Blockchain based Digital Forensics Investigation Framework in the Internet of Things and Social Systems

Shancang Li, *Senior Member, IEEE*, Tao Qin, and Geyong Min

***Abstract*—The decentralised nature of blockchain technologies can well match the needs of integrity and provenances of evidences collecting in digital forensics across jurisdictional borders. In this work, a novel blockchain based digital forensics investigation framework in the Internet of Things (IoT) and social systems environment is proposed, which can provide proof of existence and privacy preservation for evidence items examination. To implement such features, we present a block enabled forensics framework for IoT, namely IoT forensic chain (IoTFC), which can offer forensic investigation with good authenticity, immutability, traceability, resilience, and distributed trust between evidential entitles as well as examiners. The IoTFC can deliver a gurantee of traceability and track provenance of evidence items. Details of evidence identification, preservation, analysis, and presentation will be recorded in chains of block. The IoTFC can increase trust of both evidence items and examiners by providing transparency of the audit train. The use case demonstrated the effectiveness of proposed method.**

***Index Terms*—Digital forensics, provenance, evidence items, Internet of Things**

I. Introduction

Over the years, the emerging technologies such as the social networks, Internet of Things (IoT), the fifth generation of communication (5G), the decentralized blockchain technologies, *etc.* have become an indispensable part of modern life [1], [2], [3], [4]. New technologies make our lives easier, faster, and more fun by creating amazing tools, devices, resources, and putting the most useful information at fingertips. However, new technologies have made it increasingly easier for criminals to conduct their activities in IoT environment, where a huge number of devices are interconnected to the Internet [5], [6]. It is reported that these new technologies make cybercrimes much more difficult to detect and prosecute than traditional crimes [7], [8]. In forensic investigation, digital evidence plays an increasingly important role that is expected to bridge persons with criminal activities [9]. As a result, it is very important to guarantee the continuous integrity, traceability, and auditability of evidences in IoT environment.

The existing digital forensics are facing new challenges in the context of cyber physical systems, including inaccessibility of data from different sources, data provenances in multiple locations, evidence transparency and traceability, data analysis of large volumes of dataset, *etc.* In the past few years, many research efforts have focused on cloud based forensic analysis [10], evidneces modelling [11], [12], [13] and assisting the law enforcement community. In the IoT environment, digital forensics are facing a number of challenges, including: (1) Defining framework for digital forensics that can face the new challenges in new environment; (2) Guaranteeing the reliability, availability, recovery of dynamic digital evidence in complicated environment; (3) Privacy concerns and new privacy laws, such as the compliances of the General Data Protection Regulation (GDPR); and more. New research in digital forensics must address these above challenges in the procedural, social, and legal field [14], [15], [16].

The blockchain technology is a distributed ledger system, which can store linked records in the form of a decentralized database in the peer-peer network. The data are stored in timestamped blocks which are linked in a chain, creating immutable, publicly visible and validated audit trail by a consensus-based proof of trust [17]. The blockchain gains its secure, immutable nature of cryptographic hash link between blocks and transactions, meanwhile, it can provide well immutability, traceability, transparency, auditability and accountability. The blockchain has been successfully applied in financial services, supply chain, energy industries, pharmaceutical, *etc.* In forensic applications, the blockchain technology is promising to address above challenges. The advantage of blockchain technologies in digital forensics is the examiner can provide self-verification for digital evidences, which can make use of hash function to effectively establish verifiable evidence chain. The blockchain makes use of cryptography to guarantee the immutability, transparency, and distributed trust within the case examination.

In this paper, a blockchain based IoT forensic framework (IoT Forensic-Chain, IoTFC) for forensic investigating in the IoT environment is proposed, which provides full data provenance architecture and assurance of examination operations. Meanwhile, it can also provide security privacy and availability together with the transparency, traceability, trust between evidence/item and investigators, and continuous integrity of each evidence item. In the following sections, detailed IoT forensic analysis procedures of recording all examination operations in blockchain networks are addressed. The evidences with its provenance data are hashed into a Merkel tree and written into block. The examination operations are also formatted

Dr. Shancang Li is with Department of Computer Science and Creative Technologies, University of the West of England, Bristol BS16 1QY, U.K. (email: Shancang.Li@uwe.ac.uk).

Prof. Tao Qin is with Department of Computer Science, Xi'an Jiaotong University, Xi'an 710049, China. (email: Qin.Tao@xjtu.edu.cn).

Prof. G. Min is with the Department of Computer Science, College of Engineering, Mathematics, and Physical Sciences, University of Exeter, Exeter, EX4 4QF, U.K. (e-mail: g.min@exeter.ac.uk).

into transactional evidences are liked with related evidence items using Merkle tree. The main aim of this work to extract as much as possible potential digital evidences and reduce investigation costs in IoT environment.

The rest of the paper is organized as follows: in Section II, a comprehensive review for the resent research on IoT forensic analysis is provided, and a blockchain enabled IoT forensic-chain architecture is proposed in Section III. In Section IV, IoT blockchain forensic applications is provided and a use case is provided; Section V discusses the research challenges and trends and concludes the paper.

II. Related Works and New Challenges in Digital Forensics

This section briefly overviews previous works related to digital forensic investigation in complex digital environment and the use of blockchain in digital forensics.

A. Related Works

In the past few years, lots of research efforts have been conducted in IoT forensics [18], [19], [20], including digital evidences identification, collection, storage, analysis, and distribution in IoT environments [19], which is very different with the existing computer forensics. The IoT systems contain many smart devices, heterogeneous networks, and diverse applications, where huge volumes of data and heterogeneous technologies create new challenges for forensic investigation [21], [22], [23]. Since 2017, the emerging blockchain technology have been applied in digital forensics to document evidence items, interaction actions, and preserving evidence in the blockchain [24], [25], [26], [27], [28], [29]. The blockchain enabled forensic investigation also presents promises in tracing of criminals and helping anticipate unauthorized actions in cyber environment [30]. The RFC 2337 provides a guide for evidence collection and archiving in Internet environment [31]. The NIST SP 800-86 [32] introduces digital investigation analysis techniques, strategies for reducing the amount of overhead.

Many forensic investigation methods and analysis models have been proposed forensic investigators and practitioners based on their expertise and experiences [23], [30]. However, currently there no international standards available that formalised these developed forensic investigation processes. Specifically, in the complex digital environments, like the Internet of Things (IoT), cloud computing, and the networked digital cyber physical environment, many challenges are facing by the existing forensic investigation methods. Cebe *et al.* developed a lightweight application objected blockchain framework: Block4Forensic [24], which integrated digital forensic processes and data privacy together and can provide efficient vehicle related digital investigation.

In [33], Zhang *et al.* proposed a provenance process model for the digital investigation using blockchian in cloud environment, which aimed at enhance the interaction trust between stakeholders in cloud forensics. Al-Nemrat *et al.* in [34] investigated the possibility to introduce blockchain technologies in the investigation of financial fraud in e-governance, and the results shows that the blockchain technologies can effectively financial fraud related online product reviews. The blockchain technologies can ensure integrity, trust, immutability and authenticity in untrusted software development. In [35], blockchain is used to provide the auditability, traceability in software development and a role-based access control mechanism for unauthorized data accesses is developed.

In [36], Hossain *et al.* proposed a forensic investigation framework based on the blockchain, which aimed at detecting criminal incidents in the Internet of Things (IoT) environment and collecting interactions from different entities in IoT. The proposed framework can well model the interaction transactions, but it is inefficient in data collection and data analysis in large scare IoT systems. Lone *et al.* proposed a digital forensic chain based on the popular blockchain platform Ethereum [37]. The proposed forensic chain model was implemented over `Ethereum`, which can provide integrity, transparency, authenticity for data collected from multiple sources. Lots fo research efforts have been done on the digital investigation in heterogeneous environment [23], lightweight security solutions over IoT devices [29], digital witness [38], and more.

It is clear that the latest digital forensic analysis and research works are falling into two categories: (1) focusing on assisting the law enforcement community; (2) focusing on specific forensics applications. This work aims at developing a blockchain based digital forensic framework that can be used in complex cyber environment (such as IoT, cyber physical systems, *etc.*) and a use case will be provided to demonstrate the effectiveness of proposed method.

B. Digital Forensics Challenges in IoT Environment

In digital forensics, hash function is widely applied to keep the digital integrity and repeatability by generating a digital fingerprint (hash digest) for a digital asset to prevent changing. However, in existing digital forensic applications, it is only used guarantee the integrity of whole disk drive or data validation, *e.g.*, the EnCase imager uses both MD5 and SHA1 to guarantee the integrity of the image, and FTK Imager computes the acquisition hash of the imaged data when the acquisition is finished. A big concern is that the hash verification/validation is only for the image files or some specific files, but not for examination events, or each evidence items. The existing DF solutions significantly rely on the experiences of the investigators [14], [18].

Here this work summarised the challenges in existing digital forensics investigation as follows:

1) Trustworthy: Trusted insider threats to evidence in the IoT environment, how to improve the trustworthy of evidence item in digital forensics.

2) Integrity: Continuous integrity check for evidence items and examination events in digital investigation. In traditional investigation, no support provided for forensics activities/events between evidence items and examiners/tools and/or data or images/objectives.

3) Improved Provenance: . In IoT environment, the above hash functionality is expected to provide hash validation for all evidence pieces, findings, and all behaviours in examination by creating a hash tree.

A hash tree can be created by repeatedly hashing transactional evidences or its hash value until aggregate into a single root hash, in this work HE denotes the hash value of an evidence, and

$$HE_1 = Hash(TransanctionalEvidence\#1) \quad (1)$$

$$HE_2 = Hash(TransanctionalEvidence\#2) \quad (2)$$

$$H_{12} = Hash(HE_1|HE_2) \quad (3)$$

$$H_{root} = Hash(H_{12}|\cdots) \quad (4)$$

4) Scalability: In a hash tree, a parent node is able to support up to 1000 children-nodes, in digital forensics it means it can support up to 1000 events/activites/evidence items. In IoT environment, a hash tree is capable of up to 10^{3n} hash digests (n is the deep level of a hash tree) and can supports large number of evidence items/events [32].

5) Availability and resiliency: Each node in blockchain has a complete copy of the whole hash tree, which is guaranteed to be accurate. This property makes it extremely resilient store digital evidence data or events in forensic investigation. Once an evidence item is identified and written to a blockchain, an examiner can have a very high degree of confidence that the evidence item will be accessible in question.

To address the above challenges, in next section this paper proposed a blockchain enabled digital forensics framework for the IoT, named as IoT forensics chain (IoTFC).

III. Blockchain enabled Digital Forensics Investigation (IoTFC)

The blockchain technology can offer forensic applications with substantial benefits for the whole procedure of digital forensics investigation procedures, including the data collection, preserving, evidence validating, data analysis, and the presentation of the finding. Specifically, the blockchain can improve the transparency in each individual stage, *e.g.*, it can assistant examiner to accurately identify the data sources in the early investigation stage, reduce the data storage, and improve transactional analysis efficiency, and subsequently can reduce the costs of the investigation.

A. Motivation and Objectives

The proposed IoTFC mainly achieves the following objectives:

1) Comprehensive view of evidence items: the decentralized ledger system can provide a comprehensive view of evidence items back to their evidential sources or links to related evidence items. This will be very helpful in many investigation scenarios when a large number of evidence sources and activities are involved. In IoTFC, the blockchain is used to provide distributed trust to all participants in forensic investigation.

2) Continuous integrity: The continuous integrity, value and/or ownership of specific evidence items is still a challenge in digital investigation. Many cases are caused by the data breaches and a large number IoT devices are interconnected. How to ensure the integrity of these evidences is a basic object of IoTFC. In many scenarios, the trusted insider threats are increasing, and key evidence information were lost or compromised due to the unstable evidence systems. The cryptographic hash functions (such as SHA1, SHA256, *etc.*) are widely used in forensics imaging process aimed at the integrity of specific evidence items, however, for the whole evidence chain, current a continuous integrity check or validation mechanism is missing.

3) Immutability and Auditability: The nature of the blockchain technology can offer digital forensics immutability and auditability, which are key features required in digital forensic chain of evidence.

4) Tamper-proof Environment: Evidence items are collected and then written to the blockchain network, which guarantees the full provenance of each evidence item. All evidence items on the blockchain are shared among the participants. The IoTFC establish a public timestamped log for all examiners on the IoT without the presence of a trusted third part. All evidences items are chained cannot be tempered.

5) Full Provenance: Report of evidence items may have significant implications for criminal justice system [5], providing complete provenance of each evidence item is very important in IoTFC. This should include the full provenance of the item. In forensic investigation, an examiner should provide exact location for each evidence item inters of their full provenance and an independent investigator could locate that evidence. For example, in a windows xp based examination, the examiner should be able to provide logic and physical sector (LS/PS) for all evidence items. In some case, for large files (such as `pagefile.sys` in Windows XP), it is useful to provide the file offset position of the evidence item.

6) Traceability: In many applications, the traceability that offered by blockchain is criticised as a potential privacy issue and encryption solutions have been applied to protection them. However, the blockchain in IoTFC can monitor glitches and provide nice traceability from the scene-to-court along the evidence chain, which is able to restrict the access to all recorded information (*i.e.*, evidence items, examiners, timestamps, tools, *etc.*) in blockchain.

B. Data Acquisition

In IoT environment, the overwhelming majority of data is captured digitally at source, where the evidence will be in the form of digital assets which could be collected from sensors, devices, cloud storage, and at sources. In the context of criminal evidence, it is difficult to restrict access to a digital asset. Fingerprinting digital evidence is a way to generate a digital fingerprint of each piece of digital evidence. The hash algorithms are widely used to generates digital fingerprint, which is unique to the digital evidence and the nature of hashing functions means that even the most minute alteration of the underlying digital evidence completely changes its hash

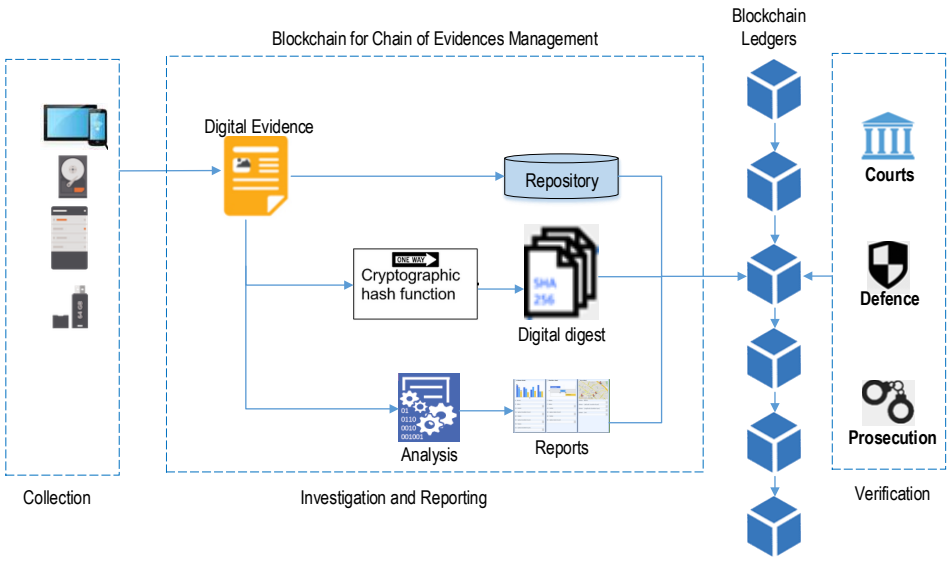

Fig. 1. Blockchain for Chain of Evidence Management in IoT (IoTFC)

digest. To narrow the source devices, this work uses our proposed features based devices fingerprinting methods [39] to identify and fingerprint the devices involved in the case, by doing this, the proposed method does not have to acquire data from all devices in the IoT system, but only focus on the devices that related to the case. The basic procedures include:

- Blockchain can make the data acquisition and validation more accurate and informative by integrate the transactional evidences and addition information;
- For each transactional evidence item, its provenance as well as all related examining events can be traced back to it origination;
- The IoTFC uses blockchain to build a close-loop system that provides significant forensic analysis benefits in an efficient and economic way.

C. Forensic-Chain Framework

IoTFC is a blockchain based forensic solution for digital investigation forensic chain of custody, as shown in Figure 1, which allows the system to create a distributed ledger for recording and storing transactional evidences (examining events/findings, and additional information). These transactional evidences will be shared by all authorized participants via the blockchain network. The cryptographic nature of blockchain guarantees the immutability, timestamping, resilience, traceability, and distributed trust of evidences. The framework consists following critical components:

1) Users and IoT Devices: The users include the users, owners, or examiner that related in this investigation. The devices in this framework include all devices, sensors, or IoT infrastructures involved in the case, which can be identified using our developed feature based device identification [40].

2) Merkle Tree: As discussed above, a Merkle tree is actually a hash tree that allows for efficient and secure verification of transactional evidences in the investigation. It can summarise all the transactional evidences, examination addition information in a block by producing a digital signature for the entire set of items, thereby enabling a user to verify whether or not a transaction is included in a block. Figure 2 shows an example of Merkle trees of nodes, in which the a `Transactional evidence` could be a file, folder, memory, *etc*.

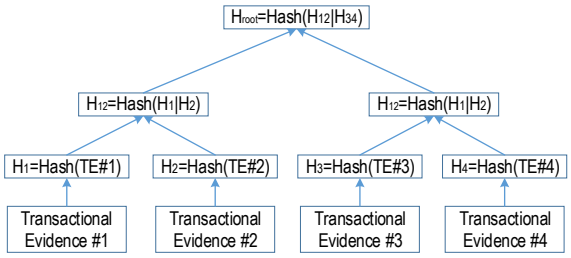

Fig. 2. Rough sketch of the structure of a Merkle Tree

In the IoT forensic context, the blockchain's capability in combination with cryptographic hashing and encryption can fingerprinting transactional evidence items and examination events, which is naturally tamper-proof and secure.

- The evidence items that could be encrypted and can only be accessed by authorized parties on the blockchain but would simultaneously record the timestamps, date, full provenances, *etc*. All this would be completed automatically through smart contract.
- A blockchain browser is used to view the evidence blockchain, will more specific restrictions are defined according to the analysis requirements.

Figure 3 shows an example of Merkle tree in the IoTFC, in which the H_{root} is the hash root `Merkle tree`, H_{12} is the hash of concatenation of hash of two transactional evidence

items #1 and #2. In IoTFC, a digital forensic workstation keeps the IoTFC and it can be easily verified by other nodes or itself. All participating parties in IoTFC are capable of quickly verifying the hash values. However, when failure happens, a distributed consensus is applied in IoTFC. in this work, major voting is used to guarantee the uniqueness of evidential blocks.

3) Block: In the blockchain network of IoTFC, evidences item can be verified based on its fingerprinting. In each block, the block header contains follow attributes: *pre block hash, version, nonce, timestamp, block state*, and `Merkle root`, as shown in Figure 3. The `Prev.Hash` represents the hash value of the block header of pre-block and a nonce. The transactional evidence item represents the evidence item record and it is hashed into a Merkle tree.

4) Smart Contract: Smart contract, also called blockchain contract, is digitalized contract that is executable for a computer. The smart contract are usually stored stored in blockchain network and supervised by the blockchain network nodes. It can help user automatically exchange information, data, business process without the need of middleman. Smart contracts can run, validate, and make decision automatically in the decentralized ledger in a certain security and immutability way.

The smart contract can be easily implemented on the a blockchain platform, such as Ethereum, *etc.* It has been widely used in financial service, healthcare, insurance, e-government, supply chain, *etc.* Similarly, the smart contract can benefit the digital forensics investigation from following aspects:

- Autonomy, it can define the conditions to find related evidences item in an automatic way;
- Trust, the evidence item can be encrypted on a shared ledger;
- Safety, the items can be cryptographically encrypted;
- Speed, the smart contracts can significantly reduce the examining time than manually process.
- Saving, smart contracts can save the cost without paying for middlemen, such as notary, witness, etc.
- Accuracy, the automated smart contract runs in a faster, accurate, and cheaper way.

D. Evidences Grading in IoTFC

In IoTFC, the evidence items can be defined in layers according to their relationships to the case, attributes, and how easy it can be find, in this work evidences are categoried into five grades:

- g_1. Easy to identify, such as plan text in files, unencrpted image, QR *etc.*;
- g_2. Some deliberate attempt at hiding, e.g., renaming of extension, *etc.*;
- g_3. Hard to identify, e.g., plain text held other than in files system, volume slack, *etc.*;
- g_4. Difficult to identify, i.e., encrypted data in a file, password protexted xls file, *etc.*;
- and g_5. Very difficult to identify, such as encrypted data held other than in the files system, steganography, *etc.*

E. Evidence Item Bookmarking and Blockchaining

In forensics examination, a bookmark is a group of files referencing in the cases. An examiner can create as many bookmarks as needed in a case. It provides additional analysis features, includes `hashing, job Options, indexing/tools, miscellaneous, etc.` The bookmarks can also assist carving the data by identifying file headers and footers in mainly unallocated clusters. The bookmarks can enable intuitive forensics activity retrieves packet data and ingests other contents, which is driven by searching, session reconstruction, and forensics intelligence to help security incident investigations.

In existing examination tools, such as `Autopsy, FTK, EnCase`, *etc.*, the search results that related to the investigation can be bookmarked for deeper inspections and final determination. The bookmarks can fine-tune the inspection from following aspects:

- Inspect each bookmarked evidence items through the visualization and analysis tools
- Attach case notes to the bookmarked documents/items and make final decisions on each items about its relevance to the case.
- If a record is not relevant, remove the bookmark.

Evidence items, examination event/actions, and additional information (`e.g., examiner, tools, workstation, timestamps, etc.`) are formatted as `transactional evidence` shared by all participating parties over the blockchain network, where the IoTFC makes use of cryptography for protecting these transactional evidences. Smart contracts are designed to create/record transactional evidences based on examination details, like address to whom evidence is transferred to, current state of evidence, permission level, data and time, *etc.* Further any subsequent access to digital evidence also gets recorded securely on blockchain by smart contracts triggered by corresponding forensic investigation.

IV. IoT Blockchain Forensics Applications

In the IoTFC, the links between each entities, such as *evidence item, devices, users, social system account etc.*, can be easily identified using the Merkle tree. To guarantee the integrity and auditability of digital evidences are very important due to it moves along different levels of hierarchy in chain of investigation. Basically, the IoTFC can enhance existing digital forensic investigation from following three ways:

- Use the smart contract make some evidence analysis be done automatically, such as file signature analysis, email analysis, etc., to improve investigation efficiency and reduce the data exchanges between parts;
- Improve the transparency of investigation and provide better auditability;
- Reduce examination costs and resource uses;
- Establish connection with trusted third parties.

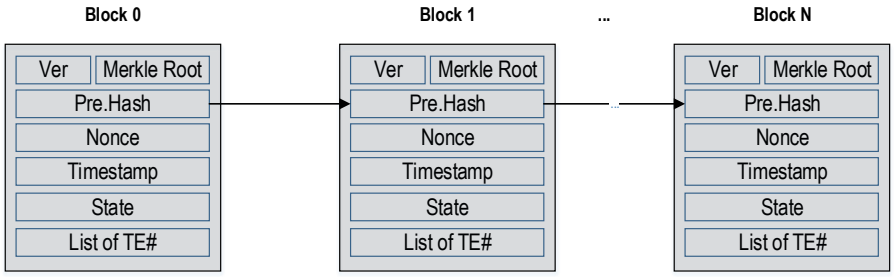

Fig. 3. Chained blocks

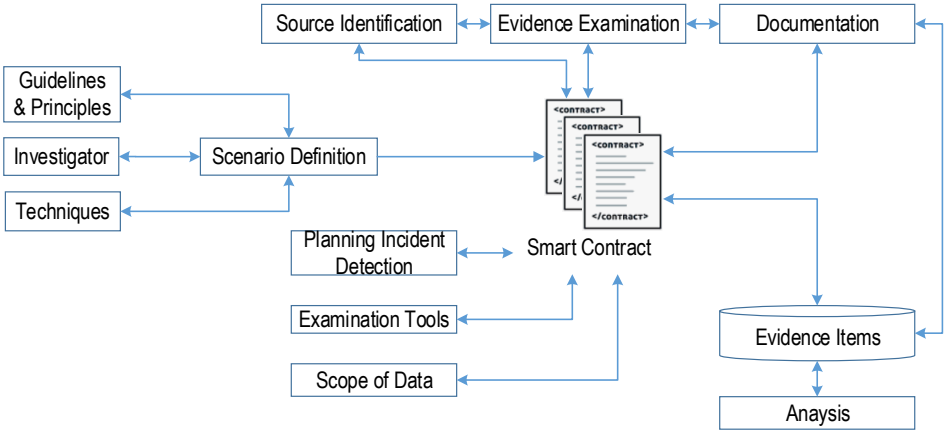

Fig. 4. Evidence identification as an application of smart contract

Fig. 5. JSON Script for Evidence Blocks

A. IoTFC Use Case

This subsection will introduce the use of blockchain in processing digital forensic evidences. Digital evidences can be presented in the form of digital assets in all stages of digital forensic investigation. In this work, we employ the proposed method for a case investigation of an illegal home grower, in which *Ric* and *Shaw* are in the illegal home growing/selling of bad things. Frightened by the government regulations, *Ric* and *Shaw* became more security consious how they communicate and save files. Investigators imaged the computher that *Ric* and *Shaw* are using. After using the Encase 7.5 and Autopsy Media Viewer, a number of key independent evidence items are identified, including

- JPG images masquerading as a .docx file
- JPG images of bad thing (*e.g. skittles.jpg, etc.*)
- Steganography software found on the computer
- Pidgin IM client, chat log conversation file
- Outlook email log files, password protected zip files
- Other suspicious files (*e.g., pwd.txt, etc.*)

As shown in Figure 6, based on the Encase analysis, a stegged JPG image is identified that include key information (sales document). IoTFC shows good performance for organising the evidence items and establishing the links between them. For each evidence item, following informations are identified to generate a transactional evidence (TE):

1) Contemoraneous head: `Examiner, Exam commenced, Software used, version and licensing`;
2) Event: `Action, Done?, Time, Notes, Screen capture`.
3) Evidence item: `No., Description of item, significane to case, full provenance, method discovery`.

Using the IoTFC framework, all case related information can be easily chainned in the blockchain system, such as evidence items, examination event/actions, and additional information (`e.g., examiner, tools, workstation, timestamps, etc.`) are formatted as `transactional evidence` and can be accessed by all participants in the blockchain network. Smart contracts are designed to automatically create/record transactional evidences based on examination details, including source devices, ownership, states, users, or other examination related additional information. By using emerging artificial intelligence (AI) technology, the smart contract can learn rules or knowledge from past cases and create new business logic to improve the investigation efficiency.

As an example, in the investigation, a steged file *skittles.jpg* was identified and all inforation and actions related this file are formed into a transactional evidence $TE\#a1$, which will be written into the blockchain. In fact, further investication is needed for this file. With `JP Hide and Seek Steg` software we extraced an excel file named `output.xlsx` using a password *"Nwkbvceg"* hidden in `pwd.txt`. In this investigation, all information and actions were written into $TE\#b1$ which linked with $TE\#a1$. In fact, the `output.xlsx` recoded the profit sales and details of deals, which proved that *Shaw* was selling illegal stuff and can be presented on the court. All investigation findings and actions can be written into the blockchain and we summarised the procedures as follows.

1) Evidence Identification and Acquisition: this stage involves following four main steps:

Step 1. Identification of digital evidence. The proposed IoTFC uses a one-way HASH function (SHA1) for identifying and fingerprinting digital evidence. If more than one version of digital assets were found, each claiming to be definitive and a digital fingerprint for each digital evidence will be generated, the contents and examination events will be defined as `transactional evidence records`;

Step 2. Together with additional information and timestamps, the fingerprinted records will be written into block of evidence and then append onto the end of the blockchain;

Step 3. In the peer-peer blockchain network, each participant will hold a complete copy of the evidence blockchain. Once an evidence block is written onto the blockchain, each participant can have a very high degree of confidence that the information will be accessible and trace back. Provenance of each evidence item will be guaranteed with a very high degree.

For example, if an evidence item might contain multiple pieces from different sources, each piece and its source will be fingerprinted with hash function to forms transactional evidence item in blockchain. Similarly, entirety of the full evidence chains will be formed in blockchain. When transactional evidences need to be transfer from one party to another, digital signed new records will be created and appended into the blockchain.

2) Analysis: in this stage, the smart contract will be used to create analysis results. Possibly, more interface to intelligent, *EnScript of EnCase, LogRhythm* and more will be provided to use the analysis tools in forensic area, Figure 5 shows an example of smart contract based evidence item analysis. For network events related analysis, more interfaces are provided, such as intrusion analysis, log file analysis, *etc.*

3) Presentation: this stage will be based on the findings in analysis stage, as mentioned above, all evidence can be easily traced back to its originality. All report, or presentation will be based on the blockchain and be appending to the blockchain.

The IoTFC framework well supports the collaboration from different departments. Collaboration between law enforcement, government, and industry will also be considered in building the evidence blockchain. The IoTFC can provide quickly each investigator some special tools, provenances of item, and its origination. As shown in Figure 6, in first stage, all data are imaged and all acquisition related information are written into blockchain; In identification stage, an suspicious image file is located in the acquisition and all identification events/findings in this staged are also written into blockchain; In analysis stage, OpenSteg is used to extract a steged text file, both of the image file and text file are fingerprinted using hash function and all analysis events are recorded in a block; In presentation stage, all findings, report, and related events/behaviors are written into the blockchain. It can be found that all information such as original files, findings, examining events, together with the additional information (such as examiners, examination tools, platforms, *etc.*) are

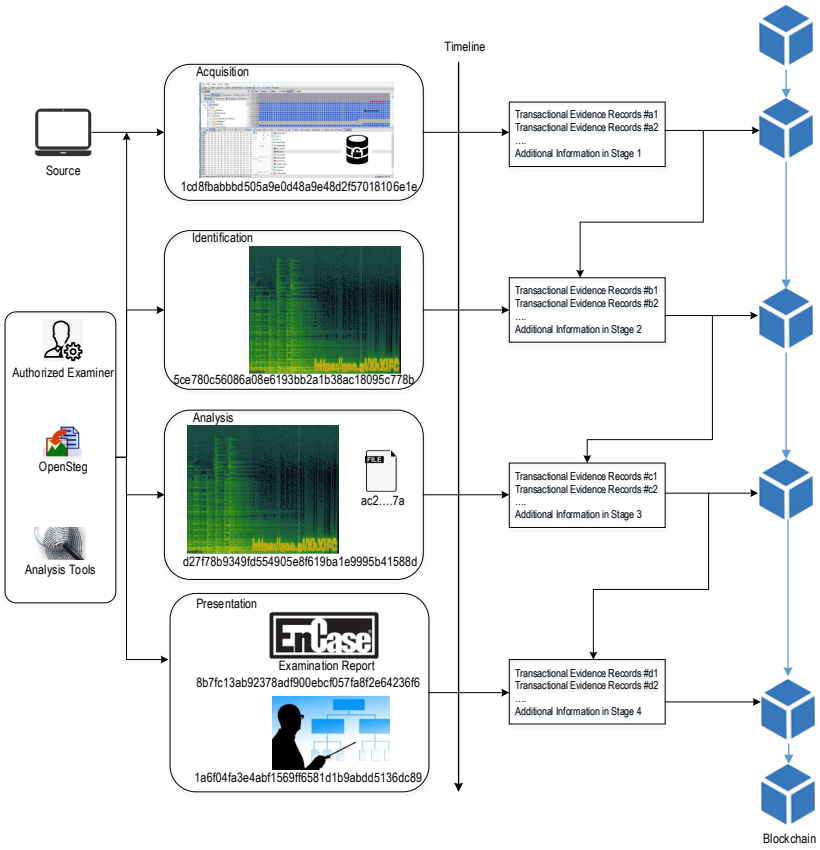

Fig. 6. Use case: Steganography based forensic analysis in IoTFC

fingerprinted and recorded in the blockchain. The IoTFC is an effective digital forensic framework that can provide nice properties: immutability, timestamping, resilience, transparency, and distributed trust.

V. Discussion

A. Self-validation in IoTFC

In IoTFC environment with signature tokens on each evidence item, the examiner could simply conduct hashset comparisons to find `well-defined`, `bad`, `not-sure` or `suspicious` files for further examination. This can speed up the investigation and incident response. As a forensic ready environment, the IoTFC can be applied in IoT environment. An examiner or maintainer can response for the remote incident through hashed and timestamped phtotos, documents, ease of time-line analysis, and IoT forensics artefact storage with flawless chain of custody procedure.

The proposed IoTFC can significantly reduce the processing time in imaging-hash procedure, which can significantly reduce examine-time and provide accurate and quick response for eradication and remediation.

B. Bottlenecks of Blockchain

A permissionless blockchain stores data on a global ledger, which is validated by many unrelated participants, or nodes, that are financially motivated to keep one true version. The nature of immutability of blockchain cryptographically guarantees the transactional evidences in IoTFC can never be replaced or reversed. However, there is always the chance that one entity gains a 51% majority of computing power and thus gets to make the rules but this is difficult/expensive to achieve.

C. IoTFC in Cyber Crimes

Cyber threats are dramatically on the rise in IoT, it is not just data ex-filtration, but data integrity is a growing concern. Cyber forensics is maturing but more works need to be done. Hashing is improving with timestamps and blockchaining. Blockchain based digital forensic chain of custody has great potential to bring substantial benefits to forensic applications, by maintaining integrity, transparency, authenticity, security, and auditability of digital evidence to achieve the desired end. Collecting, preserving and validating evidence can be strengthened with the help of forensic chain. The blockchain technology can also improve the law enforcement collaboration for a better track, monitor, and capture cyber criminals.

Many solutions for this bottleneck are being proposed and trialled, including increasing block size, having few nodes, side chains, random selection of block verifiers, *etc.* This paper proposes a digital forensics solution over the distributed ledger technology, which could provide a way to maintain the integrity of evidence that is digital from source and to

strengthen trust in the authorities involved in its handling and attestation.

VI. Conclusion

This work conducted preliminary forensic research on the blockchain-based forensic investigation framework by considering the diversity of devices, evidence items, data formats, and more in the complicated IoT environment. The main idea is to retrieve artifacts from IoT devices and further write to blockchain-based IoTFC after analysing the connections between evidence items, provenance, traceability, and auditability of each evidence item.

References

[1] F. Wang and L. Xu and W. Gao, "Comments on SCLPV: Secure Certificateless Public Verification for Cloud-Based Cyber-Physical-Social Systems Against Malicious Auditors," *IEEE Transactions on Computational Social Systems*, vol. 5, no. 3, pp. 854–857, Sep. 2018.

[2] A. Shah, R. Ganesan, S. Jajodia, and H. Cam, "Understanding trade-offs between throughput, quality, and cost of alert analysis in a csoc," *IEEE Transactions on Information Forensics and Security*, pp. 1–1, 2018.

[3] Y. Wu, G. Min, and L. T. Yang, "Performance analysis of hybrid wireless networks under bursty and correlated traffic," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 1, pp. 449–454, Jan 2013.

[4] S. Li, L. Da Xu, and S. Zhao, "5g internet of things: A survey," *Journal of Industrial Information Integration*, 2018.

[5] S. Wang and X. Wang and P. Ye and Y. Yuan and S. Liu and F. Wang, "Parallel Crime Scene Analysis Based on ACP Approach," *IEEE Transactions on Computational Social Systems*, vol. 5, no. 1, pp. 244–255, March 2018.

[6] Y. Liu and S. Hu, "Cyberthreat Analysis and Detection for Energy Theft in Social Networking of Smart Homes," *IEEE Transactions on Computational Social Systems*, vol. 2, no. 4, pp. 148–158, Dec 2015.

[7] G. Min, Y. Wu, and A. Y. Al-Dubai, "Performance modelling and analysis of cognitive mesh networks," *IEEE Transactions on Communications*, vol. 60, no. 6, pp. 1474–1478, June 2012.

[8] L. D. Xu, W. He, and S. Li, "Internet of things in industries: A survey," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 4, pp. 2233–2243, Nov 2014.

[9] L. M. Cullell. (2019) Digital forensics and blockchain. [Online]. Available: https://medium.com/@blockxlabs/digital-forensics-and-blockchain-bf3af5e7153c

[10] Y. Teing, D. Ali, K. Choo, M. T. Abdullah, and Z. Muda, "Greening cloud-enabled big data storage forensics: Syncany as a case study," *IEEE Transactions on Sustainable Computing*, pp. 1–1, 2018.

[11] S. Li, L. D. Xu, and X. Wang, "Compressed sensing signal and data acquisition in wireless sensor networks and internet of things," *IEEE Transactions on Industrial Informatics*, vol. 9, no. 4, pp. 2177–2186, Nov 2013.

[12] D. Zhao, L. Wang, Z. Wang, and G. Xiao, "Virus propagation and patch distribution in multiplex networks: Modeling, analysis and optimal allocation," *IEEE Transactions on Information Forensics and Security*, pp. 1–1, 2018.

[13] D. Zou, J. Zhao, W. Li, Y. Wu, W. Qiang, H. Jin, Y. Wu, and Y. Yang, "A multi-granularity forensics and analysis method on privacy leakage in cloud environment," *IEEE Internet of Things Journal*, pp. 1–1, 2018.

[14] S. Li and K. R. Choo and Q. Sun and W. J. Buchanan and J. Cao, "IoT Forensics: Amazon Echo as a Use Case," *IEEE Internet of Things Journal*, pp. 1–1, 2019.

[15] A. Paradise and A. Shabtai and R. Puzis and A. Elyashar and Y. Elovici and M. Roshandel and C. Peylo, "Creation and Management of Social Network Honeypots for Detecting Targeted Cyber Attacks," *IEEE Transactions on Computational Social Systems*, vol. 4, no. 3, pp. 65–79, Sep. 2017.

[16] G. Mezzour and W. Frankenstein and K. M. Carley and L. R. Carley, "A Socio-Computational Approach to Predicting Bioweapon Proliferation," *IEEE Transactions on Computational Social Systems*, vol. 5, no. 2, pp. 458–467, June 2018.

[17] J. Lee. (2018) Leveraging blockchain for forensic applications. [Online]. Available: https://www.blockchaindailynews.com/Leveraging-blockchain-for-forensic-applications_a25271.html

[18] S. Li, S. Zhao, P. Yang, P. Andriotis, L. Xu, and Q. Sun, "Distributed consensus algorithm for events detection in cyber physical systems," *IEEE Internet of Things Journal*, pp. 1–1, 2019.

[19] M. Hossain and Y. Karim and R. Hasan, "FIF-IoT: A Forensic Investigation Framework for IoT Using a Public Digital Ledger," in *2018 IEEE International Congress on Internet of Things (ICIOT)*, July 2018, pp. 33–40.

[20] M. Hossain and R. Hasan and S. Zawoad, "Trust-IoV: A Trustworthy Forensic Investigation Framework for the Internet of Vehicles (IoV)," in *2017 IEEE International Congress on Internet of Things (ICIOT)*, June 2017, pp. 25–32.

[21] M. Chernyshev, S. Zeadally, Z. Baig, and A. Woodward, "Internet of things forensics: The need, process models, and open issues," *IT Professional*, vol. 20, no. 3, pp. 40–49, May 2018.

[22] D. Quick and K. R. Choo, "Iot device forensics and data reduction," *IEEE Access*, vol. 6, pp. 47 566–47 574, 2018.

[23] L. Caviglione, S. Wendzel, and W. Mazurczyk, "The future of digital forensics: Challenges and the road ahead," *IEEE Security Privacy*, vol. 15, no. 6, pp. 12–17, November 2017.

[24] M. Cebe, E. Erdin, K. Akkaya, H. Aksu, and S. Uluagac, "Block4forensic: An integrated lightweight blockchain framework for forensics applications of connected vehicles," *arXiv preprint arXiv:1802.00561*, 2018.

[25] S. Wu, Y. Chen, Q. Wang, M. Li, C. Wang, and X. Luo, "Cream: A smart contract enabled collusion-resistant e-auction," *IEEE Transactions on Information Forensics and Security*, pp. 1–1, 2018.

[26] L. V. D. Horst, K. R. Choo, and N. Le-Khac, "Process memory investigation of the bitcoin clients electrum and bitcoin core," *IEEE Access*, vol. 5, pp. 22 385–22 398, 2017.

[27] H. Ritzdorf, C. Soriente, G. O. Karame, S. Marinovic, D. Gruber, and S. Capkun, "Toward shared ownership in the cloud," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 12, pp. 3019–3034, Dec 2018.

[28] G. Tziakouris, "Cryptocurrenciesa forensic challenge or opportunity for law enforcement? an interpol perspective," *IEEE Security Privacy*, vol. 16, no. 4, pp. 92–94, July 2018.

[29] Z. Liu and H. Seo, "Iot-nums: Evaluating nums elliptic curve cryptography for iot platforms," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 3, pp. 720–729, March 2019.

[30] A. Valjarevic and H. Venter, "A harmonized process model for digital forensic investigation readiness," in *IFIP International Conference on Digital Forensics*. Springer, 2013, pp. 67–82.

[31] D. Brezinski and T. Killalea, "Guidelines for evidence collection and archiving," Tech. Rep., 2002.

[32] K. Kent, S. Chevalier, T. Grance, and H. Dang, "National institute of standards and technology.guide to integrating forensic techniques into incident response. nist sp 800-86," *Online [Aug. 2006]*, 2006.

[33] Y. Zhang, S. Wu, B. Jin, and J. Du, "A blockchain-based process provenance for cloud forensics," in *2017 3rd IEEE International Conference on Computer and Communications (ICCC)*, Dec 2017, pp. 2470–2473.

[34] A. Al-Nemrat, "Identity theft on e-government/e-governance amp; digital forensics," in *2018 International Symposium on Programming and Systems (ISPS)*, April 2018, pp. 1–1.

[35] D. Ulybyshev, M. Villarreal-Vasquez, B. Bhargava, G. Mani, S. Seaberg, P. Conoval, R. Pike, and J. Kobes, "(wip) blockhub: Blockchain-based software development system for untrusted environments," in *2018 IEEE 11th International Conference on Cloud Computing (CLOUD)*, July 2018, pp. 582–585.

[36] M. M. Hossain, R. Hasan, and S. Zawoad, "Probe-iot: A public digital ledger based forensic investigation framework for iot." in *INFOCOM Workshops*, 2018, pp. 1–2.

[37] A. H. Lone and R. N. Mir, "Forensic-chain: Ethereum blockchain based digital forensics chain of custody," *Scientific & practical cyber security journal— ISSN 2587-4667*, 2018.

[38] A. Nieto, R. Roman, and J. Lopez, "Digital witness: Safeguarding digital evidence by using secure architectures in personal devices," *IEEE Network*, vol. 30, no. 6, pp. 34–41, November 2016.

[39] S. Li, S. Zhao, Y. Yuan, Q. Sun, and K. Zhang, "Dynamic security risk evaluation via hybrid bayesian risk graph in cyber-physical social systems," *IEEE Transactions on Computational Social Systems*, vol. 5, no. 4, pp. 1133–1141, Dec 2018.

[40] Z. Liu, X. Guan, S. Li, T. Qin, and C. He, "Behavior rhythm: A new model for behavior visualization and its application in system security management," *IEEE Access*, vol. 6, pp. 73 940–73 951, 2018.