Research Article
Volume 15:3, 2024

Journal of Forensic Research

ISSN: 2157-7145 Open Access

Advancing Forensic Science: AI and Knowledge Graphs Unlock New Insights

Sundararaj S. Iyengar[1]*, Seyedsina Nabavirazavi[1], Hemant Rathore[2], Yashas Hariprasad[1] and Naveen Kumar Chaudhary[3]

[1]Knight Foundation School of Computing and Information Sciences, Florida International University, 11200 SW 8th St, Miami, Florida, USA
[2]Computer Science & Information Systems Department, BITS Pilani, K K Birla Goa Campus NH 17B, Bypass, Road, Zuarinagar, Sancoale, Goa, India
[3]School of Cyber Security & Digital Forensics, National Forensics Sciences University 6M56+XP8, Police Bhavan Rd, Sector 9, Gandhinagar, Gujarat, India

Abstract

This paper introduces an AI-powered Knowledge Graph for large forensic data investigations, combining machine learning and deep learning to create a sophisticated digital investigation tool. Traditional forensic methods often suffer from a lack of synergy among experts, leading to missed insights and delayed judicial processes. Our Knowledge Graph addresses this by autonomously identifying connections between offenders or victims and analyzing crime event patterns using machine learning-based knowledge signatures and spatial cascadability metrics.

The paper details the creation of a Knowledge Graph from diverse forensic data, highlighting the challenges of data handling and standardization. It showcases the application of this approach in four real-world datasets, demonstrating its effectiveness in forensic reasoning. The results indicate that AI-enabled knowledge graphs can significantly enhance digital investigations. Additionally, the use of spectral analysis for examining real-world interconnections highlights the system's autonomous capabilities. This AI-driven approach promises more efficient digital investigations and could play a crucial role in reducing security breaches in global businesses.

Keywords: Digital forensics • Knowledge graphs • Autonomous forensic systems • Artificial intelligence

Introduction

Every day, the field of forensics produces enormous amounts of data from several investigations throughout the globe [1,2]. Traditionally the forensic investigation process is mostly human-driven as shown in Figure 1 and thus cannot cope with next-generation challenges in the domain. Law enforcement personnel, researchers and scientists are unable to swiftly sort through data in order to find answers to urgent problems while it is in its raw form. New and improved digital forensic investigation tools are essential as forensics moves into the age of Artificial Intelligence, Machine Learning, and Deep Learning. On the other hand, Embedded artificial intelligence will provide enhanced opportunities for near real-time collection and analysis of forensic evidence [3,4]. Super internet growth and advances in rapid communications technologies, coupled with increased bandwidth and the proliferation of telecommunications and computing devices are driving exponential growth in the networks as well as the information transiting these networks and being stored or archived for future use (Figures 1 and 2).

Existing literature suggests that proper mining of digital forensic data has resulted in the identification of the culprit in the past [5,6]. For example, the famous case of Larry J. Thomas vs the State of Indiana used the culprit's Facebook images to corroborate the evidence that led to the conviction. The infamous case of The Craigslist Killer also used Facebook data and emails to produce digital evidence to identify the culprit. The BTK Killer case was solved based on the mining metadata of a Word document that helped to reveal the

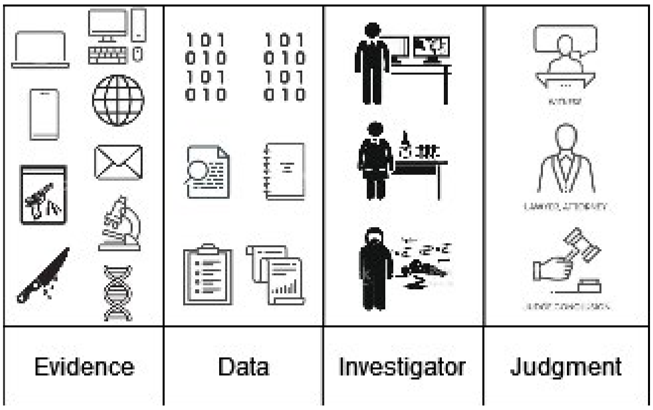

Figure 1. Traditional forensics process.

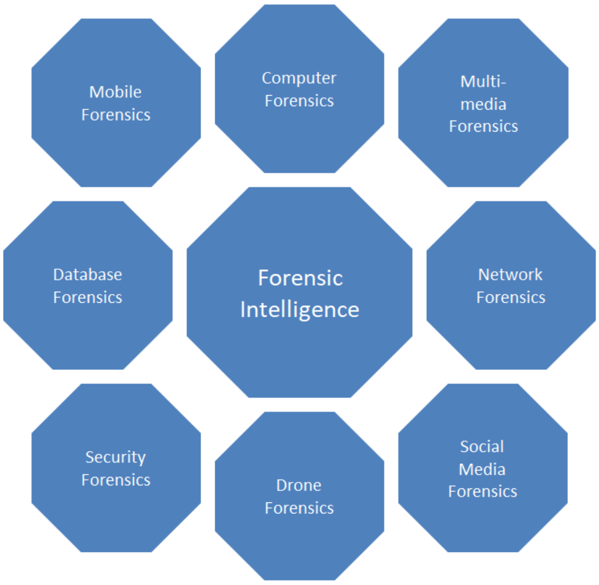

Figure 2. Different forensic domains.

**Address for Correspondence:* *Sundararaj S. Iyengar, Knight Foundation School of Computing and Information Sciences, Florida International University, 11200 SW 8th St, Miami, Florida, USA, E-mail: iyengar@cis.fiu.edu*

Received: 01 June, 2024, Manuscript No. jfr-24-137781; **Editor Assigned:** 03 June, 2024, PreQC No. P-137781; **Reviewed:** 17 June, 2024, QC No. Q-137781; **Revised:** 22 June, 2024, Manuscript No. R-137781; **Published:** 29 June, 2024, DOI: 10.37421/2157-7145.2024.15.615

true identity of the killer.

Today, we are living in a digital age where we are generating tons of digital data that if required can be used for forensic investigation. For example, geotags embedded in files like pictures, videos, etc, and location data from cell towers, Wi-Fi connections, etc. can produce ample forensics evidence. One famous case study is when a Russian soldier named Alexander Sotkin posted selfies on his Instagram account in June 2014. The forensic examination of those images revealed geotag metadata and locations that showed the movement of the Russian army from its military base to eastern Ukraine and back. Similarly, today smart cars collect and store a lot of digital parameters to improve customer experience. A hit-and-run case in 2017 was solved using forensic analysis of this data. The police performed forensic analysis of data from the accused car's navigation system, infotainment system, telematics system, etc. to generate digital evidence that helped them in solving the case.

Currently, criminals are exploring new and faster means of infiltrating and exploiting individual and corporate networks, penetrating new security systems, and violating both the security and privacy of our society. As law enforcement struggles to apprehend and convict these criminals, it is increasingly important for forensic investigators to be equipped with new tools and mechanisms to identify forensic evidence stored in the networks in near real-time. Law enforcement personnel and scholars may now more quickly and effectively solve crimes by creating a knowledge graph a visible, linked network illustrating the links between current data and knowledge entities as shown in Figure 2. Furthermore, in any forensics case, the government or a court of law may make better decisions thanks to these nuanced solutions.

Knowledge signature graphs are networks that demonstrate the relationships between various items by integrating the information [7,8]. They support platforms like Google search, social media websites, streaming media, etc with ease that is used by almost everyone. Knowledge graphs offer a wide range of applications in forensic science due to their capacity to establish complicated and overlapping relationships, such as representing hundreds of nodes in a simulated network [9]. When utilized appropriately, they can provide information on new target security threats, reveal how hackers and network assaults work, or pinpoint the impacts of a particular virus that is introduced into the system.

Intelligent digital forensic data

Knowledge graphs can be created using data from various sources, including crime scene information, forensic lab, and investigative agencies, court case records, criminal databases, databases of software attacks, social media profiles, journal articles, public repositories, third-party tools, and private and experimental data [10,11]. Knowledge graphs should be created with the aim in mind to make the greatest use of the available data. This involves synchronizing and retrieving information on data sets leveraging sophisticated ontologies. Semantic technology can be used to convert unstructured text into structured data, classify it, and extract relationship data. Deeper insights, linkages, and a reduction in complexity will be possible as a result. Knowledge graphs can improve scientific rigor by implementing domain-specific ontologies and employing cross-checked IDs. Stakeholders will have more faith in the conclusions since the automated AI process will be made more explainable and less of a closed black box.

Making sure data is FAIR, i.e., Findable, Accessible, Interoperable, and Reusable is a crucial step in creating datasets for knowledge graphs. The tools and instructions used to query the knowledge network will be significantly harder in the absence of extensive, standardized, comparable data. Data in a knowledge graph may be transferred, precisely defined, and formatted in a way that makes it interoperable. This gives graph models a strong foundation.

Discovering the connections in data

Knowledge graphs can continually ingest new data from specified data sources, making them a dynamic source of knowledge that may be updated in real time or as required. As a result, they can develop based on a semantic network of incoming data. Law enforcement officials can respond to queries like what entities might be targets for an attacker, which ones are being targeted the most, or if an attack might be recycled to hack into another business with a similar network pathway by thoroughly mining data and utilizing latent knowledge (Figure 3).

Knowledge graphs have a wide range of applications in digital forensics and cyber security. Their strength lies in discovering and utilizing connections between data and knowledge entities to create solutions but to fully benefit from this strategy, sound data practices and reliable sources are required. The utilization of data in knowledge graphs can hasten the detection of crimes, produce insights or predictions about how an investigation will turn out, and eventually hasten the capture of offenders and shield them from further attacks.

Proposed framework

Artificial Intelligence techniques like Named Entity Recognition (NER) [12], Natural Language Processing (NLP), and Machine Learning (ML) can be used to recognize, comprehend, and connect data to create a knowledge graph. Knowledge graphs, also known as triples, depict certain connections between data and knowledge items in a form that computers can understand. These triples, which specify particular connections between two entities, are either automatically derived from existing semantics or taken from them. For example, the words bugs, or worms is an error in the programming that is causing a glitch or an unexpected problem for the end user, the named entity recognition can be used to recognize the terms bugs and worms as a programming error and not as an insect. While the submission of evidence to existing large language models (LLMs) like ChatGPT remains unfeasible, these LLMs can function as auxiliary tools for tasks such as the creation of forensic scripts and error validation [13].

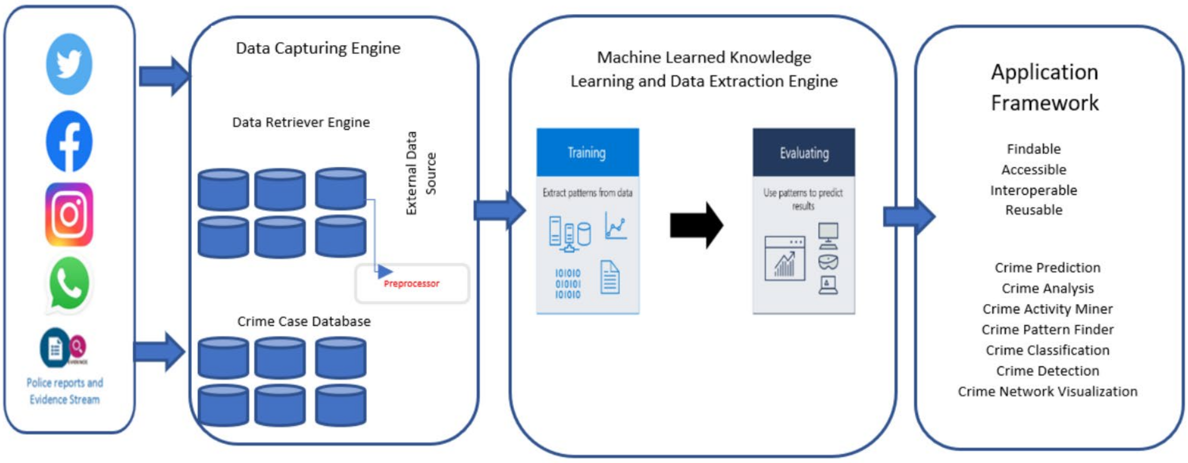

Figure 3. Proposed framework architecture of AI-powered knowledge graph.

Building the knowledge signature graphs

Knowledge graphs may therefore be used to depict, explain, and map complex, overlapping relationships. Law enforcement personnel may uncover pertinent information and insights more quickly and accurately using this rich model since it produces a lot more relevant information than simply keyword-searching books. Any knowledge graph needs well-formatted data that is also pertinent to the application. To be genuinely successful, data must be maintained and sourced wisely. However, without context, research data, study reports, photographs, and other writings frequently lack significance, posing a problem for computers that need a certain amount of data to begin learning. Figure 3 illustrates the proposed knowledge graph-based network architecture, two major categories of data sources are considered in this framework 1) Social media and 2) Police reports and evidence streams.

The data capturing layer contains a data puller engine that inputs real-time data from two major categories through a secured Application Program Interface (API) developed through data crawlers. A different API is implemented on each network media that binds to the data-capturing layer through a set of controllers to govern and screen the data. These APIs pull data based on the pre-configured timers and are also intelligently configured to pull based on the network traffic on social media surfaces. This also includes a separate preprocessing data layer that focuses on removing errors and inconsistency from the data, missing links, and maintaining the integration of data. Holding to valid and quality data is more vital than having a large pool of inefficient data; hence, the machine and deep learning models are plugged into this section. This framework automatically gathers information from various multimedia into the internal structured repository and organizes the data as soon as it rests on disks. This process involves different tasks related to image cropping, audio cleansing, video segmentation, malware prediction, script injections, and extracting logos, weapons, biometric features, and emojis. The framework also processes files of different formats, such as audio, video, image, text, etc. files that are stored locally in internal repositories.

The next layer is knowledge harnessing is performed using machine and deep learning models, where the data are modeled based on AI-powered databases. The training on crime scene database and social media data repositories is conducted using suitable feature extraction. Further, the regression analysis is performed on internal and trained data sets to ensure the quality of the machine and deep learning models. A mapping of crime scene data is collated with the available social media with a trained data set with combinational data patterns. The same is supplied to the prediction model to predict the dissimilar data properties. The knowledge learning layer covers the application of machine and deep learning techniques and reasons databases to analyze the graph data to extract the required information. Here, criminal behavior can be modeled using various machine and deep learning algorithms. Additionally, the system alerts for abnormal behaviors or patterns that can be modeled using various anomaly detection techniques. Thus, it is necessary to train appropriate models to obtain high detection rates and ensure a low false positive rate. The application layer contains a dashboard that allows users to customize different processes related to prediction and crime investigation (Figure 4).

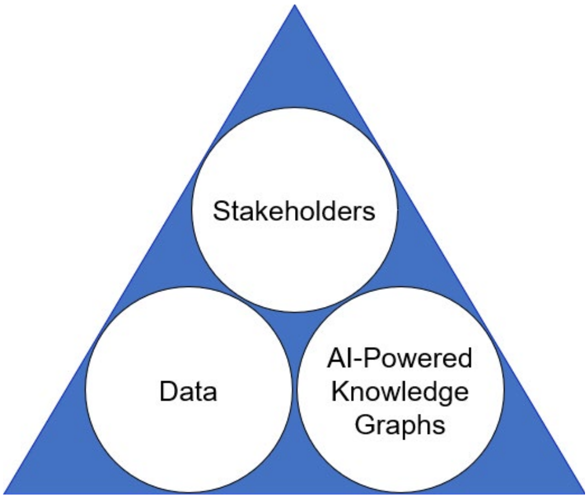

Figure 4. Explainability in AI-powered knowledge graphs.

Methodology

Technically creating a knowledge graph involves extracting relevant entities from data and representing the relationships between them. Our framework includes the following steps.

Entity extraction: The module employs mode-specific tools to recognize entities in the evidence. In textual data, named entities that refer to the key subjects of a piece of text should be extracted. Named entities include names, locations, events, companies, and times. There are various publicly available libraries to employ for extracting such named entities from natural text, including Stanford NER, spaCy, NLTK, and Polyglot [14]. Extracting entities from visual data is achieved by localization techniques that have been evolving over the past decades in the field of computer vision employing neural architectures such as Vision Transformers (ViTs) and Region-based CNNs (R-CNN) [15]. System and network log files also provide useful entities including IP addresses, MAC addresses, hostnames, usernames, timestamps, process IDs, device names, and location identifiers. Several utilities are available for entity extraction from log files such as grep, awk, sed, Microsoft Log Parses, etc.

Relationship extraction: There are two types of relationships within our framework. Either two entities are from the same modality where we record an intra-domain relationship, or they are from different modalities where we record an inter-domain relationship. For example, a company name and the name of the owner may be collected as entities with an intra-domain relationship while the name of the owner has an inter-domain relationship with their image.

Types of relationships: Two types, contextual and evidence-related. Contextual is where George (who was in the incident) lives. Evidence-related is the location of George in the incident. Also inter domain or intra-domain.

Knowledge graph construction: Use a graph database or a graph representation library in Python (e.g., NetworkX) to construct the knowledge graph. Add nodes for each entity and edges for the relationships between them.

Query and visualization: Once the knowledge graph is constructed, you can query it to retrieve specific information or visualize it to gain insights into the relationships between entities.

Explainability

The literature suggests that the AI model's explainability is an essential criterion for its deployment in the real world [16,17]. The AI-powered knowledge graphs will incorporate explainability in data as well as AI models as shown in Figure 4. In data, explainability can be improved with respect to samples, distribution, and features. In the model, case-based reasoning (ProtoDash), feature-based explanation (LIME, SHAP, etc.) can be used to improve explanations. Finally, quantitative metrics like faithfulness, monotonicity, etc., can further improve the explainability of AI-powered knowledge graphs.

Adversarial robustness

The literature on other domains, like image classification [18], object recognition [19], spam detection [20], malware detection [21], etc., suggests that AI models are susceptible to adversarial attacks. The adversary can target the AI-powered knowledge system to reduce its performance as illustrated in Figure 5. The threat modeling of adversarial attacks against AI-powered knowledge systems can be described using adversary's Goal, Knowledge, and Capabilities against the target system. The adversary can design attacks with the GOAL to disrupt the integrity, availability, and privacy of the AI system. The adversary's KNOWLEDGE about the target system can be defined based on information about the following three parameters (1) dataset, (2) feature set, and (3) classification function used to construct the AI system. The white-box scenario assumes that the attacker/adversary has complete knowledge of all three parameters of the target system. In contrast, the black box scenario assumes no knowledge is available about any parameter of the target system. The grey-box scenario assumes that the attacker/adversary has partial information about the target system. The CAPABILITY can be defined based on the adversaries' ability to influence/modify the test data (exploratory influence) or training data (causative influence). Privacy attacks can also be developed for model stealing attacks, model inversion attacks, or membership interference attacks (Figure 5).

Experimental Results and Discussion

Datasets

Our experiments are conducted on four benchmark anonymized datasets. We employ the Stanford Network Analysis Platform (SNAP) [22] real-world dataset collection for our forensics framework. The chosen datasets encompass a spectrum of key information, ranging from social media friendships to cryptocurrency transactions, thereby giving rise to diverse forms of knowledge graphs, including both directed and undirected structures.

email-Eu-core [23]. The dataset captures eight hundred days of email communication between institution members, representing a temporal network.

CollegeMsg [24]. The dataset consists of private messages exchanged within an online social network aliated with the University of California, Irvin, representing users' behavior and interaction.

Stablecoin ERC20 Transactions [25]. The dataset captures cryptocurrency transactions which employ the ERC-20 standard within the Ethereum blockchain. It captures the transactions of leading stablecoins by market capitalization, namely USDT, USDC, DAI, UST, PAX, and additionally, WLUNA.

Gowalla [26]. The dataset contains Gowalla friendship and check-in information. Over six million check-ins are included in the dataset.

Implementation and metrics

In our empirical investigation, we systematically acquired data from forensic datasets to construct a knowledge graph tailored for a specific forensic application. This is achieved through the instantiation of a Data-Capturing class, dedicated to retrieving pertinent information from databases, and a Data-Extraction class, designed to discern and organize the acquired data into the resulting knowledge graph. Finally, a Forensics-Tool module analyzes the extracted graph to identify specific patterns associated with an incident. The Forensics-Tool categorizes the network into multiple node groups, serving as an effective method for identifying outliers. It employs the k-means clustering algorithm for the forensic task. It is also capable of deriving key characteristics of the entire graph, including the highest centrality values.

Results on benchmarks

For clarity, we present a pruned version of the knowledge graphs in the manuscript. The email-Eu-core directed and temporal network has a diameter of seven and high connectivity (Figures 6 and 7).

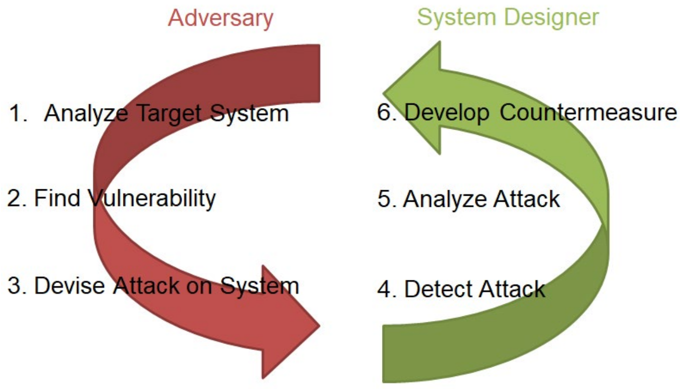

Figure 5. Adversarial robustness in AI-powered knowledge graphs.

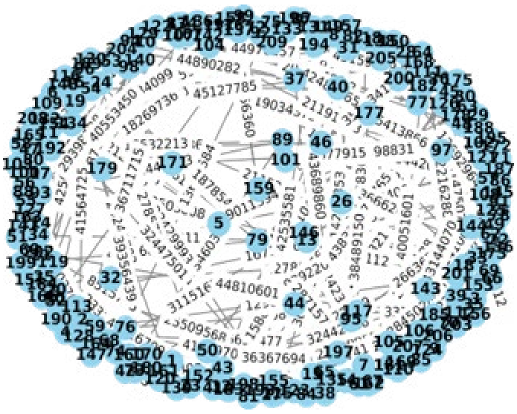

Figure 6. Email-Eu-core knowledge graph. Each edge represents an email communication between two members.

Figure 7. CollegeMsg knowledge graph. Each edge represents a message communication between two members.

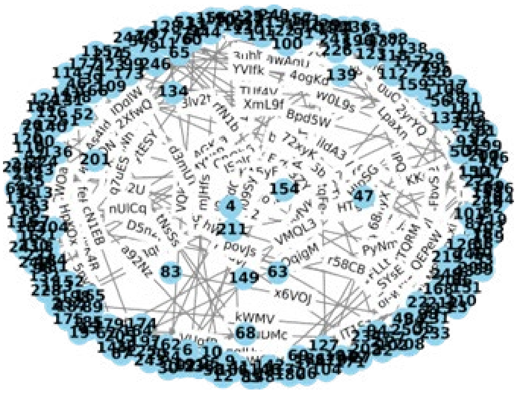

Figure 8. Stablecoin ERC20 knowledge graph. Each edge represents a crypto transaction.

Stablecoin ERC20 is a larger dataset and requires more pruning before representation (Figure 8).

Within the Gowalla dataset context, we delineate two separate graphs: one delineating user friendships Figure 9 and another portraying their approximate geo-locational proximities Figure 10. The Gowalla undirected network has a diameter of fourteen with a higher rate of leaves. The subsequent graphs present the outcomes of clustering obtained through the forensic tool (Figures 9 and 10).

The forensic module indicates that in the email-Eu network Figure 11, node 50 has the highest degree centrality of 0.028, node 60 has the highest eigenvector centrality of 0.66, and node 46 has the maximum betweenness centrality of 0.0003.

Similarly, in the CollegeMsg network Figure 12, node 15 has the highest degree centrality of 0.016, node 27 has the highest eigenvector centrality of

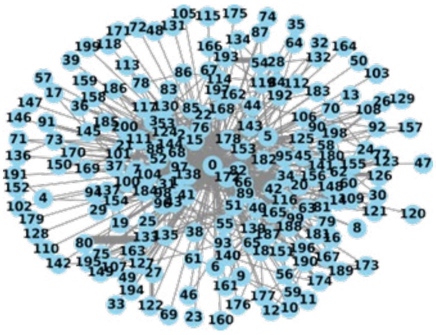

Figure 9. Gowalla friendship knowledge graph.

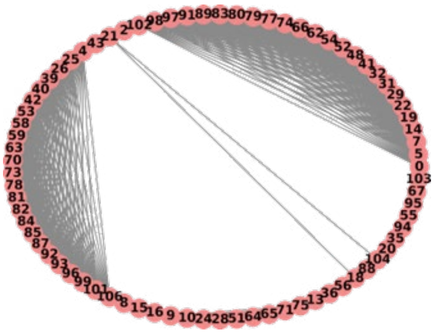

Figure 10. Gowalla geo-location proximity knowledge graph.

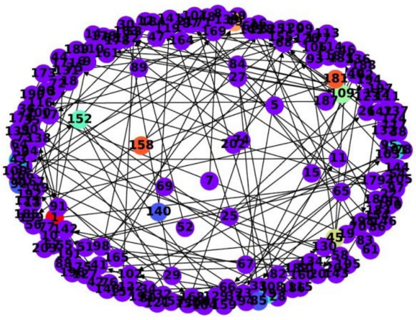

Figure 11. Email-Eu network after clustering where k=10. Nodes {158, 181} are an example of outliers.

0.44, and node 29 has the maximum betweenness centrality of 0.0002.

Within the Stablecoin network Figure 13, node 34 has the pinnacle of degree centrality with a value of 0.02, node 50 has the highest eigenvector centrality at 0.45, and node 34 has the utmost betweenness centrality of 0.0002. Given the uniformity of these metrics across the initial three datasets, our emphasis will pivot toward analyzing the outcomes of clustering for more extensive forensic analysis. In light of the clustering outcomes, identifiable groups of outliers emerge within the datasets. Instances of unexplained outliers, lacking justifications such as organizational hierarchy or internal team structures, warrant thorough investigation in the event of an incident (Figures 11-13).

In addition to employing a general methodology, context-specific measures can prove effective. For instance, after the derivation of a transaction graph from Stablecoin ERC20 and the identification of anomalies, a forensic expert may establish connections between user pseudonyms and online identities to facilitate further investigative procedures [27] (Figures 14 and 15).

Mysteries and challenges

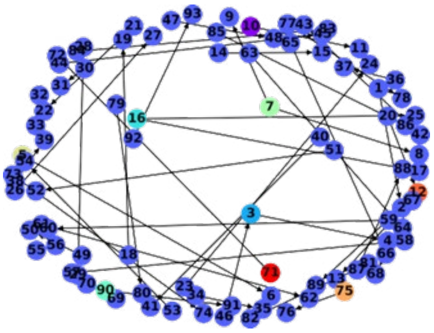

Figure 12. Pruned collegemsg network after clustering where k=10. Nodes 16, 7, and 3 are examples of outliers.

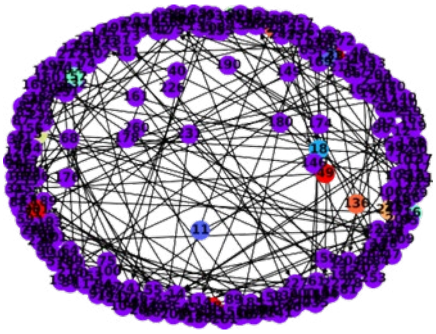

Figure 13. Stablecoin ERC20 network after clustering where k=10. Nodes 18, 136, and 11 are examples of outliers.

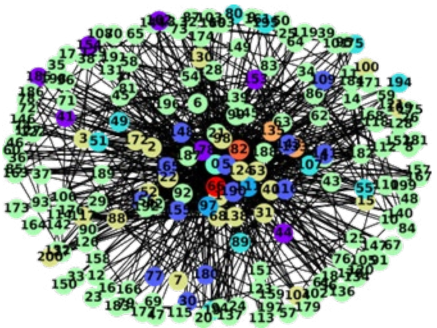

Figure 14. Gowalla friendship network after clustering where k=10. Nodes 82, 135, and 66 are examples of outliers.

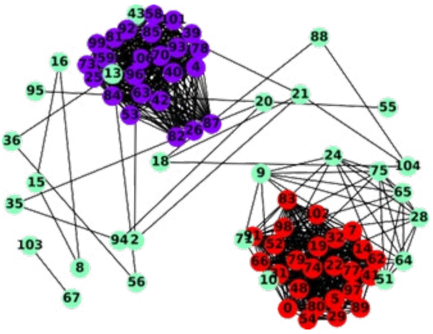

Figure 15. Gowalla geo-location proximity network after clustering where k=3.

The most astonishing problem in digital forensics is the hidden and untraversed challenges that the forensics community has to deal with on a daily basis due to outdated tools and isolated data repositories. Hence, due to the limitations of the tools and data, proper investigation, reasoning, and prediction cannot be performed. Further, the intra-communication of crime databases and investigations is not in practice due to various legal compliances; this is another setback for digital forensics. A small crime consumes a long time for investigation because of the less clarity of the case, as computers, mobiles, and network-based crimes are becoming more sophisticated and complex. A single clue for a case is insufficient to catch the core of crime as services and actions are connected to various flows. The misery of fast-running technology is that it creates headloads by introducing dissimilar technology and data patterns [28]. Specifically, any investigations based out of the cloud are still time-consuming and non-compliant as there are foundational practices to discover data, action, traces, and actors.

While image forensics has obtained the greatest benefit from employing machine learning methods [29], the prime challenges of digital forensics are:

- Lack of knowledge in rising technology
- Development of standards
- Explosion of complexity
- Rise of anti-forensics techniques
- Legitimacy
- Privacy-preserving investigations
- Technology diversity
- Techno-legal complexities
- Lack of specialists and experts

Drawing from the framework and experimental results showcased in the paper, it proposes a groundbreaking paradigm shift in addressing the challenge of extracting solutions from forensic data, presenting a novel approach to unraveling mysteries within this domain.

Conclusion

This paper introduces an innovative framework that utilizes artificial intelligence-powered knowledge graphs to unravel mysteries within forensic data. Of particular significance is its elucidation of a reasoning process applicable to the increasingly expansive and complex landscape of multi-modal forensic data. The document provides a comprehensive framework for the creation of knowledge graphs from forensic data, delving into the intricacies of managing and extracting insights from the burgeoning volume of diverse forensic data. Additionally, the paper thoroughly explores the challenges associated with handling and standardizing various types of forensic data, offering valuable insights into overcoming these obstacles in the field of digital forensics.

Acknowledgment

This research was sponsored by the Army Research Office and the NSF, and was accomplished under Grant Numbers W911NF-21-1-0264 and 2018611 respectively. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Office or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

The authors extend their sincere gratitude to all individuals who contributed through valuable discussions in the early stages of this paper.

References

1. Horsman, Graeme and James R. Lyle. "Dataset construction challenges for digital forensics." *Forensic Sci Int Digit Investig* 38 (2021): 301264.
2. Quick, Darren and Kim-Kwang Raymond Choo. "Impacts of increasing volume of digital forensic data: A survey and future research challenges." *Digit Invest* 11 (2014): 273-294.
3. Hariprasad, Yashas, K. J. Latesh Kumar, L. Suraj, and S. S. Iyengar. "Boundary-based fake face anomaly detection in videos using recurrent neural networks." *In Proceedings of SAI Intelligent Systems Conference*, Cham: Springer International Publishing, (2022). 155-169.
4. Kumar, KJ Latesh, Yashas Hariprasad, K. S. Ramesh, and Naveen Kumar Chaudhary. "AI Powered Correlation Technique to Detect Virtual Machine Attacks in Private Cloud Environment." *In AI Embedded Assurance for Cyber Systems*, Cham: Springer International Publishing, (2023). 183-199.
5. Pouyanfar, Samira, Saad Sadiq, Yilin Yan, and Haiman Tian, et al. "A survey on deep learning: Algorithms, techniques, and applications." *ACM Computing Surveys (CSUR)* 51 (2018): 1-36.
6. Wang, Cliff, Sundararaja S. Iyengar, and Kun Sun, eds. "AI Embedded Assurance for Cyber Systems." Springer Nature, 2024.
7. Gutierrez, Claudio and Juan F. Sequeda. "Knowledge graphs." *Commun ACM* 64 (2021): 96-104.
8. Hogan, Aidan, Eva Blomqvist, Michael Cochez and Claudia d'Amato, et al. "Knowledge graphs." *ACM Comput Surv* 54 (2021): 1-37.
9. Rodrigues, Fillipe Barros, William Ferreira Giozza, Robson de Oliveira Albuquerque and Luis Javier García Villalba. "Natural language processing applied to forensics information extraction with transformers and graph visualization." *IEEE T Comput Soc Sy* (2022).
10. Singaram, Jayakumar, S. S. Iyengar, and Azad M. Madni. "Deep Learning Networks."
11. Shi, Bin, and Sundararaja S. Iyengar. "Mathematical theories of machine learning-Theory and applications." Springer International Publishing, (2020).
12. Silalahi, Swardiantara, Tohari Ahmad and Hudan Studiawan. "Transformer-based named entity recognition on drone flight logs to support forensic investigation." *IEEE Access* 11 (2023): 3257-3274.
13. Scanlon, Mark, Frank Breitinger, Christopher Hargreaves and Jan-Niclas Hilgert, et al. "ChatGPT for digital forensic investigation: The good, the bad, and the unknown." *Forensic Sci Int Digit Investig* 46 (2023): 301609.
14. Vychegzhanin, Sergey and Evgeny Kotelnikov. "Comparison of named entity recognition tools applied to news articles." In 2019 Ivannikov Ispras Open Conference (ISPRAS) IEEE (2019): 72-77.
15. Zou, Zhengxia, Keyan Chen, Zhenwei Shi and Yuhong Guo, et al. "Object detection in 20 years: A survey." Proceedings of the IEEE 111 (2023): 257-276.
16. Balasubramanian, Vineeth N. "Toward explainable deep learning." Commun ACM (2022): 68-69.
17. Rajabi, Enayat and Kobra Etminani. "Knowledge-graph-based explainable AI: A systematic review." *J Inf Sci* (2022): 01655515221112844.
18. Goodfellow, Ian J., Jean Pouget-Abadie, Mehdi Mirza and Bing Xu, et al. "Generative adversarial networks. arXiv 2014." arXiv preprint (2014): arXiv:1406.2661 10.
19. Serban, Alex, Erik Poll and Joost Visser. "Adversarial examples on object recognition: A comprehensive survey." *ACM Comput Surv* 53 (2020): 1-38.
20. Rao, Sanjeev, Anil Kumar Verma and Tarunpreet Bhatia. "A review on social spam detection: Challenges, open issues, and future directions." *ESWA* 186 (2021): 115742.
21. Yan, Senming, Jing Ren, Wei Wang and Limin Sun, et al. "A survey of adversarial attack and defense methods for malware classification in cyber security." *IEEE Commun Surv Tut* 25 (2022): 467-496.
22. Leskovec, Jure and Rok Sosič. "Snap: A general-purpose network analysis and graph-mining library." *TIST* 8 (2016): 1-20.
23. Paranjape, Ashwin, Austin R. Benson and Jure Leskovec. "Motifs in temporal networks." In Proceedings of the tenth ACM international conference on web search and data mining (2017): 601-610.
24. Panzarasa, Pietro, Tore Opsahl and Kathleen M. Carley. "Patterns and dynamics of users' behavior and interaction: Network analysis of an online community *JASIST* 60 (2009): 911-932.

25. Shamsi, Kiarash, Friedhelm Victor, Murat Kantarcioglu and Yulia Gel, et al. "Chartalist: Labeled graph datasets for utxo and account-based blockchains." *Adv Neural Inf Process Syst* 35 (2022): 34926-34939.

26. Cho, Eunjoon, Seth A. Myers and Jure Leskovec. "Friendship and mobility: user movement in location-based social networks." In Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining (2011): 1082-1090.

27. Pocher, Nadia, Mirko Zichichi, Fabio Merizzi and Muhammad Zohaib Shafiq, et al. "Detecting anomalous cryptocurrency transactions: An AML/CFT application of machine learning-based forensics." *EM* 33 (2023): 37.

28. Miller, Jerry, Lawrence Egharevba, Yashas Hariprasad, and Kumar KJ Latesh, et al. "Cyber Security Attack Detection Framework for DODAG Control Message Flooding in an IoT Network." In *International Conference on Information Security, Privacy and Digital Forensics*, Singapore: Springer Nature Singapore, (2022). 213-230.

29. Nayerifard, Tahereh, Haleh Amintoosi, Abbas Ghaemi Bafghi and Ali Dehghantanha. "Machine learning in digital forensics: A systematic literature review." arXiv preprint (2023): arXiv:2306.04965

How to cite this article: Iyengar, Sundararaj S., Seyedsina Nabavirazavi, Hemant Rathore and Yashas Hariprasad, et al. “Advancing Forensic Science: AI and Knowledge Graphs Unlock New Insights.” *J Forensic Res* 15 (2024): 615.