PROBLEMS OF CYBERCRIME INVESTIGATION IN UZBEKISTAN

Jumayev Bekzod Baxtiyorovich
Academy of the General Prosecutor's Office Republic of Uzbekistan
Independent researcher

Annotation: As technology has become pivotal a part of life, it has also become a part of criminal life. Criminals use new technology developments to commit crimes, and investigators must adapt to these changes. Many people have, and will become, victims of cybercrime, making it even more important for investigators to understand current methods used in cyber investigations. The two general categories of cyber investigations are digital forensics and open-source intelligence. Cyber investigations are affecting more than just the investigators. They must determine what tools they need to use based on the information that the tools provide and how effectively the tools and methods work. Tools are any application or device used by investigators, while methods are the process or technique of using a tool. This survey compares the most common methods available to investigators to determine what kind of evidence the methods provide, and which of them are the most effective. To accomplish this, the survey establishes criteria for comparison and conducts an analysis of the tools in both mobile digital forensic and open-source intelligence investigations. We found that there is no single tool or method that can gather all the evidence that investigators require. Many of the tools must be combined to be most effective. However, there are some tools that are more useful than others. Out of all the methods used in mobile digital forensics, logical extraction and hex dumps are the most effective and least likely to cause damage to the data. Among those tools used in open-source intelligence, natural language processing has more applications and uses than any of the other options.

Key words: cybercrime; open-source intelligence; mobile forensics; digital forensics; cyber investigations problems.

Ever since the creation of the Internet, people have been finding ways to conduct illegal activities using it as a tool. In order to counteract these actors, technologies and methods have been developed to track these criminals. It is critical for security and law enforcement professionals to understand these technologies and how they are developing, so they can better perform in their job roles. Internet crime is something that affects anyone who uses a computer, thus making it critically important to counteract it in any way possible.

Some of the most common technologies and methods for tracking cyber criminals are digital forensics and online investigations, which leverages open-source

intelligence (OSINT). Within these areas, there are many different technologies and techniques that can be used to gather data on the malicious actor. This data can then be aggregated to determine who committed the crime and build a case against the individual. This paper will cover a survey of these technologies and the methods associated with them.

Digital forensics is a key field used in combating cybercrime because it can be useful if the case is presented in court. Digital forensics helps investigators piece together evidence and determine the timeline of events in a crime. It is mainly made up of network forensics and memory/disk analysis. By analyzing information found on disks and through networks, investigators can learn about other potential conspirators in the crime. This could help them track down these individuals and stop them before another crime is committed.

Much of the tracking of criminals is done online. The different layers of crime on the Internet can be broken up into three categories: (1) the surface or open web, (2) the deep web and (3) the dark web. These areas of the web contain a host of information that can be valuable to investigations, so it is important to understand the methods that allow investigators gather and use this information. For example, investigators can utilize information regarding cryptocurrency transactions on the dark web to learn about criminal activity.

The changes and developments in this field are occurring rapidly and it is important for security professionals to keep up to date. Some new developments are coming from automation and machine learning. By automating their tools, investigators can speed up their process and reach their goal sooner. AI forensics will in the future help combat the growing trend of AI crime. This paper will also cover a summary of the developing technologies in this area and how they could change investigations in the future.

Having this information compiled in a single document allows for easy comparison of methods and the information that these methods provide. None of the methods available to investigators are able to gather all the information they require for a case, making it even more important to understand how this information is gathered and how to fill the information gaps. If investigators are able to gain a complete picture of a crime, then they will be able to take action against the criminal or potentially stop a future crime from occurring.

Digital forensics is the practice of collecting and organizing information found on an electronic device for investigative purposes. It is important to know both the technologies and the methods and frameworks investigators use in this field.

Digital forensics can be broken into four areas: host forensics, mobile forensics, network forensics, and cloud forensics. Each of these four areas provides investigators

with different kinds of information, with very little overlap. By breaking this field up into these four areas, this paper can analyze the methods for each without covering the same technique twice. This makes it ideal to categorize the methods into these areas because many of the techniques for gathering and analyzing the information from these sources are unique to each source.

2.1. Host Forensics

Host forensics is often called digital forensics because it encompasses forensics done on "normal" devices, such as desktops, servers, and other non-specialized sources of data. This method has been long established, but the tools used are constantly evolving as technology is progressing.

Investigators can also utilize the method of weighting forensic evidence with blockchain technology. This can help with certifying the validity of digital evidence with it is presented in a court. This weighting system first collects evidence in a blockchain that records when the evidence was collected and who was in possession of it at the time. This data can then be categorized by relevance to the case and a timeline of events can be created [1]. This method allows investigators to confidently show that evidence was processed correctly and was not tampered with. It can also be helpful with IoT forensics because of the large amount of information gathered in those investigations [2]. An example to demonstrate how weighted forensic evidence can be used is if, after investigators collect evidence for a case, they need some way to prove to the jury that the evidence has not been tampered with. Because of the structure of blockchain and its unchangeable nature, investigators can document the chain of custody for the evidence, showing that it was never unaccounted for.

Something investigators must take into account when performing forensics is the operating system of the device in question. Each operating system performs tasks differently and stores information in different places in the system, which affects all areas of digital forensics [3]. It is critical for investigators to be familiar with the many different types and versions of operating systems in order for them to be able to gather all relevant evidence.

Another challenge that investigators face with host forensics is the randomization of kernel addresses. In order to face this problem, investigators can use four approaches to derandomize this information: brute force code, patched code, unpatched code, and read only kernel data. The brute force method simply scans the entire kernel code. For the patched code option, the kernel must know where to apply patches. The signature from this gives investigators insight into the organization of randomized address locations. The unpatched code signatures come from the code that has been identified

as having not been patched. Finally, for read only kernel data, static pointers can help investigators shift data to find offsets, which will lead them to the proper address [4].

As technology has developed, mobile devices have become more common. This means that mobile forensics is a critical part of investigations and should be understood by anyone in the field. Mobile forensics is distinct from any other kind of forensics because of the difference in "hardware, software, power consumption, and overall mobility" [3]. Furthermore, mobile devices are presumed to have personal data, which could be critical to an investigation.

There are four investigation phases in mobile forensics investigations: preservation, acquisition, examination analysis, and reporting. These phases are depicted in **Figure 1**.

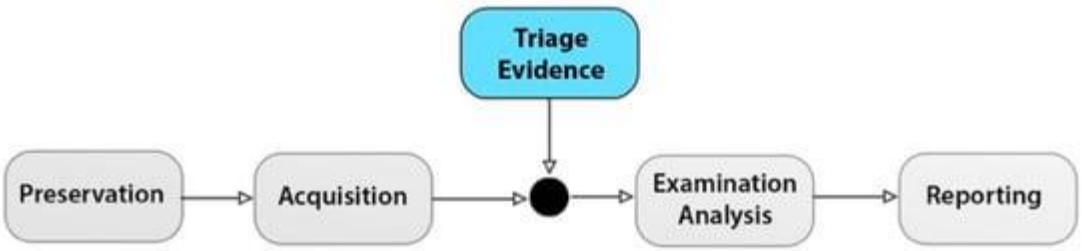


Figure 1. Mobile forensic investigative phases [3].

The preservation phase is where mobile devices are taken by investigators and tracked to ensure that the data on them is not tampered with. The acquisition phase is where the data on the mobile device is copied to another device for the analysis that occurs in the examination analysis phase. Finally, the reporting phase is where all the information investigators uncover in the examination analysis phase is documented [3]. Each of these phases must be followed properly to ensure the integrity of any investigation involving mobile devices.

There are common collection methods, also called data extraction, used in mobile forensics. Data from mobile devices must be extracted during investigations. There are five levels of data extraction: manual, logical, hex dumps, chip-offs, and micro reads [5]. Each of these options allow investigators to gather different information from different areas of the device with varying levels of complexity. **Table 1** shows a comparison of these methods based on the criteria described in **Section 1**.

Table 1. Comparison of the methods of mobile data extraction [5].

Finally, micro reads are the most complicated method out of these five. They use electron microscopes to analyze the logic gates in order to determine the readable data.

This method is considered a "last resort" method because it is challenging and resource exhaustive. Micro reads are not applicable to many case scenarios because of their challenging nature.

As shown in Table 1, manual extraction, although easy to perform, is the least recommended because of the risk it poses to the data's integrity. The best methods are logical extraction and hex dumps. These analyze information from different places, so they give investigators a method of gathering different evidence that the other method does not access. Logical extraction and hex dumps have medium or low complexity, making them faster and more efficient to use. Finally, both of these methods pose a low risk to data integrity because they utilize a separate workstation for data manipulation.

Network forensics is the practice of analyzing information from a host or an entire network [5]. The forensic information can be obtained through logs or traffic captures.

Three of the layers of the TCP/IP Model can provide investigators with useful information. These layers are the application, transport, and network layers. The only layer not included in this is the network interface layer, which includes ethernet frames and the physical connections of a network. Forensic information can be gathered from the application layer through logs that hosts gather. This can be information regarding failed logons or timestamps, which could be critical information in an investigation. The transport and network layer are where firewalls are classified. Firewalls, if properly configured, can contain log data of traffic that has been dropped from the network [6]. This can give investigators information about potentially malicious traffic that has been seen by the firewall. Figure 2 shows the relationship between the layers and the information that investigators can gather by showing the flow of traffic though the network model and the devices it affects.

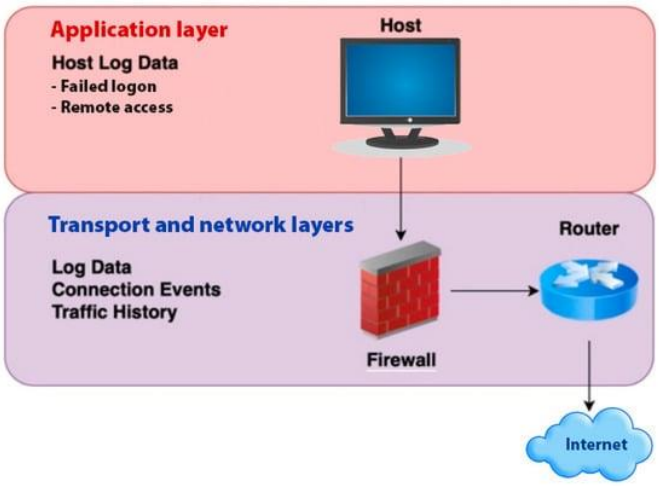
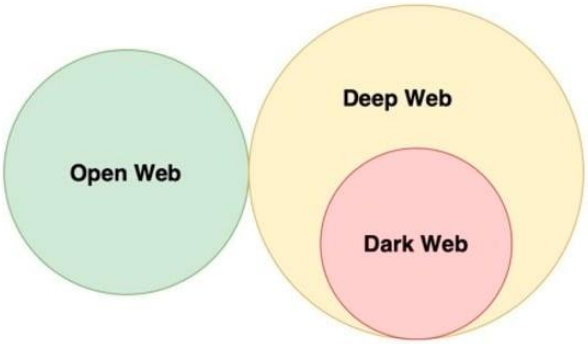
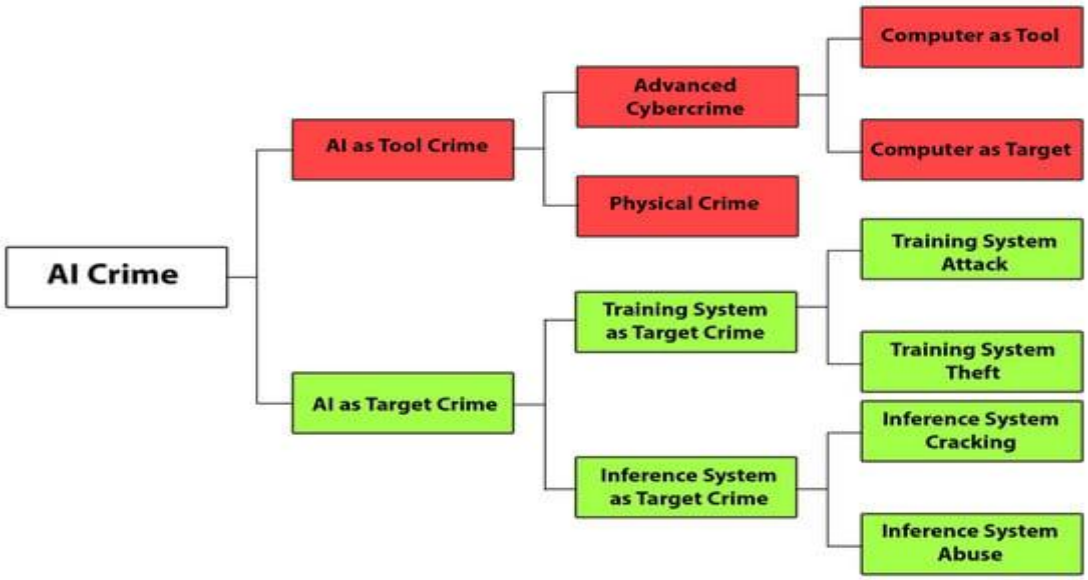


Figure 2. Evidence gathered from network layers.

Online investigations are the process of gathering, structuring, and using information that can be obtained online. These can be performed by law enforcement, security professionals, or any individual. The main method for gathering information in online investigations is Open-source intelligence (OSINT). This method is the aggregation and use of the information that is gathered using other methods described in sections below. The information gathered for this type of investigation shows relationships, identities, or events that are relevant to the cyber investigation.



Machine learning can also be applied to the criminal aspect of investigations. **Figure 5** shows a taxonomy for AI crimes. This taxonomy shows how AI can be used by criminals as a tool, but also as a target of their crimes. If criminals can harm a victim's AI systems, it could cause a lot of damage to the victim and their systems. Also, criminals can essentially teach their AI systems to attack the victim's systems, which causes the attack to be faster and more sophisticated than attacks done by individuals [24]. Because this method can be used by criminals, it is critical for investigators to understand this approach and know how to handle it during investigations.



With the advance of technology, criminal investigations rely more on technology as threat actors are becoming more advanced. There are many tools and methods investigators can implement to assist in their jobs. None of these tools can perform all the functions that investigators require. Therefore, it is necessary to become familiar with different tools, how they function, and what information they can provide to investigators.

The field of digital forensics has been around for some time, but it is still evolving. The four main areas, host, mobile, network, and cloud forensics are still critical to digital investigations, especially as new methods are being developed. There are multiple ways to extract data in mobile forensics. The most effective methods are logical extraction and hex dumps. Digital forensics' main area of growth is in the field of cloud forensics, especially as many services are becoming cloud based.

Open-source intelligence and online investigations are not a new method, but investigators are always applying new technologies to these methods. Multiple methods in this field can be applied to online investigations in order to gain as much intelligence on threat actors as possible. Each of these methods provide some information, but none of them provide all the information necessary for an investigation. However, out of all the methods available to investigators, one stands out as having the most applications: natural language processing. This method can be applied to many different cases and provides investigators with several different kinds of information for these cases.

Automation and machine learning are advancing the field technology and cyber investigations are also being affected by these technologies. Automation is helping investigators speed up the process of collecting evidence, while machine learning is helping investigators identify and classify this evidence. Automation also presents challenges to this field in regard to legal assumptions and implications [18].

This is a growing field and there are many opportunities for further research. Automation and machine learning provide potential areas of further research as these technologies become more sophisticated. Open-source intelligence techniques were also found to be underrepresented in the research field, opening another area for future research.

References

1. Billard, D. Weighted Forensics Evidence Using Blockchain. In Proceedings of the 2018 International Conference on Computing and Data Engineering, Shanghai, China, 4–6 May 2018; pp. 57–61. [Google Scholar] [CrossRef][Green Version]
2. Zhang, L.; Li, F.; Wang, P.; Su, R.; Chi, Z. A Blockchain-Assisted Massive IoT Data Collection Intelligent Framework. *IEEE Internet Things* **2021**, *15*. [Google Scholar] [CrossRef]
3. Barmpatsalou, K.; Cruz, T.; Monteiro, E.; Simoes, P. Current and Future Trends in Mobile Device Forensics. *ACM Comput. Surv.* **2018**, *51*, 1–31. [Google Scholar] [CrossRef]
4. Gu, Y.; Lin, Z. Derandomizing Kernel Address Space Layout for Memory Introspection and Forensics. In Proceedings of the Sixth ACM Conference on Data and Application Security and Privacy, New Orleans, LA, USA, 9–11 March 2016; ACM: New York, NY, USA, 2016; pp. 62–72. [Google Scholar] [CrossRef][Green Version]
5. Chernyshev, M.; Zeadally, S.; Baig, Z.; Woodward, A. Mobile Forensics: Advances, Challenges, and Research Opportunities. *IEEE Secur. Priv.* **2017**, *15*, 42–51. [Google Scholar] [CrossRef]
6. Caviglione, L.; Wendzel, S.; Mazurczyk, W. The future of digital forensics: Challenges and the road ahead. *IEEE Secur. Priv.* **2017**, *15*, 12–17. [Google Scholar] [CrossRef]
7. Stoyanova, M.; Nikoloudakis, Y.; Panagiotakis, S.; Pallis, E.; Markakis, E.K. A survey on the Internet of things (IoT) forensics: Challenges, approaches, and open issues. *IEEE Commun. Surv. Tutor.* **2020**, *22*, 1191–1221. [Google Scholar] [CrossRef]
8. Manral, B.; Somani, G.; Choo, K.-K.R.; Conti, M.; Gaur, M.S. A Systematic Survey on Cloud Forensics Challenges, Solutions, and Future Directions. *ACM Comput. Surv.* **2020**, *52*, 1–38. [Google Scholar] [CrossRef][Green Version]
9. Khan, S.; Gani, A.; Wahab, A.W.; Bagiwa, M.A.; Shiraz, M.; Khan, S.U.; Zomaya, A.Y. Cloud Log Forensics: Foundations, State of the Art, and Future Directions. *ACM Comput. Surv.* **2016**, *49*, 1–42. [Google Scholar] [CrossRef]
10. Tavabi, N.; Bartley, N.; Abeliuk, A.; Soni, S.; Ferrara, E.; Lerman, K. Characterizing Activity on the Deep and Dark Web. In Proceedings of the Companion of The 2019 World Wide Web Conference, San Francisco, CA, USA, 13–17 May 2019; ACM: New York, NY, USA, 2019; pp. 206–213. [Google Scholar] [CrossRef][Green Version]
11. Celestini, A.; Me, G.; Mignone, M. Tor marketplaces exploratory data analysis: The Drugs Case. In *Global Security, Safety and Sustainability–The Security Challenges of the Connected World*; Springer: New York, NY, USA, 2016; pp. 218–229. [Google Scholar] [CrossRef]

12. Internet Organized Crime Threat Assessment. 2020. Available online: **https://www.europol.europa.eu/activities-services/main-reports/internet-organised-crime-threat-assessment-iocta-2020** (accessed on 12 September 2021).
13. Alonso-Fernandez, F.; Belvisi, N.M.; Hernandez-Diaz, K.; Muhammad, N.; Bigun, J. Writer Identification Using Microblogging Texts for Social Media Forensics. *IEEE Trans. Biom. Behav. Identity Sci.* **2021**, *3*, 405–426. [Google Scholar] [CrossRef]
14. Nazah, S.; Huda, S.; Abawajy, J.; Hassan, M.M. Evolution of dark web threat analysis and detection: A systematic approach. *IEEE Access* **2020**, *8*, 171796–171819. [Google Scholar] [CrossRef]
15. Edwards, M.; Rashid, A.; Rayson, P. A Systematic Survey of Online Data Mining Technology Intended for Law Enforcement. *ACM Comput. Surv.* **2015**, *48*, 54. [Google Scholar] [CrossRef][Green Version]
16. Zhang, X.; Li, W.; Ying, H.; Li, F.; Tang, S.; Lu, S. Emotion Detection in Online Social Networks: A Multilabel Learning Approach. *IEEE Internet Things J.* **2020**, *7*, 8133–8143. [Google Scholar] [CrossRef]
17. Liao, X.; Yuan, K.; Wang, X.F.; Li, Z.; Xing, L.; Beyah, R. Acing the IoC Game: Toward Automatic Discovery and Analysis of Open-Source Cyber Threat Intelligence. In Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, Vienna, Austria, 24–28 October 2016; ACM: New York, NY, USA, 2016; pp. 755–766. [Google Scholar] [CrossRef]
18. Završnik, A. Criminal justice, artificial intelligence systems, and human rights. *ERA Forum* **2020**, *20*, 567–583. [Google Scholar] [CrossRef][Green Version]
19. Raaijmakers, S. Artificial Intelligence for Law Enforcement: Challenges and Opportunities. *IEEE Secur. Priv.* **2019**, *17*, 74–77. [Google Scholar] [CrossRef]
20. Zhang, F.; Li, W.; Zhang, Y.; Feng, Z. Data Driven Feature Selection for Machine Learning Algorithms in Computer Vision. *IEEE Internet Things J.* **2018**, *5*, 4262–4272. [Google Scholar] [CrossRef]
21. Du, X.; Hargreaves, C.; Sheppard, J.; Anda, F.; Sayakkara, A.; Le-Khac, N.-A.; Scanlon, M. SoK: Exploring the state of the art and the future potential of artificial intelligence in digital forensic investigation. In Proceedings of the 15th International Conference on Availability, Reliability and Security, Virtual Event, Ireland, 25–28 August 2020; pp. 1–10. [Google Scholar] [CrossRef]
22. Shaukat, K.; Luo, S.; Varadharajan, V.; Hameed, I.A.; Xu, M. A Survey on Machine Learning Techniques for Cyber Security in the Last Decade. *IEEE Access* **2020**, *8*, 222310–222354. [Google Scholar] [CrossRef]
23. Zhang, X.; Liu, L.; Xiao, L.; Ji, J. Comparison of Machine Learning Algorithms for Predicting Crime Hotspots. *IEEE Access* **2020**, *8*, 181302–181310. [Google Scholar] [CrossRef]

24. Jeong, D. Artificial intelligence security threat, crime, and forensics: Taxonomy and open issues. *IEEE Access* **2020**, *8*, 184560–184574. [Google Scholar] [CrossRef]
25. Quick, D.; Choo, K.-K.R. Digital forensic intelligence: Data subsets and Open-Source Intelligence (DFINT+OSINT): A timely and cohesive mix. *Future Gener. Comput. Syst.* **2018**, *78*, 558–567. [Google Scholar] [CrossRef]
26. Amatoa, F.; Castiglione, A.; Cozzolino, G.; Narduccib, F. A semantic-based methodology for digital forensics analysis. *J. Parallel Distrib. Comput.* **2020**, *138*, 172–177. [Google Scholar] [CrossRef]
27. Watson, S.; Dehghantanha, A. Digital forensics: The missing piece of the Internet of Things promise. *Comput. Fraud. Secur.* **2016**, *2016*, 5–8. [Google Scholar] [CrossRef]
28. Wolfe, H. Evidence analysis. *Comput. Secur.* **2003**, *22*, 289–291. [Google Scholar] [CrossRef]
29. Louw, D. Forensic psychology. In *International Encyclopedia of the Social & Behavioral Sciences*, 2nd ed.; Elsevier: Amsterdam, The Netherlands, 2015; pp. 351–356. [Google Scholar] [CrossRef]
30. Rogers, M. The role of criminal profiling in the computer forensics process. *Comput. Secur.* **2003**, *22*, 292–298. [Google Scholar] [CrossRef]