International Journal on Science and Technology (IJSAT)

E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

A Modern Approach to IP and Email Tracing for Cybercrime Investigations: Challenges and Future Enhancements

Mahek Patel[1], Prof. Nimisha Chaudhari[2]

Department of cyber security, GIT Gandhinagar University

Abstract

Cybercrime investigations rely on the accurate analysis of IP addresses and email metadata to identify the origin of malicious activities. However, modern attackers increasingly use VPNs, proxies, spoofed email headers, forwarding services, and cloudbased infrastructure, making conventional tracing methods less effective. This review paper examines the latest techniques in IP tracing, email header analysis, and email authentication mechanisms such as SPF, DKIM, and DMARC, and evaluates their performance based on recent studies. The paper highlights key challenges including anonymization layers, inconsistent logging practices, header manipulation, and the limitations of existing authentication standards. Through a comparative analysis of current research, the study identifies a critical gap: existing methods often operate independently and lack integrated forensic intelligence. To address this, the paper outlines potential future enhancements, including hybrid AI-driven correlation models, improved metadata preservation, enhanced authentication protocols, and collaborative forensic platforms. These advancements aim to improve accuracy, reduce attribution errors, and support more reliable cybercrime investigations.

Keywords — IP tracing, email forensics, spoofing, cybercrime investigation, DMARC, attribution, digital evidence.

1. Introduction

Cybercrime continues to grow in complexity as attackers exploit the openness of digital communication networks. Email-based threats, phishing campaigns, and identity spoofing remain among the most common attack vectors, while IP-based attacks often rely on anonymization tools to hide the origin of malicious traffic. As a result, investigators depend heavily on IP tracing and email header analysis to establish evidence, track attacker behavior, and identify the source of cyber incidents.

However, modern Internet infrastructure—such as NAT, VPNs, proxies, cloud-based routing, and forwarding services—makes attribution extremely challenging. Likewise, traditional email tracing is often weakened by spoofed headers, manipulated routing information, and limitations of authentication protocols like SPF, DKIM, and DMARC. To address these issues, researchers have proposed several enhanced techniques aimed at improving accuracy, reducing false positives, and supporting digital investigations.

This review paper examines recent literature on IP tracing and email forensics, highlights major challenges, and presents future enhancements including hybrid correlation models and AI-driven forensic intelligence.

2. Literature Review

2.1 Email Authentication & Spoofing Detection

Several works focus on understanding and mitigating email spoofing and on evaluating authentication protocols.

- Sethuraman et al. provide a comprehensive examination of email spoofing techniques and the weaknesses that attackers exploit in practice, highlighting how spoofed headers and domain impersonation remain major threats despite authentication standards [2].
- Surveys and tool comparisons (Altulaihan et al. [3]) outline available forensic methods and emphasize the need for standardized accuracy benchmarks and better tool integration.
- Practical evaluations of SPF, DKIM, and DMARC show that while these protocols raise the bar against naïve spoofing, they suffer from real-world limitations. DMARC helps domain owners specify policy, but forwarding and mailing-list transformations commonly break authentication results or produce false negatives [5], [8].
- Several studies propose improved server-level and memory-forensics–augmented detection mechanisms that combine header checks with machine learning or live-memory artifacts to detect manipulated or spoofed messages [1], [4], [6], [7]. These approaches improve detection in controlled settings but raise concerns about scalability and operational cost.

2.2 IP Tracing & Network Forensics

Research on IP traceback, geolocation, and network-flow correlation highlights technical limitations when attackers use anonymization services.

- Surveys on IP traceback mechanisms summarize a range of techniques (packet-marking, logging-based traceback, flow correlation) and their trade-offs in terms of storage, cooperation requirements, and deployability [8], [10].
- Geolocation and IP database studies point out accuracy issues and variability across providers, especially for mobile and cloud-hosted IPs [9], [11].
- Studies on cloud-based routing, NAT, and TOR exit-node behavior show that anonymization and multi-hop routing severely limit the direct usefulness of an observed IP address for attributing a human actor [12], [13].
- Network-flow correlation and hybrid flow techniques can improve traceback when combined with server logs and ISP cooperation, but they require greater logging retention and cross-organizational data sharing [12], [19].

2.3 Forensic Readiness, Digital Evidence, and Legal/Operational Challenges

Multiple works address operational constraints, data retention, and cross-jurisdictional issues.

- Studies on digital evidence preservation and cloud forensics stress the need for standardized logging, longer retention periods, and automated evidence collection mechanisms to support investigations [13], [15].
- Reviews of cybercrime investigation practices point to gaps in institutional readiness — many organizations lack the procedures or expertise to preserve forensic-quality logs, which diminishes investigative success [14], [16].
- Legal and multi-jurisdictional hurdles frequently hinder timely access to ISP or cloud provider logs, increasing the complexity and delay for cross-border cases [13], [14].

2.4 AI, Machine Learning & Hybrid Forensic Approaches

Recent literature explores AI/ML integration into forensic pipelines and proposes hybrid models.

- Surveys and empirical studies show that ML is effective for email classification, anomaly detection, and attribution when trained on large, representative datasets [17], [19].
- Hybrid proposals combine content analysis, header/routing metadata, network-flow correlation, and behavioral features into unified models that output confidence scores or cluster related events [18], [20]. These systems can successfully correlate disparate evidence types and highlight likely linkages among incidents.
- However, existing work often remains at prototype level; common limitations include dataset scarcity, generalizability, interpretability of ML decisions, and the computational costs of running large-scale correlation engines in operational settings [17], [18], [20].

Synthesis & Identified Gaps

Across the surveyed literature, clear patterns emerge:

1. **Fragmentation of methods:** Most studies focus on either email-level analysis or network-level traceback; few comprehensively link the two. This isolation reduces the effectiveness of attribution in real-world multi-layer evasion scenarios [2], [8], [19].
2. **Authentication shortfalls:** SPF/DKIM/DMARC adoption improves trust, but forwarding, mailing lists, and domain misconfigurations frequently disrupt authentication—leaving a gap that attackers exploit [5], [8].
3. **Operational constraints:** Lack of forensic readiness (logging, preservation), cross-jurisdictional access, and resource limitations in many organizations reduce the practical applicability of advanced techniques [13], [15].
4. **Need for hybrid, explainable AI:** While ML-enhanced detectors and hybrid correlation engines show improved performance, studies emphasize the need for larger, diverse datasets, standardized benchmarks, and methods that provide interpretable evidence for investigators [17], [18], [20].

3 Problem Statement

Although IP tracing and email forensics are essential tools for cybercrime investigations, existing methods often fail when attackers use layered anonymization, spoofing, forwarding, or cloud-based services. Current approaches work in isolation, lack integration, and offer limited accuracy in real-world scenarios. There is a need for more advanced, combined, and intelligent tracing methods that can reliably identify the true origin of cybercrime activities.

4 Aim & Objective

To review modern approaches to IP and email tracing for cybercrime investigations and identify future enhancements that improve attribution accuracy and reliability.

Objectives:

- To study existing techniques used in IP tracing and email header analysis.
- To evaluate the effectiveness of email authentication standards such as SPF, DKIM, and DMARC.
- To compare and analyze recent research contributions from 2018–2024.
- To identify key challenges in tracing cybercriminal activity.
- To propose future enhancements, including AI-driven hybrid forensic frameworks.

5 Research Question / Identified Gaps

1. How effective are current IP tracing techniques in detecting the real source of cyber-attacks?
2. What limitations exist in email header analysis and authentication mechanisms?
3. How do anonymization tools such as VPNs and proxies impact cybercrime attribution?
4. Can combining IP logs, email metadata, and AI-based analysis improve tracing accuracy?
5. What future technologies or enhancements can strengthen digital investigations?

Key Gaps

- Lack of **integration** between IP tracing and email forensics.
- Limited accuracy when attackers use **multi-layer anonymization**.
- Email authentication protocols fail in **forwarding or mailing-list scenarios**.
- Absence of **AI-driven correlation models** across network and email data.
- Insufficient real-world datasets for testing forensic methods.

6 Proposed Solution (Scope for future work)

1. **Develop a** hybrid forensic system **that combines IP tracing and email analysis together.**
2. **Use** AI and machine learning **to detect spoofing patterns and suspicious behavior.**
3. **Improve** email authentication **(SPF, DKIM, DMARC) to work even after forwarding or mailing-list changes.**
4. **Store and preserve** better logs and metadata **for investigation (ISP logs, server logs).**
5. **Create** automated tools **that analyze email headers, routing paths, and IP data in one platform.**
6. **Encourage** collaboration between ISPs, email providers, and law-enforcement **for faster investigations.**
7. **Explore** blockchain-based **logging to prevent tampering with digital evidence.**
8. **Build** large datasets **for training and testing AI forensic models.**

7 Conclusion

This review paper analyzed modern methods of IP and email tracing used in cybercrime investigations, highlighting significant improvements as well as persistent challenges. While techniques such as IP geolocation, packet traceback, email header analysis, and authentication protocols like SPF, DKIM, and DMARC provide valuable forensic insights, they often fail in scenarios involving anonymization, spoofing, forwarding, or cloud-based routing.

Recent research demonstrates the need for integrated and intelligent forensic systems. The identified research gap shows that current approaches operate independently, reducing accuracy in real-world cases. Therefore, future enhancements must focus on hybrid AI-based forensic frameworks, improved metadata preservation, stronger authentication standards, and collaborative forensic ecosystems. These innovations can significantly enhance attribution accuracy and support more effective cybercrime investigations.

References

1. A. K. Jain, R. Kumar and D. Singh, "Email header analysis for detection of spoofed emails," International Journal of Computer Applications, vol. 182, no. 5, pp. 1–7, 2018.
2. S. C. Sethuraman, A. Choudhary and N. Chauhan, "A comprehensive examination of email spoofing," Computers & Security, vol. 134, 102013, 2024.
3. A. Altulaihan et al., "Email security issues, tools, and techniques used in email forensics," Sustainability, vol. 15, no. 13, pp. 10612, 2023.
4. J. L. Mane, P. Patil and S. Sawant, "Reliable email spoofing detection using enhanced cybersecurity approaches," in Proc. International Conference on Intelligent Computing, 2024.
5. A. Herzberg and H. Shulman, "DMARC: From spoofing prevention toward domain-based message authentication," IEEE Security & Privacy, vol. 16, no. 3, pp. 72–78, 2019.
6. R. K. Yadav and V. Gupta, "Detection of phishing and spoofed emails using header analysis and machine learning," International Journal of Engineering and Advanced Technology, vol. 9, no. 3, pp. 1402–1407, 2020.
7. A. Oest et al., "PhishFarm: A scalable framework for measuring the effectiveness of email authentication," in Proc. USENIX Security Symposium, 2019.
8. A. Bhandari and V. Varadharajan, "A comprehensive survey on IP traceback mechanisms," Journal of Network and Computer Applications, vol. 167, 102739, 2020.
9. M. Bargh and S. E. Lee, "IP geolocation techniques: A review," IEEE Communications Surveys & Tutorials, vol. 23, no. 4, pp. 2828–2856, 2021.
10. A. K. Sood and R. Enbody, "IP address spoofing and traceback: A survey," ACM Computing Surveys, vol. 52, no. 2, pp. 1–36, 2019.
11. S. Alqahtani and K. Elleithy, "Network forensics: Techniques, tools, and trends," IEEE Access, vol. 9, pp. 122–140, 2021.
12. X. Wang, Q. Li and Y. Liu, "Tracing cyber-attacks using network flow correlation techniques," Future Generation Computer Systems, vol. 108, pp. 453–463, 2020.
13. K. Raghav and M. Tripathy, "Digital forensics challenges in cloud computing: A survey," Journal of Information Security and Applications, vol. 54, 102563, 2020.
14. S. Zargari and A. Benison, "Cybercrime investigation: A review of challenges and models," Digital Investigation, vol. 37, 301200, 2021.
15. S. R. Chithralekha and B. B. Moni, "Role of digital forensics in cybercrime investigation," Forensic Science International: Digital Investigation, vol. 44, 301418, 2023.
16. M. K. Rogers and K. Seigfried-Spellar, "Cybercrime motives and behavioral patterns," Computers in Human Behavior, vol. 118, 106683, 2021.
17. G. Xu, W. Yu, D. Griffith and N. Golmie, "A survey on machine learning for cybersecurity," IEEE Communications Surveys & Tutorials, vol. 21, no. 1, pp. 287–327, 2019.
18. P. Prakash, Y. Li and M. Kumar, "AI-assisted digital forensics: Techniques and applications," IEEE Access, vol. 10, pp. 11532–11549, 2022.
19. N. Abedi and S. M. Shamsuddin, "Email classification and forensic analysis using machine learning," in Proc. International Conference on Computational Science and Technology, 2022.
20. A. Hoque, N. Hasan and M. Rahman, "Hybrid models for cyber-attack attribution using AI and log correlation," Expert Systems with Applications, vol. 229, 120601, 2023.