Block4Forensic: An Integrated Lightweight Blockchain Framework for Forensics Applications of Connected Vehicles

Mumin Cebe, Enes Erdin, Kemal Akkaya, Hidayet Aksu, and Selcuk Uluagac

The authors propose a permissioned blockchain framework among the various elements involved to manage the collected vehicle-related data. They first integrate VPKI into the proposed blockchain to provide membership establishment and privacy. They design a fragmented ledger that will store detailed data related to vehicle such as maintenance information/history, car diagnosis reports, and so on.

Abstract

Today's vehicles are becoming cyber-physical systems that not only communicate with other vehicles but also gather various information from hundreds of sensors within them. These developments help create smart and connected (e.g., self-driving) vehicles that will introduce significant information to drivers, manufacturers, insurance companies, and maintenance service providers for various applications. One such application that is becoming crucial with the introduction of self-driving cars is forensic analysis of traffic accidents. The utilization of vehicle-related data can be instrumental in post-accident scenarios to discover the faulty party, particularly for self-driving vehicles. With the opportunity of being able to access various information in cars, we propose a permissioned blockchain framework among the various elements involved to manage the collected vehicle-related data. Specifically, we first integrate vehicular public key infrastructure (VPKI) to the proposed blockchain to provide membership establishment and privacy. Next, we design a fragmented ledger that will store detailed data related to vehicles such as maintenance information/history, car diagnosis reports, and so on. The proposed forensic framework enables trustless, traceable, and privacy-aware post-accident analysis with minimal storage and processing overhead.

Introduction

Today's vehicles are becoming much smarter with special-purpose sensors, control units, and wireless adapters to monitor their operations and communicate with their surroundings [1]. These contemporary *smart vehicles* are now considered as a comprehensive cyber-physical system (CPS) with communication, control, and sensing components [2]. For instance, electronic control units (ECUs) and onboard units (OBUs) can receive data from various onboard sensing devices to take certain actions. The connections among the control units and sensor devices are made via different types of networks, including a controller area network (CAN) bus, a local interconnect network (LIN) bus, FlexRay, Bluetooth, and so on. Such developments along with capabilities to sense and communicate with the surroundings are enabling further developments such as the creation of autonomous vehicles, also known as self-driving cars, which will revolutionize our lives.

The penetration of Internet of Things (IoT) technologies in vehicles enables collection of enormous data from vehicles for various applications. For instance, most vehicles that are manufactured in the last decade have onboard diagnostics (OBD) ports which are used for retrieving vehicle controller diagnostics. These ports are typically interfaced with a WiFi, Bluetooth, or serial connection to supply data outside.

Another major development is the deployment of event data recorders (EDRs) by leading manufacturers including GM, Ford, and so on. EDRs are meant to store incident data based on triggering events. Finally, future vehicles will be equipped with OBUs to enable connectivity among vehicles and roadside units (RSUs) to provide collision avoidance and congestion control. Such safety features will be realized with wireless dedicated short-range communications (DSRC), which will not only enable broadcasting of basic safety messages (BSMs) [15] (i.e., vehicle-to-vehicle, V2V) but also provide the means to communicate with the infrastructure such as traffic lights and railroad crossings. (i.e., vehicle-to-infrastructure, V2I). Although BSM is the name of a special message in the DSRC specification, here it is used as a generic name allocated to all safety-related messages [1, 15].

Capabilities such as collecting data within and around vehicles can have a significant impact on vehicular forensics, which aims to investigate the reasons behind the accidents. This field will become even more important with the proliferation of self-driving cars, which are prone to failures and cyber attacks [3]. Typically, after an accident, investigator specialists analyze the causes of the accident so that disputes among parties can be resolved. The investigators look at many different aspects including inspection of the accident site and vehicles. Site inspection contains physical evidence including scrub marks, position of vehicles, tire conditions, and so on. In addition to physical evidence, digital data supplied from OBD ports and EDRs introduce valuable complementary evidence for supporting dispute resolution. Eventually, by enabling the capture, storage, and transfer of the vehicle data, the puzzle including drivers, insurance companies, manufacturers, and law enforcement authorities can be solved [3, 4].

Digital Object Identifier:
10.1109/MCOM.2018.1800137

The authors are with Florida International University.

Even after utilizing EDR and OBD data, accident investigation lacks certain features that are absolutely needed for comprehensive dispute resolution. These can be listed as follows:

- The obtained data does not include a comprehensive history of the vehicle due to limited storage (i.e., the data is overwritten after a while).
- The parties do not have direct control over the extracted data; therefore, they should trust third parties, which incurs questions about the integrity of data.
- There is no system for integrating data from all parties including other vehicles, road conditions, manufacturers, and maintenance centers.
- There is no vehicular forensics solution to resolve a hit and run case other than third-party information such as surveillance cameras and eyewitnesses.

Therefore, in this article, we address these points by proposing the Block4Forensic (B4F) framework, a blockchain-based vehicular forensics system that will collect vehicles' and related business components under the same umbrella. In particular, the proposed system:

- Provides a lightweight privacy-aware blockchain by gathering all related parties such as drivers, maintenance centers, car manufacturers, and law enforcement without requiring a trusted third party in case of an incident
- Introduces a vehicular forensics investigation framework that harbors all necessary data for a comprehensive vehicular forensics solution

The rest of the article is organized as follows. In the following section, we describe the preliminaries related to all concepts and provide a summary of the state of the art. Then we introduce the B4F framework. Following that, we explain BF4 with its components. The next section is dedicated to future issues in this emerging research area. Finally, we conclude the article.

Background

Vehicular Forensics: Traditional vehicular forensics deals with the physical evidence collected from an accident scene, such as photographs, measurements, and scrub marks. Usage of vehicle-generated data has attracted the interest of researchers [5]; hence, it is strengthening the hands of forensic investigators as they can find supporting evidence from the digital subsystems of a vehicle. There are many controllers and sensors in modern vehicles with different capabilities. For a better driving experience, almost every capability of the vehicle is measured and reported.

When an accident occurs, first responders arrive at the scene to identify and secure the digital devices to keep them forensically sound (preserving the integrity of evidence) by following the process shown in Fig. 1. After securing and getting access to all related devices, further examination and analysis are performed. This basically means finding incident-related data on the digital devices such as finding traces of a cyber attack and failure of a manufacturer component or the mistake of a driver and so on.

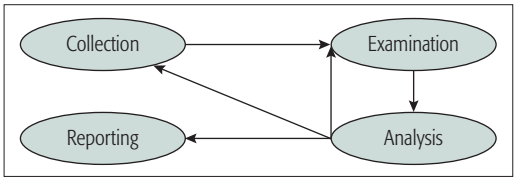

Figure 1. Digital forensics process model [6].

At the reporting phase, investigators prepare a report, and testify and present the evidence. Obviously, the most important factor in the admissibility of the report is to verify that the evidence devices have not been altered during the investigation. This may be quite challenging as there is no universal standard to collect, examine, and analyze data from digital devices on vehicles, drivers, and involved units. Therefore, a framework that will enable convenient data collection and analysis is needed. The framework should satisfy privacy of the user, and the stored data content should be clear to the user (i.e., the owner of the data).

Event Data Recorders and Onboard Diagnosis: The event data recorder (EDR), informally named the "black box," is a device placed in vehicles in order to collect data related to crashes and accidents. In the case of a dispute, investigators come up with the most probable setup. The digital data recorded by the EDR is widely used as supporting evidence in investigations for reconstructing the accident scene. When a triggering event occurs — two of those events are airbag deployment and sudden speed changes above a threshold — the EDR captures and stores the state of the vehicle in tamper-proof storage. It is known that EDR data is extracted by the investigators through the onboard diagnosis (OBD) port in an incident. Meanwhile, the ownership of EDR data and its integrity is discussed in [7] along with how this data is used by the traffic safety administrator (TSA) and other third parties for post-accident scenario reconstruction.

DSRC and Basic Safety Messages: DSRC specification defines the dedicated channels, standards, and protocols for communication between connected vehicles. Among many different messages, the basic safety message (BSM) is one of the most important ones for safety-related awareness between vehicles. Part I of the BSM includes high-priority information about a vehicle such as position, speed, size, brake status, and ID of the vehicle, and also medium-priority messages such as positional accuracy and steering wheel angle. This scheme brings additional value to the forensic investigation since the collected digital data will not be related solely to the car itself but also to the participants surrounding it.

Vehicular Public Key Infrastructure: Vehicular Network Security

In the networking layer of communication of connected vehicles, IEEE 1609.2 is utilized for message integrity and authentication [8].

The vehicular public key infrastructure (VPKI), a simplified version of which is shown in Fig. 2, utilized in IEEE 1609.2 is a highly complicated infrastructure specially tailored to the needs of the transportation system. The main certification authority (CA) generates, distributes, and

Traditional vehicular forensics deals with the physical evidence collected from an accident scene, including photographs, measurements, scrub marks, and so on. Usage of vehicle-generated data has attracted the interest of researchers; hence, it is strengthening the hands of the forensic investigators as they can find supporting evidence from the digital subsystems of a vehicle.

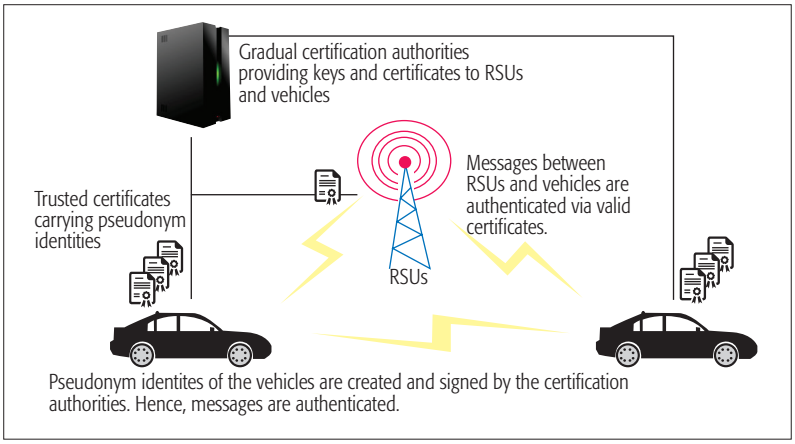

Figure 2. A simplified representation of VPKI.

revokes the digital certificates. The proposed VPKI structure also deals with privacy and security issues. According to the safety pilot model, the certificates that constitute the pseudonym identity of the vehicle are valid for only five minutes. That behavior provides anonymity for the communicating parties and also makes the system strong against targeted attacks against privacy and spoofing.

Blockchain

A blockchain is composed of blocks that are linked to each other and secured cryptographically. This establishes a strong tie between blocks that guarantees the order of blocks and provides an implicit strong timestamp mechanism. Thus, a block is prevented from any alteration without changing all of its successors. This blockchain data structure can be shared to build a distributed data structure called a *shared ledger* [9]. This working scheme of blockchain carries unique properties such as relieving central authority trust, immutability, and timestamping.

There are two types of blockchain structure: public and permissioned. For instance, Bitcoin and Ethereum fall into public blockchain category where everyone is able to read and write the ledger without any restriction (i.e., there is no membership requirement). However, in permissioned blockchains [10], the participants form a members-only club.

The process of adding a new block to the chain is carried out via a protocol, which establishes consensus among participants to confirm the new block. The implementation details of the consensus protocol (e.g., proof of work or POW) change a lot depending on the type of blockchain. For instance, in public blockchain, a consensus is typically in the form of a hash puzzle which requires finding a predefined hash value. This consensus protocol brings a significant level of security to the chain (withstanding up to 50 percent of nodes being malicious), but at the cost of computational power and time. For instance, Bitcoin's maximum throughput is 7 transactions/s, and reaching a final consensus can take an hour. On the other hand, permissioned blockchains utilize some kind of Byzantine fault-tolerant voting-based algorithm as a consensus mechanism, such as Practical Byzantine Fault Tolerance (PBFT) or Stellar Consensus Protocol (SCP), which do not require computationally expensive hash puzzles. As a result, reaching a consensus is faster, which means higher transaction throughput. However, permissioned blockchains generally require more than two-thirds of nodes to be trustworthy rather than 51 percent. More details about consensus algorithms can be found in [11].

Current State of the Art in Vehicular Forensics

The use of digital vehicular forensics is increasingly being investigated. There are commercial products targeting comprehensive data collection from the cars. The iVe project from Berla is a result of that effort, where their product has access to EDR and OBD port. They also retrieve data from the infotainment and telematics systems. The data is collected on cloud storage. Authors in [4] offer a similar solution. EDR and OBD ports are accessible by design and the data is stored in the cloud. Although the authors in [12] do not directly aim implementation for digital forensics, they offer a framework mainly discussing guidelines named "forensic by design." The idea of blockchain utilization for vehicular security is offered in [13]. The authors sketch possible use cases for insurance companies or wireless software updates for smart cars; however, their discussion lacks practical issues such as membership management and scalability. For a proper investigation, non-repudiation is of great importance. There is an implicit consensus in the research community that public key cryptography produces reliable solutions for that issue [14]. However, there is a need for a comprehensive applicable and scalable framework for vehicular forensics research.

A Blockchain Framework for Vehicular Forensics

The ultimate aim of vehicular forensics is to resolve disputes and determine the faulty parts in the case of an accident. Developments in *connected vehicles* provide new opportunities for forensic analysis by taking advantage of the IoT and CPS features. Utilizing produced sensors' data with decision entities would allow building a comprehensive vehicular forensic analysis. Considering involving multiple parties, including manufacturers, drivers, insurance companies, law enforcement, and so on, we first identify the key features for an effective and trustworthy vehicular forensics framework.

Desired Features of Envisioned Forensic Analysis

The following key features are desired for Vehicular Forensics.

Integrity: The integrity of forensic data is very important for resolving disputes.

Non-Repudiation: The parties should be held responsible for their actions by providing proof of integrity.

Relieve Single Point of Trust: The system should remove the assumption of trust reliance solely on a single authority and provide accountable trustworthiness for each participant.

Comprehensive Forensic Analysis: The system should provide a comprehensive mechanism for accident analysis by providing access to historical data even before the accident. For example, the behavioral pattern of the vehicle after main-

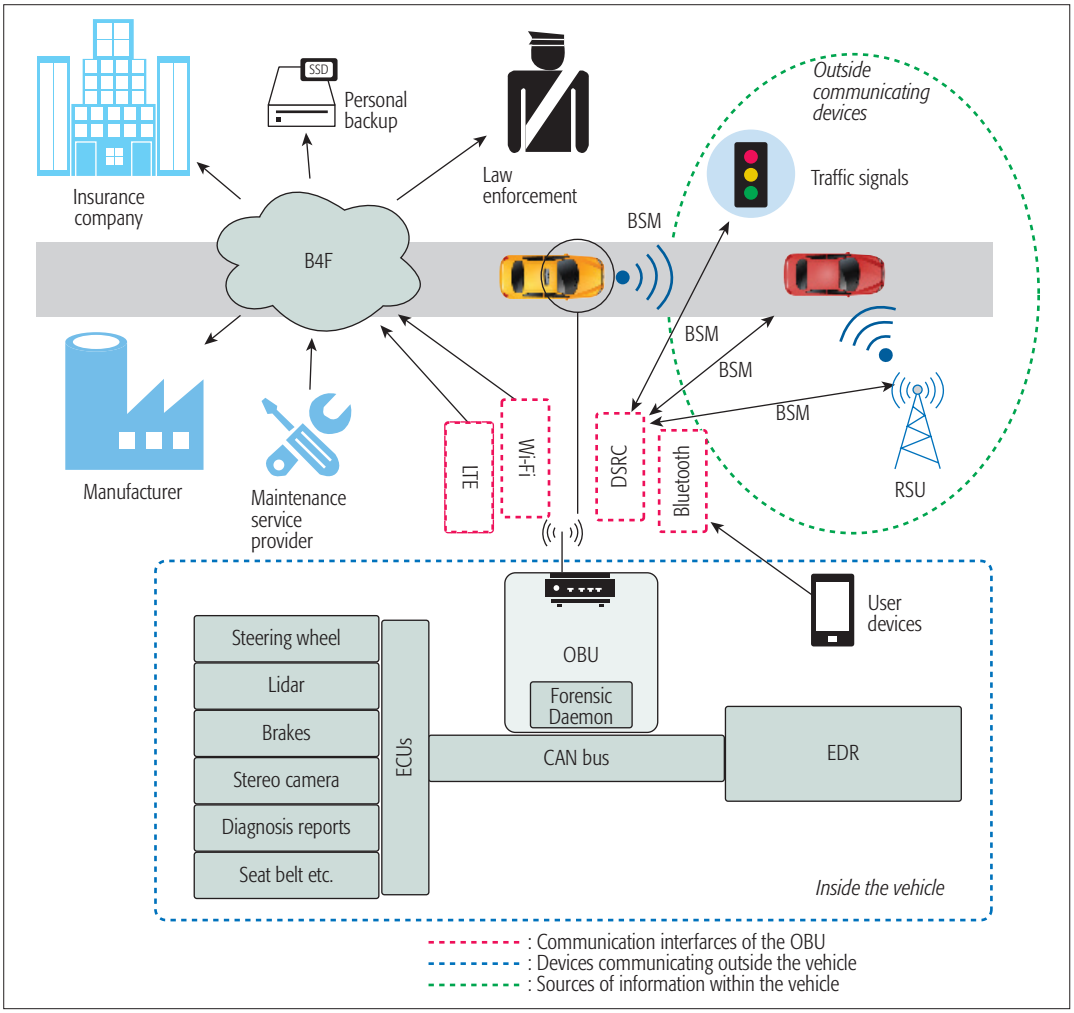

Figure 3. An overview of the forensic system model with its stakeholders.

tenance (e.g., steering ability, braking distance) or a previously reported malfunctioning component of a vehicle can provide important clues to determine the faulty party.

Lightweight: The system should have minimum overhead on endpoints since it includes multiple parties that may have different capabilities and resources.

Privacy: The system should preserve the privacy of the participants while also providing the flexibility for the participants to selectively reveal their data as they wish.

B4F Framework

To enable the vehicular forensics vision, we introduce a novel blockchain forensic framework as shown in Fig. 3. The framework connects the following stakeholders: vehicles, maintenance service providers (e.g., mechanics), vehicle manufacturers, law enforcement, and insurance companies. The key features of the envisioned vehicular forensics system mentioned in the previous subsection guided us while building the blockchain-based vehicular forensics system.

At the heart of design, there is a special forensic daemon, which is stationed within the OBU and constantly retrieves data from EDR, BSMs (i.e., messages received from other vehicles), and onboard sensors/IoT devices through a CAN Bus. The *forensic daemon* periodically shares the EDR and BSM data with the insurance company through an encrypted channel. Note that only related BSMs are shared when an EDR triggering event occurs. On the other hand, the car manufacturers collect regular car diagnostic reports. A cryptographic hash of these data is submitted to Blockchain for removing the single trust issue. Both insurance companies and manufacturers collect those data for analysis. Moreover, maintenance records are kept at the maintenance service providers, and a hash of each record is submitted to Blockchain in the same manner. As an optional extension to the framework, all of the mentioned data can also be stored in personal cloud storage.

Stored data will be used in post-accident scenarios by allowing the parties to disclose their data selectively to determine the faulty party. Law enforcement authorities play an investigative role for post-accident scenarios while parties disclose their data with proof of integrity.

Potential Accident Scene

An investigator working on an accident scene needs to collect all pieces of clues to reconstruct the accident scene. Once the accident scene is reconstructed, the faulty party can be determined accordingly.

Stored data will be used in post-accident scenarios by allowing the parties to disclose their data selectively to determine the faulty party. Law enforcement authorities play an investigative role for post-accident scenarios while parties disclose their data with proof of integrity.

In a modern vehicle, many important sub-structures like steering wheel motor, braking system, throttle, tire pressure monitoring system, seat belt buckle status even windshield wipers are controlled and monitored via the CAN bus. Thus, the CAN bus may deliver invaluable data in terms of vehicular forensics to the OBU which can be retrieved by the forensic daemon.

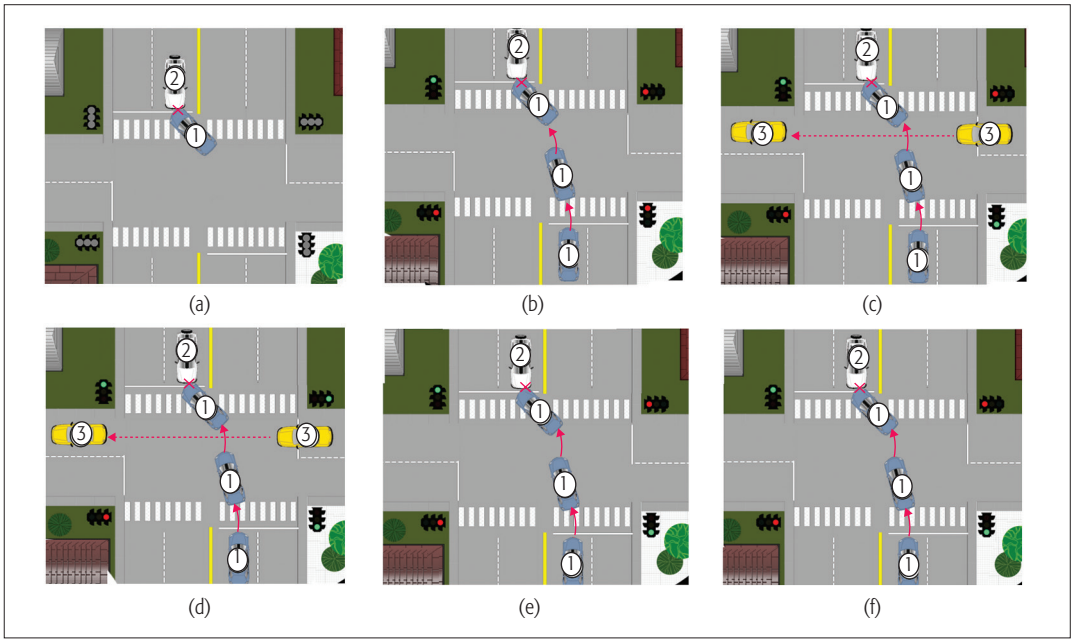

Figure 4. A Hypothetical accident scene and possible reconstructions of the accident: a) reported accident scene; b) faulty driver; c) hit and run; d) faulty signaling; e) faulty maintenance; f) faulty manufacturer.

Here, we discuss how digital data provided by B4F assist an investigation. Assume that an accident scene where Vehicle 1 (V1) collided with Vehicle 2 (V2) at an intersection with traffic lights as illustrated in Fig. 4a. The data provided by B4F may enable various forensically sound scene reconstructions as listed below.

Reconstructed Scene (b): BSM messages include the traffic light status and cars' last positions. In this scenario, BSM messages reveal that V1 started to turn left when the red light was on, as shown in Fig. 4b. Lights' statuses are being disseminated by smart traffic lights; thus, when the accident happens, B4F would have stored the last BSM messages from the traffic lights. Here, data clearly point out that V1 is the faulty party.

Reconstructed Scene (c): Timestamped data in B4F reveals the existence of another vehicle at the accident scene. Drivers of V1 and V2 started crossing the road when the light turned green. At that time V3 did not stop at the red light and caused V2 to lose control and hit V1. B4F data uncovers the existence of V3 and resolves such a hit case where the faulty party is a third car that runs out of the incident area.

Reconstructed Scene (d): Similar to scene (c), data reveals the existence of V3. However, this time none of the cars violate the rules as the traffic light for V3 is also green. BSM data supplied by smart traffic lights would reveal faulty signaling as the cause of the accident.

Reconstructed Scene (e): In this scenario, B4F data indicates that none of the drivers has violated the traffic rules. However, by investigating the car diagnostic report history on B4F, the investigator finds out that after maintenance, the vehicle has a pulling problem while braking. Due to this faulty operation in V1, the driver lost control of the car and hit V2. The history of previous vehicle maintenance records helps to resolve this complicated scenario and suggests the maintenance provider as the faulty party.

Reconstructed Scene (f): In this scenario, B4F data shows that V1 was on autopilot at the accident time. Moreover, the diagnostic records report a failed sensor. Thus, V1 autopilot software with faulty input caused the accident, which suggests the car manufacturer as the faulty party.

Various parties might be involved in an accident as exemplified above. Forensic data provided by B4F provides a fast and efficient accident scene reconstruction, which helps any investigation significantly.

B4F Components

In this section, we first describe the forensics elements and data types, and then we move on to elaborate on the specific elements of B4F that relate to the blockchain structure, its membership management, and storage issues.

Forensic Daemon

Here, we explain how the proposed *forensic daemon* interacts with different components of a vehicle. Note that our *forensic daemon* runs as an application in an OBU thanks to existing software development kits (SDKs) for custom application development.

The OBU has read access to the vehicle network infrastructure. The backbone of the vehicle network is the CAN bus. In a modern vehicle, many important substructures like steering wheel motor, braking system, throttle, tire pressure monitoring system, seat belt buckle status, and even windshield wipers are controlled and monitored via the CAN bus. Thus, the CAN bus may deliver invaluable data in terms of vehicular forensics to the OBU that can be retrieved by the *forensic daemon*.

Additionally, through WiFi or Bluetooth interfaces, the *forensic daemon* can receive data from the driver about his/her health status via wearables. Similarly, road conditions and weather data can be retrieved from RSUs or a driver's smartphone that has applications related to such data.

The *forensic daemon* will collect data on predefined occasions based on basic or custom rules.

After adding a timestamp, it will sign the data using the pseudonym certificate, which is readily available in the OBU. In the case of an investigation, submitted data will be disclosed for investigation by the user.

Forensic Data Types and B4F Process

In this subsection, we detail the interaction between the vehicle and the B4F framework. There are three types of data in our framework. The first one is event data, which are incurred in the case of an incident triggered by the predefined conditions in EDR. The second one is the diagnosis data, which are produced by the vehicle periodically or in the case of a failure. Finally, there are maintenance data, which contains information about the maintenance report and is kept by both the maintenance service and the user. Maintenance data are signed by both the vehicle and the maintenance provider and hence are multi-signature data. We have two data submission processes in the B4F. As described below, the content of forensic data along with time and pseudonym vehicle ID is signed by vehicle and submitted to the corresponding parties such as insurance companies, manufacturers, and personal cloud storage. While the content of the forensic data is kept between two parties, the hash of this data is stored in the shared ledger on blockchain. B4F implements a gossip network where each vehicle selects a random set of validators to gossip about the hash of data. To ensure that messages are valid, every message is signed by the pseudonym identity of the vehicle; validators check that the signature is valid before relaying it. The randomly chosen leader proposes a block in submitted transactions and distributes its block of pending transactions through the gossip protocol again. B4F establishes a Byzantine agreement to reach the final conclusion.

Blockchain Structure

To address the requirements above, we propose utilizing *permissioned* blockchain technology and implement *shared* and *fragmented* ledgers to securely and efficiently exchange information between the collaborating parties.

In our proposed blockchain, we have four different types of nodes: leader, validator, monitor units, and client as shown in Fig. 5.

A leader is selected randomly every block time among the validator nodes (i.e., manufacturers, maintenance centers, insurance companies). The client (i.e., vehicle) provides signed transactions to the B4F to ensure that messages cannot be forged.

The randomly chosen leader proposes a block to the network based on the transactions it has received. To reach a consensus on a proposed block, validators run Byzantine agreement protocols such as PBFT. These protocols are resilient to malicious actions of the leader and participants [11]. Monitor units are law enforcement authorities who do not directly participate in the validation process but keep a replica of the shared ledger to be able to participate in post-accident disputes.

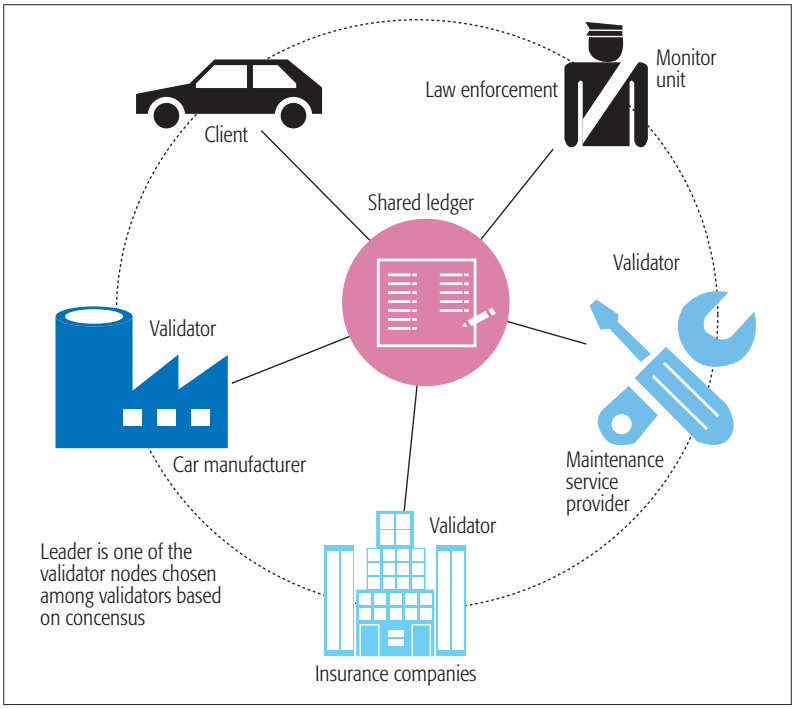

Figure 5. Permitted blockchain participants.

This proposed framework is geared for increasing the level of trust among network participants and thus will eliminate the need for a trusted third party.

Due to the use of hashes, the overhead of the building and storing replicated shared ledger among parties is minimized. Note that the integrity of data can be verified by comparing its hash value with the corresponding hashes that are stored on the Blockchain.

Integrated Membership Management and Privacy via Pseudonym Certificates

In a public blockchain, anyone can participate as either a client or a validator (e.g., miners in cryptocurrencies). However, in the case of a permitted blockchain, access permission is strictly controlled by membership service, and only granted users are able to make transactions. The identities issued by membership service are unique and cannot be altered. Thus, there is no support to protect privacy between interacting peers. Leveraging permitted blockchain impedes the use of anonymous identities in contrast to identities used in public Blockchains such as Bitcoin. This is particularly important in our case since vehicle owners would like to protect their privacy while sharing data with their manufacturers and insurance companies. On the other hand, the huge number of network participants (e.g., millions of vehicles on the roads) expose membership management as a challenge in the realization of a permitted blockchain.

Thus, we use pseudonym identities from the VPKI model suggested in IEEE 1609.2 as a token for clients to satisfy anonymity (i.e., vehicles) in the proposed B4F. According to the VPKI scheme, the vehicle has different pseudonym identities for different time intervals (i.e, every five minutes); thus, every transaction will be submitted with a different identity, which protects the

Blockchain is a shared ledger that maintains a growing list of blocks that are chained to each other. Each participant stores a copy of the entire history. In our case, the data are immense and thus the shared ledger can grow dramatically and may cause both communication and storage overhead.

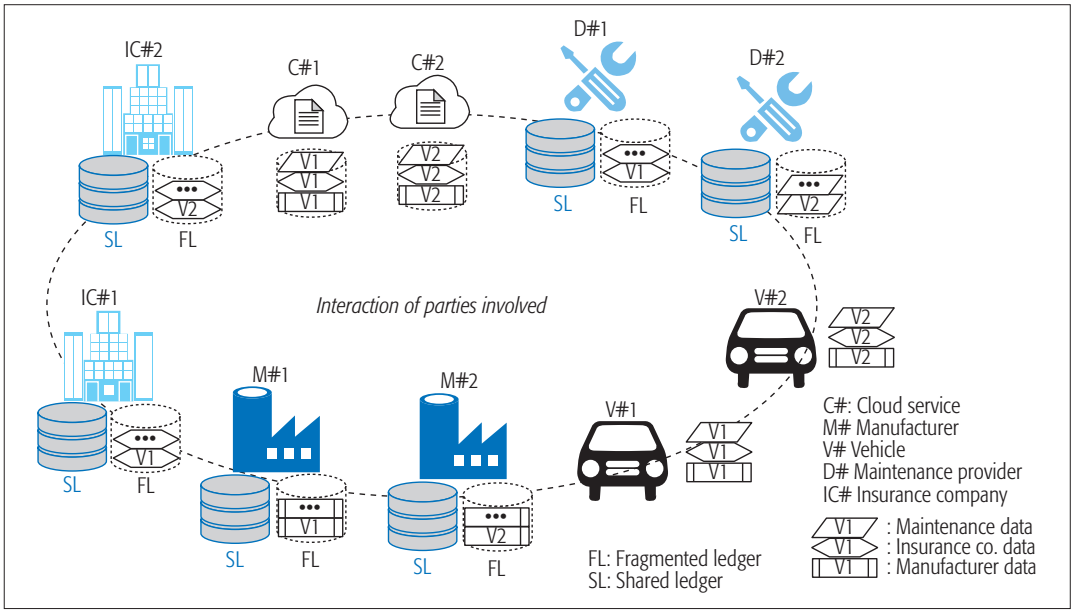

Figure 6. An overview of the proposed ledger structure. SL: shared ledger; FL: fragmented ledger, which can hold different data.

user privacy as defined in the attack model of IEEE 1609.2. However, regulations and policies should be assessed for proper disclosure of the user data. In addition, exploiting the VPKI scheme also addresses the above mentioned membership management challenge. Any vehicle that has a valid pseudonym identity can make transactions on the proposed Blockchain since participants of B4F recognize valid certificates produced by VPKI. Validator nodes check the validity of the certificate and timestamp of the submitted data (i.e., hash of the forensic data). If the timestamp belongs to the certificate validity period (i.e., every five minutes), the transaction is confirmed. The consensus on valid transactions is achieved by a computationally inexpensive voting-based Byzantine agreement scheme among validators.

Lightweight Fragmented Ledger for Forensic Participants

Blockchain is a shared ledger that maintains a growing list of blocks which are chained to each other. Each participant stores a copy of the entire history. In our case, the data are immense, and thus the shared ledger can grow dramatically and may cause both communication and storage overhead.

To address this issue, we utilize a *fragmented ledger* instead of storing all forensic data in a shared ledger. The motivation comes from the observation that each party has already stored a different fragment of required data. For instance, a maintenance provider may not be interested in the content of periodic EDR data, and thus there is no need to keep that content in a shared ledger. On the contrary, as insurance companies keep EDR data in their fragmented ledger, keeping proof of that data in the shared ledger is sufficient. Therefore, in B4F, all participants of the network will have a consensus on the shared ledger. However, each participant maintains just related information that differs from others, as shown in Fig. 6. Specifically, the difference between the shared and fragmented ledgers will be in forensic data details. The shared ledger does not carry any information related to the forensic content of EDR&BSM data, car diagnostic reports, provided maintenance, and so on.

Additionally, note that the user may want to refuse to submit maintenance or manufacturer data content. Instead, s/he keeps it in personal cloud storage. However, based on regulations and policies, in the case of an incident, the authorities will require the user to disclose this data as needed, the integrity of which is satisfied by the Blockchain.

Future Research Issues

As there is growing research on various aspects of connected vehicles, their applications will proliferate in coming years, such as driverless cars and automated fleets. This may result in increased disputes as a result of incidents. Therefore, we believe that there is a vast opportunity to pursue additional research with respect to vehicular forensics in general and our framework in particular. We list them below:

- There will be a need to analyze the storage and communication overhead of the B4F framework by implementing it using an OBU SDK.
- A punishment/incentive/avoidance mechanism should be investigated to prevent members becoming malicious actors. In this regard, a detection mechanism should be developed to discover malicious participants.
- The B4F provides a lightweight solution by just keeping hash values. While this ensures integrity and immutability of forensic data, the availability of this data depends on the individual storage and shared counterparts. There is no mechanism for ensuring availability of critical forensic data on blockchain. Therefore, this warrants further research.
- Due to increased availability of data and blockchain technologies in various domains for forensic purposes, researchers would need to consider a forensic-by-design principle when proposing new systems and mechanisms.

- Regulations for enforcing the participation of various entities to forensic blockchains and development of policies to use such data in criminal cases are potential research issues.

Conclusion

In this article, we propose constructing a blockchain infrastructure to provide comprehensive forensic services for accident investigations. To address the issues regarding the overhead of storage and membership management of blockchain, we propose using VPKI in permitted blockchain and a fragmented ledger, which enables storage of hashed data in the shared ledger while the details are stored in fragmented ledgers as non-hashed data. In addition, the use of pseudonyms for identities helps preserve the privacy of users.

Acknowledgments

This work is partially supported by the U.S. National Science Foundation (Awards: NSF-CAREER-CNS-1453647, NSF-1663051).

References

[1] IEEE Standard for Wireless Access in Vehicular Environments (WAVE), 2016.
[2] C. Berger and B. Rumpe, "Autonomous Driving 5 Years after the Urban Challenge: The Anticipatory Vehicle as a Cyber-Physical System," arXiv preprint arXiv:1409.0413, 2014.
[3] Z. A. Baig *et al.*, "Future Challenges for Smart Cities: Cyber-Security and Digital Forensics," *Digital Investigation*, vol. 22, 2017, pp. 313.
[4] H. Mansor *et al.*, "Log Your Car: The Non-invasive Vehicle Forensics," *Proc. IEEE Trustcom/BigDataSE/I SPA*, 2016, pp. 974–82.
[5] D. K. Nilsson and U. E. Larson, "Conducting Forensic Investigations of Cyber Attacks on Automobile In-Vehicle Networks," *Proc. 1st Int'l. Conf. Forensic Applications and Techniques in Telecommun., Info., and Multimedia and Wksp.*, ICST, 2008, p. 8.
[6] U. Karabiyik and K. Akkaya, "Digital Forensics in IoT and WSNs," *The Philosophy of Mission-Oriented Wireless Sensor Networks*, Springer, 2018.
[7] N. Gabriel, A. Niedzicka, and C. Krysiuk, "The Use of Event Data Recorder (EDR)-Black Box," *Advances in Science and Technology Research J.*, vol. 8, no. 21, 2014.
[8] IEEE Standard for Wireless Access in Vehicular Environments, Security Services for Applications and Management Messages WAVE 1609.2-2016, 2016.
[9] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System," 2008.
[10] C. Cachin, "Architecture of the Hyperledger Blockchain Fabric," *Proc. Wksp. Distributed Cryptocurrencies and Consensus Ledgers*, 2016.
[11] S. Bano *et al.*, "Consensus in the Age of Blockchains," arXiv preprint arXiv:1711.03936, 2017.
[12] N. H. Ab Rahman *et al.*, "Forensic-by-Design Framework for Cyber-Physical Cloud Systems," *IEEE Cloud Computing*, vol. 3, no. 1, 2016, pp. 50–59.
[13] A. Dorri *et al.*, "Blockchain: A Distributed Solution to Automotive Security and Privacy," *IEEE Commun. Mag.*, vol. 55, no. 12, Dec. 2017, pp. 119–25.
[14] J. Li, H. Lu, and M. Guizani, "ACPN: A Novel Authentication Framework with Conditional Privacy-Preservation and Non-Repudiation for VANETs," *IEEE Trans. Parallel and Distributed Systems*, vol. 26, no. 4, 2015, pp. 938–48.
[15] J. Cho *et al.*, "Efficient Safety Message Forwarding Using Multi-Channels in Low Density VANETs, *IEEE GLOBECOM*, Dec. 2014, pp. 70-75.

Biographies

Mumin Cebe is a Ph.D. student in the Department of Electrical and Computer Engineering at Florida International University. He works at the Advanced Wireless and Security Lab (ADWISE). He conducts research in the areas of blockchain, wireless networking, and security/privacy that relates to the Internet of Things and cyber-physical systems, particularly in smart grids and vehicular networks.

Enes Erdin is a Ph.D. student in the Department of Electrical and Computer Engineering at Florida International University and is an NSF CyberCorps Fellow. He conducts research in the areas of hardware security, blockchain technology, and cyber-physical systems.

Kemal Akkaya is a professor in the Department of Electrical and Computer Engineering at Florida International University. He leads the Advanced Wireless and Security Lab and is an Area Editor of the *Elsevier Ad Hoc Network Journal*. His current research interests include security and privacy, and protocol design. He has published over 120 papers in peer reviewed journals and conferences. He received the "Top Cited" article award from Elsevier in 2010.

Hidayet Aksu received his Ph.D. degree from Bilkent University in 2014. He is currently a postdoctoral associate in the Department of Electrical & Computer Engineering at Florida International University. Before that, he worked as an adjunct faculty member in the Computer Engineering Department of Bilkent University. He conducted research as visiting scholar at IBM T. J. Watson Research Center, New York, in 2012–2013. He also worked for the Scientific and Technological Research Council of Turkey (TUBITAK).

Selcuk Uluagac leads the Cyber-Physical Systems Security Lab at Florida International University, focusing on security and privacy of the Internet of Things and cyber-physical systems. He has Ph.D. and M.S. degrees from Georgia Institute of Technology, and an M.S. from Carnegie Mellon University. In 2015, he received the U.S. National Science Foundation CAREER award and the U.S. Air Force Office of Sponsored Research's Summer Faculty Fellowship, and in 2016, a Summer Faculty Fellowship from the University of Padova, Italy.

To address the issues regarding the overhead of storage and membership management of blockchain, we proposed using VPKI in permissioned blockchain and a fragmented ledger which enables storage of hashed data in the shared ledger while the details are stored in fragmented ledgers as non-hashed data.