud assisted connected cars and vehicular internet of things." In Proceedings of the 23nd ACM on symposium on access control models and technologies, pp. 193-204. 2018.