OSINT & ITS APPLICATION IN DIGITAL FORENSIC

Seinab Mohamed Ibrahim

DECEMBER 13, 2024

Table of Contents

Abstract

This paper explores the evolution and applications of Open-Source Intelligence (OSINT) in digital forensics. Emerging from traditional intelligence, OSINT now analyzes vast amounts of publicly available online data. The paper highlights how AI and machine learning enhance OSINT's efficiency for cybersecurity, criminal profiling, and law enforcement. It discusses key developments, such as OSINT's integration with law enforcement frameworks and tools like Wireshark for network forensics, while addressing challenges like ethical concerns, unreliable data, and fragmented methodologies. The review underscores OSINT's role in addressing digital threats and advancing investigations, emphasizing the need for unified standards and robust ethical guidelines.

Introduction

Open-source intelligence (OSINT) emerged from traditional intelligence-based practices but has grown significantly with advancements in technology and the internet. Indeed, OSINT today is working within the digital domain, studying web pages, social media, and other accessibly online sources.

With the growing of net, new crime like-Hacking, Identity theft and financial fraud, started to happen. OSINT was designed to monitor and investigate these crimes, providing law enforcement and cybersecurity professionals with updated knowledge and leading tools against online threats.

It is widely used in the investigation and evidence collection as well as situational awareness for preventing crime. Organizations and law enforcement can also use this information to help predict and prevent future potential threats by profiling individuals and using historical data to understand how to prevent future risks from developing, improving security in both digital and physical environments.

Critical review

With the use of Open-Source Intelligence (OSINT) the cornerstone of digital forensics, the ambit of assessing, classifying, and ultimately addressing challenges is extended through dissecting publicly accessible data ranging from social media integrations to graphics, and even videos displayed on the web. OSINT provides actionable insights using large amounts of online data and information that can keep pace with the ever-changing digital environment. In this review, we have critically assessed OSINT applications, merits, drawbacks, and impact of current methodologies and available tools.

Yadav et al. (2023) underline the relevance of OSINT in the fight against Cybersecurity threats and in the vicinity of digital forensics. AI and machine learning can also be integrated, allowing the freaking efficient analysis of public data — where it has been especially effective on sentiment analysis, criminal profiling, and predictive threat detection. Nonetheless, the scalability and reproducibility of these approaches are hampered by data reliability issues and the absence of standardized datasets. In this paper, we demonstrate the necessity of a common framework for integration of heterogeneous data modalities.

OSINT for Law Enforcement — Fugitive Tracking through Geolocation & Social Media Analysis Waghmare (2023) By using OSINT tools in conjunction with Interpol's real-time updates, investigators can retrace movements and anticipate where fugitives may be located. Given their promise, there are significant ethical concerns with such methods — including the risk of misuse against marginalized groups. It is imperative to have robust legal protections and ethical systems in place to mitigate such risks.

For example, Soepeno (2023) shows the use of OSINT in network forensics ethics using a tool to capture data packets in real-time called Wireshark. Wireshark is good for the network but has limited ability. This underscores the demand for cross-disciplinary tools that complement network-level analysis with more granular OSINT tools.

A recurring theme in this literature is that OSINT is being fundamentally transformed by automation and advanced analytics. WIreshark-type tools shine in network-centric endeavors, while AI-based technologies provide significant value in geospatial intelligence and the general scoping of investigative requirements. Yet, ethical and legal concerns remain such as privacy violations and misuse of OSINT frameworks. The use of unstructured and heterogeneous data sources also increases fears about accuracy and reliability.

OSINT's multifaceted applications extend to numerous fields of work ranging from cybersecurity, where it discovers vulnerabilities, repels threats, and monitors dark web activity. OSINT is applied by law enforcement in criminal profiling, locating

fugitives, and dismantling organized crime, and by corporations in helping to fortify information security and combat social engineering attacks. While there are numerous applications, variability of tool performance and lack of adaptations for specific contexts are still major challenges.

The paths taken by open-source intelligence (OSINT) investigators can be as varied as the investigators themselves, and technological and methodological fragmentation combined with a lack of standardization in tooling and practice impede OSINT's interoperability and scalability. Mitigating these challenges necessitates concerted initiatives to establish guidelines frameworks, datasets and ethical standards. OSINT changes the face of digital forensics as well as its technological maturity, but also its ethical issues, gaps in methodology, and lack of standardization. Further studies must tackle these GARs if OSINT wants to remain responsive to future digital threats.

(authors)	Year	Focus Area	Key Contributions	Limitations
Yadav et al.	2023	OSINT in cybersecurity and digital forensics	Comprehensive analysis of OSINT tools and integration with AI to enhance cybercrime investigations.	Lack of a unified framework for data integration; challenges in data reliability.
Waghmare	2023	Fugitive tracking using OSINT	Integration of OSINT with Interpol frameworks for real-time monitoring, geolocation tracking, and digital footprint analysis.	Ethical concerns regarding potential misuse; inconsistent adoption across jurisdictions.
Soepeno	2023	Network forensics with Wireshark	Demonstrated Wireshark's effectiveness in analyzing realtime network traffic and detecting anomalies.	Limited application scope to network traffic analysis; lacks broader forensic capabilities.

Yadav et al.	2023	AI and machine learning in OSINT	Highlighted AI's role in enhancing OSINT applications, including sentiment analysis and criminal profiling.	Absence of standardized datasets and limited scalability due to languagespecific models
Various Authors (e.g., Nouh et al., Kandias et al.)	Multiple	Broader OSINT applications in forensics and investigations	Covered a range of OSINT use cases including social media analysis, criminal profiling, and organized crime	Ethical concerns and reliance on public data that can be outdated or manipulated.
			investigations.	

Methodology for Investigating Network Traffic Using OSINT Tools

The purpose of this analysis was to use open-source intelligence (OSINT) tools which would be Wireshark and Shodan in examining anomalous network traffic to identify if what was seen was normal or malicious activity. The investigation looked for abnormalities in communication, opened ports and external IP activity.

Investigation Process

It starts with the analysis of the network traffic from the case study that Digital Corpora provided. Using Wireshark, a PCAP file of captured network traffic was analyzed for trends, anomalies, and potential malicious activity. The Wireshark findings were further enriched with external intelligence specific to the identified IP address using Shodan.

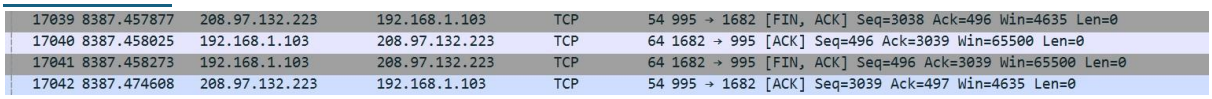
Upon reviewing the host's network activity with Wireshark, the communication was between an internal host (192.168.1.103) and external IP (208.97.132.223) over port 995 known to be used for secured email retrieval (POP3S). They shared information travel and network paths and were configured so they could exchange data with one another but were not necessarily set up for those email services. Additionally, the TCP connections teardown and teardowns with FIN and ACK flags were noted, which potentially indicate scanning or probing activity. There were also suspicion due to of some packets having a payload length of 0.

The sequence and acknowledgment numbers in the TCP connections seemed reasonable, for example, Seq=3038 and Ack=496. The TCP window size was abnormally large (65500), which could mean that the client tries to optimize the data flow, but their intention was most likely to exfiltrate data. Coupled with the absence of payloads, this suggested the possibility of clandestine communication or signalling activity.

The analysis of the external IP (208.97.132.223) on Shodan reported several open ports and services, which were: 995 (POP3S), 80 (HTTP), and 3306 (MySQL). The server also had several other vulnerabilities, such as CVE-2021-3618 (ALPACA attack) and CVE-2021-23017 (Nginx DNS resolver vulnerability). The hostname of the external IP is egleton, which is located in the United States. iad1-mysql-e2-3a. dreamhost. com. This was consistent with the suspicious external communication hypothesis, since the hostname suggests hosting services that could have been abused.

Results (or finding)

Wireshark



17039	8387.457877	208.97.132.223	192.168.1.103	TCP	54 995 → 1682 [FIN, ACK] Seq=3038 Ack=496 Win=4635 Len=0
17040	8387.458025	192.168.1.103	208.97.132.223	TCP	64 1682 → 995 [ACK] Seq=496 Ack=3039 Win=65500 Len=0
17041	8387.458273	192.168.1.103	208.97.132.223	TCP	64 1682 → 995 [FIN, ACK] Seq=496 Ack=3039 Win=65500 Len=0
17042	8387.474608	208.97.132.223	192.168.1.103	TCP	54 995 → 1682 [ACK] Seq=3039 Ack=497 Win=4635 Len=0

Figure 1

- Unusual communication on port 995 with overlapping FIN/ACK flags, a huge TCP window size (65500), and the absence of a payload. This might mean probing, data exfiltration, or covert activity. Examine the external IP and internal device for deviations.

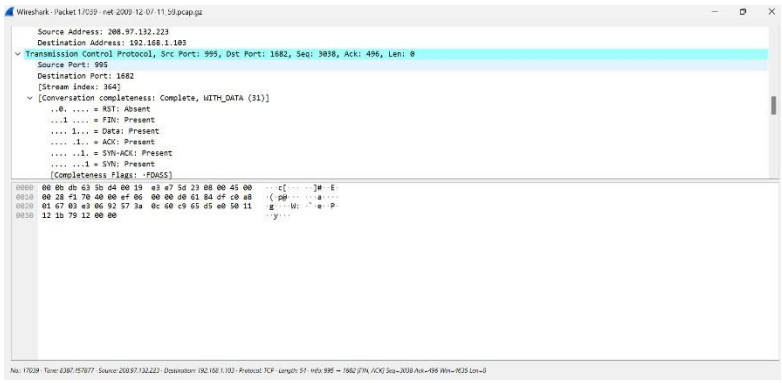


Figure 2

- this packet indicates a TCP connection from the internal IP 192.168.1.103 to the external IP 208.97.132.223 (port 995 POP3S) The connection is done with

FIN, ACK, SYN, SYN-ACK flags, but the length of the packet is 0, there is no payload. Indicates possible keep-alive signals, probing, or unusual traffic, particularly if unsolicited on this port.

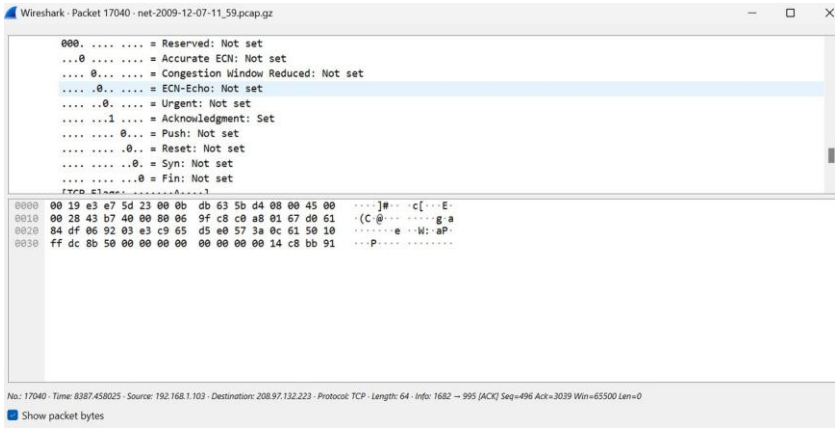


Figure 3

- The packet is TCP connection between 192.168.1.103 and 208.97.132.223 on port 995 (POP3S). The ACK flag in the packet is set, confirming receipt of previous communication. This very large window size of 65500, together with a Len=0 payload, is also not normal behavior for email traffic -- further indication that this transaction could be some form of signalling or covert activity.

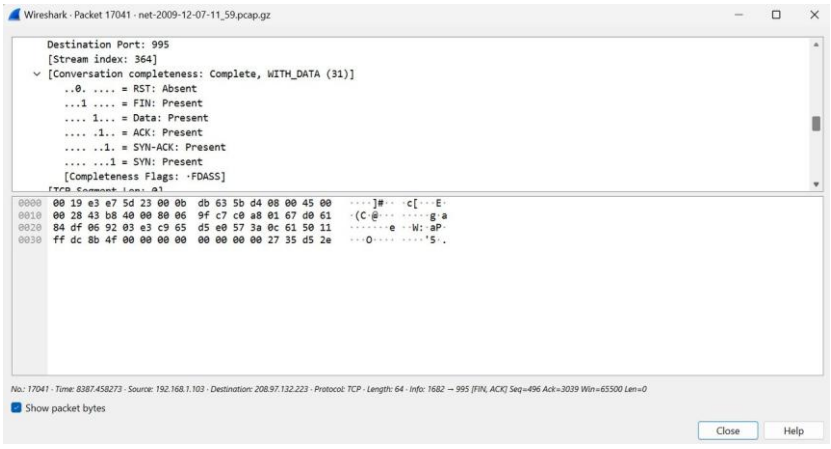


Figure 4

- This packet demonstrates a TCP connection from internal IP address 192.168.1.103 to external IP address 208.97.132.223 to port 995 (POP3S).

The connection contains FIN and ACK flags, which means the session closes. PacketLen= 0 Window= 65500 No payload data possible. So that we could be talking a “heartbeat” or signalling traffic which is potentially nasty or hidden communications.

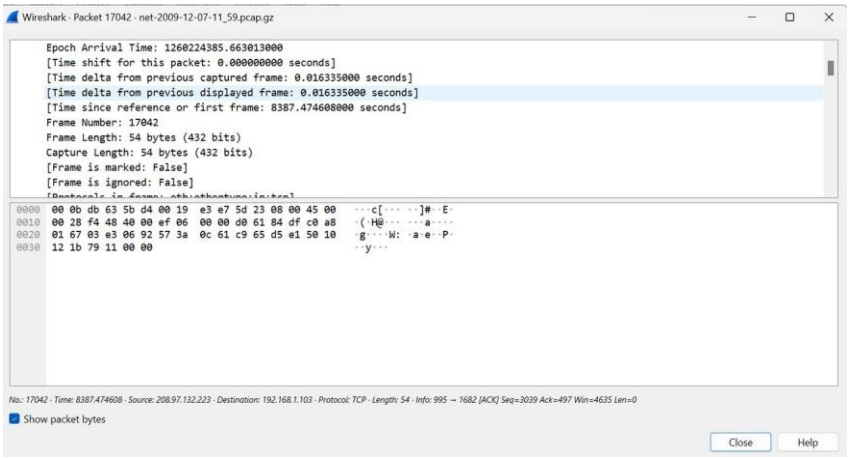


Figure 5

- Here is packet of TCP communication between external IP 208.97.132.223 and internal IP 192.168.1.103 using the port 995 (POP3S). The packet is an ACK that has sequence number 3039 and acknowledgment number 497 along with no payload (Len=0). It shows a window size of 4635, indicating that the communication is still ready. Multiple packets with no payload between the source and destination could indicate some form of signalling or keep-alive traffic rather than genuine data transfer.

Shodan

- The Shodan search reveals that IP 208.97.132.223 is hosted on a U.S. server via DreamHost and associated with MySQL-hosted names. Exposed services such as nginx and MySQL listen on ports 80, 123 and 3306, with further details outlined in CVE-2021-23017, CVE-2021-3618 and CVE-2023-44487, respectively, indicating the following potential attack vectors: protocol confusion and denial of service.

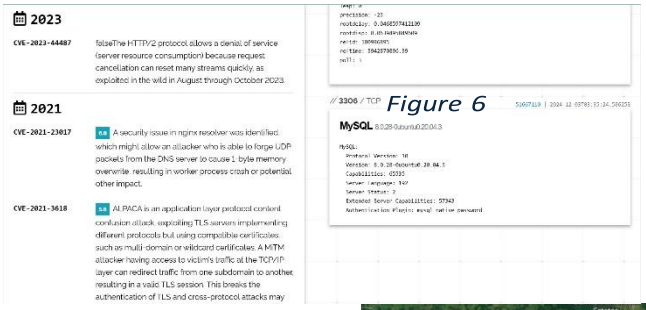


Figure 6

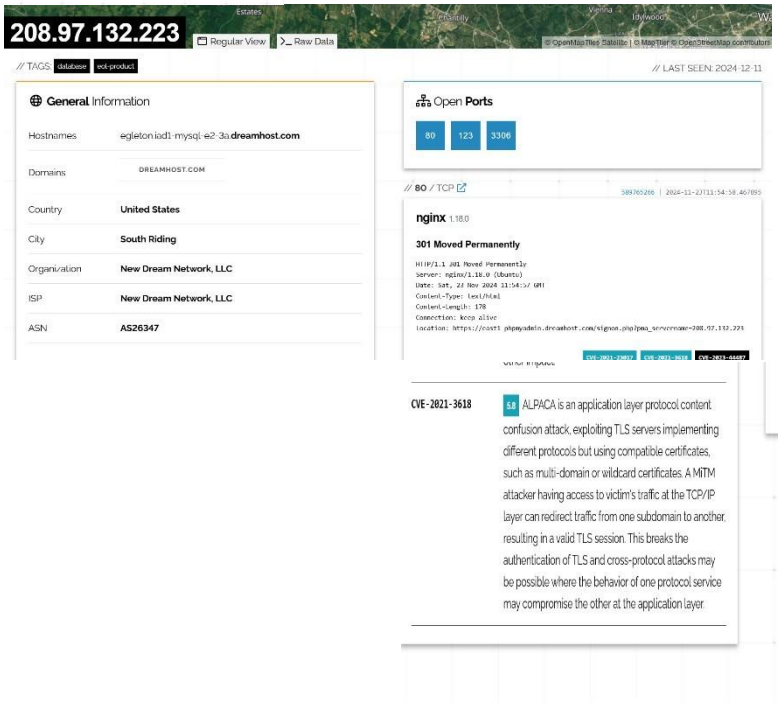


Figure 7

Figure 8

Discussion of Results (Analysis and Interpretation)

The analysis using OSINT tools, Wireshark and Shodan, provided critical insights into anomalous network behaviour and potential cybersecurity threats. The investigation differentiated between legitimate and malicious activities by examining communication patterns, open ports, and external IP activity.

Network Anomalies

Wireshark flagged unusual communication between the internal host (192.168.1.103) and the external IP (208.97.132.223) over port 995 (POP3S), typically used for secure email retrieval. Anomalies included overlapping FIN/ACK flags, abnormally large TCP window sizes (65500), and repeated zero-payload packets. These indicators suggest covert activities such as data exfiltration or signalling traffic rather than legitimate communication. Frequent connection teardowns with FIN/ACK flags further support the hypothesis of scanning or probing activity.

- TCP and Payload Analysis

By its own findings, TCP sequence and acknowledgment numbers seem within a normal range, yet the size of TCP windows indicates TCP payloads may not be human traffic. This behaviour indicates optimized data flow for exfiltration or covert communication attempts. Emails with such non-standard protocol behaviours deserve some further analysis. □ Shodan Findings

Shodan exposed many open ports (995, 80, 3306) on the external IP along with the vulnerabilities like CVE-2021-3618 (ALPACA attack) and CVE-2021-23017 (Nginx DNS resolver vulnerability). Such problems can expose the server to exploits including protocol confusion and denial of service attacks. Furthermore, the hosting service (DreamHost) association is consistent with our hypothesis of abuse for malicious purposes.

- Key Takeaways

Threat Indicators: Zero-payload packets, large TCP window sizes, and unusual flags very likely indicate malicious activity.

Vulnerability Exposure: Open Ports and Breitner Vulnerability

Wireshark Shodan integration advantages: Choosing Wireshark Shodan integration combines Wireshark packets tracker with external intelligence from Shodan □ Implications

The discoveries shine a light on the need for effective network monitoring to identify and mitigate threats. All this information, combined with tools like Wireshark and Shodan, adds to the ability to correlate anomalies with threat intelligence, which means being able to implement defences before being attacked. But the presence of vulnerabilities and anomalous
traffic patterns indicates a requirement for proactive defence strategies like patch management, secure configuration, and defensive monitoring.

Ultimately, this analysis shows the importance of OSINT tools in detecting network anomalies and solving cybersecurity problems. Forensic investigations and the protection of digital infrastructures will greatly benefit from improved detection algorithms and cross-disciplinary tool implementations.

Conclusion

This project casts a spotlight on the transformative capabilities of Open-Source Intelligence, or OSINT, in the realm of digital forensics and cybersecurity. OSINT grew to overcome sophisticated cyber threats, utilizing various tools such as Wireshark and Shodan to monitor and identify abnormal network traffic. This can pinpoint irregularities like zero-payload packets or unusually large TCP window sizes, which frequently indicate malicious behaviour. With Wireshark analysing packets and Shodan providing external intelligence, Wireshark can detect a number of useful vulnerabilities such as open ports and exploitable configurations of the server. The versatility of OSINT is also clear in its use: cybersecurity, criminal profiling, law enforcement and so on.

But there is room for improvement in some areas. The lack of standardization in OSINT methodologies reduces tool interoperability and the scalability of analyses. Confronting these challenges requires coordinated frameworks implemented equitably at scale, however, remains an ethical and legal minefield, with privacy violations and misuse top barriers to entry. To ensure responsible adoption, it is critical to establish strong ethical and regulatory mechanisms. Also, since public data is unstructured, it requires tools that enhance the validation and reliability of data.

The project highlights how OSINT can stay ahead of digital threats. Advancements in AI and machine learning will enable OSINT to digest ever-larger datasets. It is crucial for researchers, policymakers, and technology developers to work together to create standardized methodologies and ethical guidelines. Merging network-level analysis with cross-disciplinary tools will take OSINT to the next level. Its effectiveness will be further enhanced by the constant evolution of detection algorithms which are designed to recognise new threats. To sum up, OSINT is very important for the security of the digital ecosystem and its share future will be the one of the bottlenecks of the forensic analysis evolution.

References

- Ashok Yadav1 · Atul Kumar2 · Vrijendra Singh. (2023). *ink.springer.*. [Online]. Open-source intelligence: a comprehensive review of the current state, applications and future perspectives in cyber s. Last Updated: 15 March 2023. Available at: https://link.springer.com/article/10.1007/s10462-023-10454-y [Accessed 13 December 2024].

- Muhammeadali nellyullathill. (2020). *Teaching Open Source Intelligence (OSINT) Journalism: Strategies and Priorities*. [Online]. communication and journalism research. Last Updated: 30 june 2020. Available at: https://cjrjournal.in/Article/Teaching%20Open%20Source%20Intelligence%20(OSINT)%20Journalism%3a%20St [Accessed 13 December 2024].

- Pranav Waghmare. (2024). *Fugitive Tracking using Open Source Intelligence (OSINT)*. [Online]. SSRN. Last Updated: 07 march 2024. Available at:https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4719968 [Accessed 13 December 2024].

- Tiago M. Fernández-CaramésPaula, Fraga-Lamas. (2020). *Teaching and Learning IoT Cybersecurity and Vulnerability Assessment with Shodan through Practical Use Cases*. [Online]. mdpi. Last Updated: 27 May 2020. Available at: https://www.mdpi.com/1424-8220/20/11/3048 [Accessed 13 December 2024]

- Ryufath Alief Adhyaksa Putera Soepeno. (2023). *Wireshark: An Effective Tool for Network Analysis*. [Online]. researchgate. Last Updated: September 2023. Available at: https://www.researchgate.net/publication/374675769_Wireshark_An_Effective_Tool_for_Network_Analysis [Accessed 13 December 2024].