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 decentralized nature of blockchain technologies can well match the needs of integrity and provenances of evidences collecting in digital forensics (DF) across jurisdictional borders. In this paper, a novel blockchain-based DF 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 guarantee 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 the proposed method.**

***Index Terms*— Digital forensics (DF), evidence items, internet of things, provenance.**

I. Introduction

OVER the years, the emerging technologies such as the social networks, Internet of Things (IoT), the fifth generation of communication (5G), and the decentralized blockchain technologies, have become an indispensable part of modern life [1]–[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 (DF) 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, and data analysis of large volumes of data set. In the past few years, many research efforts have focused on cloud-based forensic analysis [10], evidences modeling [11]–[13], and assisting the law enforcement community. In the IoT environment, DF are facing a number of challenges, including: 1) defining framework for DF 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). New research in DF must address these above challenges in the procedural, social, and legal field [14]–[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 that 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 the 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, and pharmaceutical. In forensic applications, the blockchain technology is promising to address the above-mentioned challenges. The advantage of blockchain technologies in DF 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

Manuscript received January 18, 2019; revised May 8, 2019 and June 30, 2019; accepted July 2, 2019. Date of publication July 26, 2019; date of current version December 9, 2019. *(Corresponding author: Shancang Li.)*

S. Li is with the Department of Computer Science and Creative Technologies, University of the West of England, Bristol BS16 1QY, U.K. (e-mail: shancang.li@uwe.ac.uk).

T. Qin is with the MOE Key Lab for Intelligent Networks and Network Security, Xi'an Jiaotong University, Xi'an 710049, China (e-mail: qin.tao@xjtu.edu.cn).

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).

Digital Object Identifier 10.1109/TCSS.2019.2927431

together with the transparency, traceability, trust between evidence/item and investigators, and continuous integrity of each evidence item. In Sections II and III, the 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 into transactional evidences (TEs) are liked with related evidence items using Merkle tree. The main aim of this work is to extract as much as possible potential digital evidences and reduce investigation costs in IoT environment.

The rest of this 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 this paper.

II. Related Works and New Challenges in Digital forensics

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

A. Related Works

In the past few years, a lot of research efforts have been conducted in IoT forensics [18]–[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]–[23]. Since 2017, the emerging blockchain technology has been applied in DF to document evidence items, interaction actions, and preserving evidence in the blockchain [24]–[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 formalized these developed forensic investigation processes. Specifically, in the complex digital environments, such as 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 DF processes and data privacy together and can provide efficient vehicle related digital investigation.

Zhang *et al.* [33] proposed a provenance process model for the digital investigation using blockchian in cloud environment, which aimed to enhance the interaction trust between stakeholders in cloud forensics. Al-Nemrat *et al.* [34] investigated the possibility to introduce blockchain technologies in the investigation of financial fraud in e-governance, and the results show 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], the blockchain is used to provide the auditability, traceability in software development, and a role-based access control mechanism for unauthorized data accesses is developed.

Hossain *et al.* [36] 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 and Mir [37] proposed a DF chain based on the popular blockchain platform Ethereum. The proposed forensic chain model was implemented over `Ethereum`, which can provide integrity, transparency, authenticity for data collected from multiple sources. A lot of research efforts have been done on the digital investigation in heterogeneous environment [23], lightweight security solutions over IoT devices [29], and digital witness [38].

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

B. Digital Forensics Challenges in IoT Environment

In DF, the 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 change. However, in the existing DF 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 forensic toolkit (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 summarized the challenges in the existing DF investigation as follows.

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

2) Integrity: Continuous integrity check for evidence items and examination events in digital investigation. In the 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 behaviors in examination by creating a hash tree.

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

$$HE_1 = Hash(Transanctional\ Evidence\#1) \quad (1)$$

$$HE_2 = Hash(Transanctional\ Evidence\#2) \quad (2)$$

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

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

4) Scalability: In a hash tree, a parent node is able to support up to 1000 children nodes, in DF, 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 a large number of evidence items/events [32].

5) Availability and Resiliency: Each node in the 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 the next section, this paper proposed a blockchain-enabled DF framework for the IoT, named as IoTFC.

III. Blockchain-Enabled Digital Forensics Investigation (IoTFC)

The blockchain technology can offer forensic applications with substantial benefits for the whole procedure of DF 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 of 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 was lost or compromised due to the unstable evidence systems. The cryptographic hash functions (such as SHA1, SHA256, and so on) are widely used in forensics imaging process aimed at the integrity of specific evidence items; however, for the whole evidence chain, a current continuous integrity check or validation mechanism is missing.

3) Immutability and Auditability: The nature of the blockchain technology can offer DF immutability and auditability, which are the key features required in DF 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 criticized 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, and tools) in the 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 generate digital fingerprint,

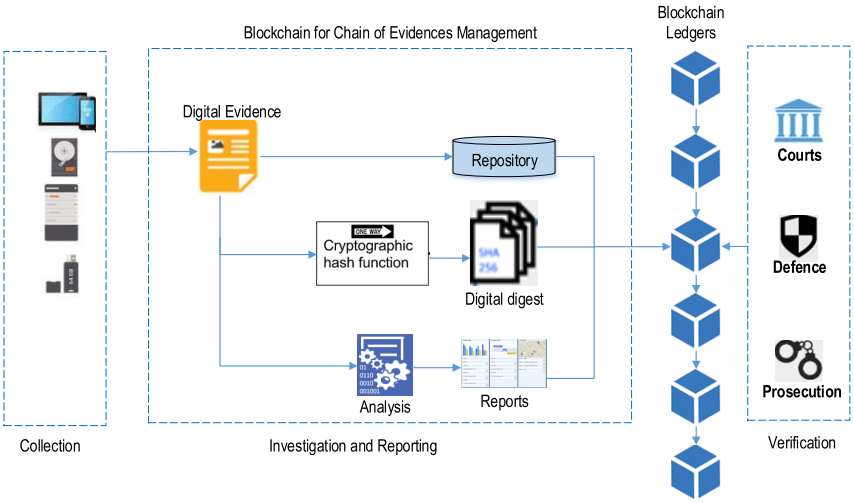

Fig. 1. Blockchain for chain of evidence management in IoT (IoTFC).

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 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 the following.

1) Blockchain can make the data acquisition and validation more accurate and informative by integrating the TEs and addition information;
2) For each TE item, its provenance as well as all related examining events can be traced back to its origination;
3) The IoTFC uses blockchain to build a close-loop system that provides significant forensic analysis benefits in an efficient and economical way.

C. Forensic-Chain Framework

IoTFC is a blockchain-based forensic solution for digital investigation forensic chain of custody, as shown in Fig. 1, which allows the system to create a distributed ledger for recording and storing TEs (examining events/findings, and additional information). These TEs 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 of following critical components.

1) Users and IoT Devices: The users include the users, owners, or examiner that are related with 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

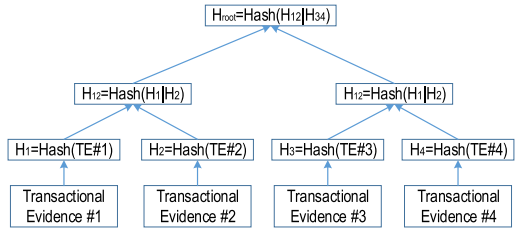

Fig. 2. Rough sketch of the structure of a Merkle tree.

of TEs in the investigation. It can summarize all the TEs, 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. Fig. 2 shows an example of Merkle trees of nodes, in which the `TE` could be a file, folder, memory.

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

1) 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, and full provenances. All this would be completed automatically through smart contract.
2) A blockchain browser is used to view the evidence blockchain, will more specific restrictions are defined according to the analysis requirements.

Fig. 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 TE items #1 and #2. In IoTFC, a DF 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.

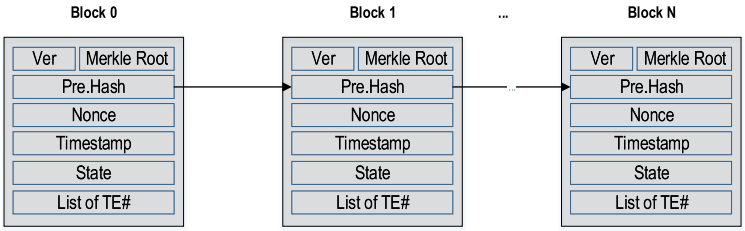

Fig. 3. Chained blocks.

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 Fig. 3. The `Prev.Hash` represents the hash value of the block header of preblock and a nonce. The TE item represents the evidence item record and it is hashed into a Merkle tree.

4) Smart Contract: Smart contract, also called blockchain contract, is a digitalized contract that is executable for a computer. The smart contract is usually stored in the 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 blockchain platform, such as Ethereum. It has been widely used in financial service, healthcare, insurance, e-government, and supply chain. Similarly, the smart contract can benefit the DF investigation from the following aspects.

1) Autonomy—it can define the conditions to find related evidences item in an automatic way.
2) Trust—the evidence item can be encrypted on a shared ledger.
3) Safety—the items can be cryptographically encrypted.
4) Speed—the smart contracts can significantly reduce the examining time than manually process.
5) Saving—smart contracts can save the cost without paying for middlemen, such as notary and witness.
6) 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 found; in this work, evidences are categorized into five grades.

1) g_1*:* Easy to identify, such as plan text in files, unencrypted image, and quick response (QR).
2) g_2*:* Some deliberate attempt at hiding, e.g., renaming of extension.
3) g_3*:* Hard to identify, e.g., plain text held other than in files system and volume slack.
4) g_4*:* Difficult to identify, i.e., encrypted data in a file and password protected xls file.
5) and g_5*:* Very difficult to identify, such as encrypted data held other than in the files system and steganography.

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, including `hashing, job Options, indexing/tools, and miscellaneous.` 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 the existing examination tools, such as `Autopsy, FTK, and EnCase`, 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 the following aspects.

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

Evidence items, examination event/actions, and additional information (e.g., `examiner, tools, workstation, and timestamps`) are formatted as `TE` shared by all participating parties over the blockchain network, where the IoTFC makes use of cryptography for protecting these TEs. Smart contracts are designed to create/record TEs based on the examination details, like address to whom evidence is transferred to, the current state of evidence, permission level, and data and time. Further any subsequent access to digital evidence also gets recorded securely on the blockchain by smart contracts triggered by the corresponding forensic investigation, as shown in Fig. 4.

IV. IoT Blockchain Forensics Applications

In the IoTFC, the links between each entity, such as *evidence item, devices, users, and social system account*, can be easily identified using the Merkle tree. To guarantee the integrity and auditability of digital evidences are very

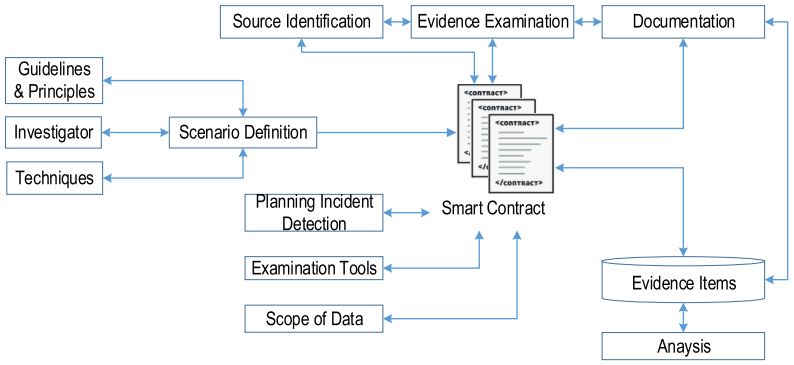

Fig. 4. Evidence identification as an application of smart contract.

important due to it moves along different levels of hierarchy in the chain of investigation. Basically, the IoTFC can enhance existing DF investigation from the following three ways.

1) Use the smart contract that makes some evidence analysis to be done automatically, such as file signature analysis and email analysis, improving investigation efficiency and reduce the data exchanges between parts.
2) Improve the transparency of investigation and provide better auditability.
3) Reduce examination costs and resource uses.
4) Establish connection with trusted third parties.

A. IoTFC Use Case

This section will introduce the use of blockchain in processing DF evidences. Digital evidences can be presented in the form of digital assets in all stages of DF 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 conscious how they communicate and save files. Investigators imaged the computer 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:

1) Joint Photographic Experts Group (JPG) images masquerading as a .docx file;
2) JPG images of bad thing (e.g., skittles.jpg.);
3) Steganography software found on the computer;
4) Pidgin instant messaging client, chat log conversation file;
5) Outlook email log files, password protected zip files;
6) Other suspicious files (e.g., pwd.txt).

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

1) *Contemporaneous Head:* `{examiner, examine commenced, software used, version and licensing}`.
2) *Event:* `{action, done?, Time, Notes, Screen capture}`.
3) *Evidence Item:* `{no., description of item, significance to case, full provenance, method discovery}`.

Using the IoTFC framework, all case related information can be easily chained in the blockchain system, such as evidence items, examination event/actions, and additional information (e.g., examiner, tools, workstation, and timestamps) are formatted as `TE` and can be accessed by all participants in the blockchain network. Smart contracts are designed to automatically create/record TEs 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 information and actions related with this file are formed into a TE $TE\#a1$, which will be written into the blockchain. In fact, further investigation is needed for this file. With `JPG Hide and Seek Steg` software, we extracted 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 summarized 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 `TE records`.

Step 2: Together with additional information and timestamps, the fingerprinted records will be written

Fig. 5. JSON script for evidence blocks.

into the 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 form TE item in blockchain. Similarly, the entirety of the full evidence chains will be formed in blockchain. When TEs 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 the forensic area, Fig. 4 shows an example of smart contract-based evidence item analysis. For network events related analysis, more interfaces are provided, such as intrusion analysis and log file analysis. Fig. 5 presents an example for JavaScript Object Notation (JSON) script for evidence item.

3) Presentation: this stage will be based on the findings in the 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 Fig. 6, in the first stage, all data are imaged and all acquisition related information are written into blockchain. In the identification stage, an suspicious image file is located in the acquisition and all identification events/findings in this stage are also written into blockchain. In the 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 the 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, and platforms) are fingerprinted and recorded in the blockchain. The IoTFC is an effective DF 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 photographs, documents, ease of time-line analysis, and IoT forensics artifact 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 QR 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 TEs 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 DF 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 trialed, including increasing block size, having few nodes, side

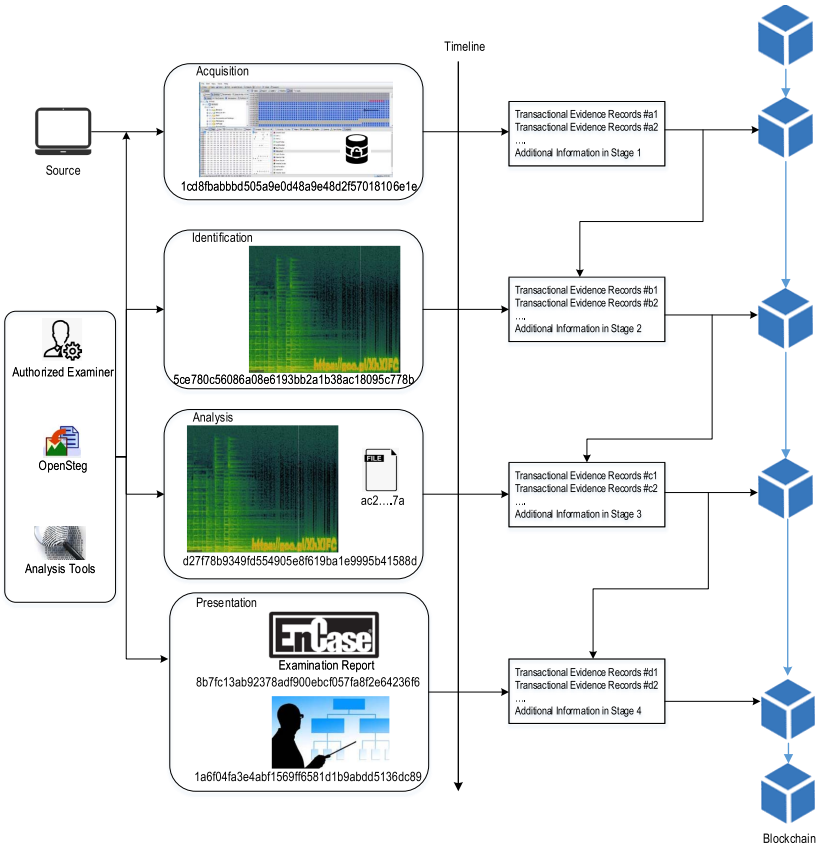

Fig. 6. Use case: steganography-based forensic analysis in IoTFC.

chains, and random selection of block verifiers. This paper proposes a DF 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 paper 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 analyzing the connections between evidence items, provenance, traceability, and auditability of each evidence item.

References

[1] Y. Zhang, C. Xu, S. Yu, H. Li, and X. Zhang, "SCLPV: Secure certificateless public verification for cloud-based cyber-physical-social systems against malicious auditors," *IEEE Trans. Comput. Social Syst.*, vol. 5, no. 3, pp. 854–857, Dec. 2018.

[2] A. Shah, R. Ganesan, S. Jajodia, and H. Cam, "Understanding tradeoffs between throughput, quality, and cost of alert analysis in a CSOC," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 5, pp. 1155–1170, May 2018.

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

[4] S. Li, L. Da Xu, and S. Zhao, "5G Internet of Things: A survey" *J. Ind. Inf. Integr.*, vol. 10, pp. 1–9, Jun. 2018.

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

[6] Y. Liu and S. Hu, "Cyberthreat analysis and detection for energy theft in social networking of smart homes," *IEEE Trans. Comput. Social Syst.*, 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 Trans. Commun.*, vol. 60, no. 6, pp. 1474–1478, Jun. 2012.

[8] L. D. Xu, W. He, and S. Li, "Internet of things in industries: A survey," *IEEE Trans. Ind. Inform.*, 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 Trans. Sustain. Comput.*, vol. 4, no. 2, pp. 204–216, Apr./Jun. 2018.

[11] S. Li, L. Da Xu, and X. Wang, "Compressed sensing signal and data acquisition in wireless sensor networks and Internet of Things," *IEEE Trans. Ind. Inform.*, 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 Trans. Inf. Forensics Security*, vol. 14, no. 7, pp. 1755–1767, Jul. 2019.

[13] D. Zou *et al.*, "A multigranularity forensics and analysis method on privacy leakage in cloud environment," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 1484–1494, Apr. 2019.

[14] S. Li, K. R. Choo, Q. Sun, W. J. Buchanan, and J. Cao, "IoT forensics: Amazon echo as a use case," *IEEE Internet Things J.*, to be published.

[15] A. Paradise *et al.*, "Creation and management of social network honeypots for detecting targeted cyber attacks," *IEEE Trans. Comput. Social Syst.*, vol. 4, no. 3, pp. 65–79, Sep. 2017.

[16] G. Mezzour, W. Frankenstein, K. M. Carley, and L. R. Carley, "A socio-computational approach to predicting bioweapon proliferation," *IEEE Trans. Comput. Social Syst.*, vol. 5, no. 2, pp. 458–467, Jun. 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 Things J.*, vol. 6, no. 2, pp. 2299–2308, Apr. 2019.

[19] M. Hossain, Y. Karim, and R. Hasan, "FIF-IoT: A forensic investigation framework for IoT using a public digital ledger," in *Proc. IEEE Int. Congr. Internet Things (ICIOT)*, Jul. 2018, pp. 33–40.
[20] M. M. Hossain, R. Hasan, and S. Zawoad, "Trust-IoV: A trustworthy forensic investigation framework for the Internet of vehicles (IoV)," in *Proc. IEEE Int. Congr. Internet Things (ICIOT)*, Jun. 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 Prof.*, vol. 20, no. 3, pp. 40–49, May/Jun. 2018.
[22] D. Quick and K. R. Choo, "Iot device forensics and data reduction," *IEEE Access*, vol. 6, pp. 47566–47574, 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, Nov./Dec. 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," 2018, *arXiv:1802.00561*. [Online]. Available: https://arxiv.org/abs/1802.00561
[25] S. Wu, Y. Chen, Q. Wang, M. Li, C. Wang, and X. Luo, "CReam: A smart contract enabled collusion-resistant e-auction," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 7, pp. 1687–1701, Jul. 2019.
[26] L. V. Der Horst, K.-K. R. Choo, and N.-A. Le-Khac, "Process memory investigation of the bitcoin clients electrum and bitcoin core," *IEEE Access*, vol. 5, pp. 22385–22398, 2017.
[27] H. Ritzdorf, C. Soriente, G. O. Karame, S. Marinovic, D. Gruber, and S. Capkun, "Toward shared ownership in the cloud," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 12, pp. 3019–3034, Dec. 2018.
[28] G. Tziakouris, "Cryptocurrencies—A forensic challenge or opportunity for law enforcement? an INTERPOL perspective," *IEEE Security Privacy*, vol. 16, no. 4, pp. 92–94, Jul./Aug. 2018.
[29] Z. Liu and H. Seo, "Iot-nums: Evaluating nums elliptic curve cryptography for iot platforms," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 3, pp. 720–729, Mar. 2019.
[30] A. Valjarevic and H. Venter, "A harmonized process model for digital forensic investigation readiness," in *Advances Digital Forensics*. Berlin, Germany: Springer, 2013.
[31] T. Killalea and D. Brezinski, *Guidelines for Evidence Collection and Archiving*, RFC Editor, 2002.
[32] J. H. Ryu, P. K. Sharma, J. H. Jo, and J. H. Park, "A blockchain-based decentralized efficient investigation framework for IoT digital forensics," *J. Supercomput.*, pp. 1–16, 2019.
[33] Y. Zhang, S. Wu, B. Jin, and J. Du, "A blockchain-based process provenance for cloud forensics," in *Proc. 3rd IEEE Int. Conf. Comput. Commun. (ICCC)*, Dec. 2017, pp. 2470–2473.
[34] A. Al-Nemrat, "Identity theft on e-government/e-governance & digital forensics," in *Proc. Int. Symp. Program. Syst. (ISPS)*, Apr. 2018, p. 1.
[35] D. Ulybyshev *et al.*, "(WIP) blockhub: Blockchain-based software development system for untrusted environments," in *Proc. IEEE 11th Int. Conf. Cloud Comput. (CLOUD)*, Jul. 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 *Proc. INFOCOM Workshops*, Apr. 2018, pp. 1–2.
[37] A. H. Lone and R. N. Mir, "Forensic-chain: Ethereum blockchain based digital forensics chain of custody," *Sci. Practical Cyber Secur. J.*, vol. 1, no. 2, pp. 21–27, 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, Nov./Dec. 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 Trans. Comput. Social Syst.*, 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. 73940–73951, 2018.

Shancang Li (M'12–SM'19) is currently with the Department of Computer Science and Creative Technologies, University of the West of England, Bristol, U.K. His current research interests include digital forensics for emerging technologies, cyber security, IoT security, data privacy preserving, Internet of Things, Blockchain technology, and the lightweight cryptography in resource constrained devices. He has authored more than 60 papers published in high-profile journals and conferences.

Mr. Li is the Associate Editor of IEEE ACCESS and *Journal of Industrial Information Integration*.

Tao Qin received the B.S. degree in information engineering and the Ph.D. degree in computer science and technology from Xi'an Jiaotong University, X'ian, China, in 2004 and 2010, respectively.

He is currently an Associate Professor with the Department of Computer Science and Technology, MOE KLINNS Lab, X'ian Jiaotong University. He is also a Visiting Scholar with the Department of Electrical and Computer Engineering, University of Massachusetts, Amherst, MA, USA. His current interests include Internet traffic analysis, traffic modeling, anomaly detection, and online social network analysis.

Geyong Min received the B.Sc. degree in computer science from the Huazhong University of Science and Technology, Wuhan, China, in 1995, and the Ph.D. degree in computing science from the University of Glasgow, Glasgow, U.K., in 2003.

He is currently a Professor of high performance computing and networking with the Department of Computer Science, College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, U.K. His current research interests include future Internet, computer networks, wireless communications, multimedia systems, information security, high-performance computing, ubiquitous computing, and modeling and performance engineering.