Privacy in the Internet of Things: A Study to Protect User's Data in LPR Systems Using Blockchain

Iago Ochôa*, Leonardo Calbusch*, Karize Viecelli*, Juan De Paz‡, Valderi Leithardt*,Cesar Zeferino*
*Laboratory of Embedded and Distributed Systems – LEDS, University of Vale do Itajaí, Itajaí – SC, Brazil
‡Departamento Informática y Automática Plaza de la Merced, 37008. – Salamanca – Spain
{iago.ochoa, leonardo.calbusch, karize.viecelli}@edu.univali.br, fcofds@usal.es, {valderi, zeferino}@univali.br

***Abstract*—Over the past decade, smart crime-fighting solutions have been adopted by the major cities around the world. In this context, license plate recognition (LPR) systems have been used by public safety forces to monitor vehicle movement. However, current systems store vehicle location data indistinctly, without differentiating vehicles that are under criminal investigation from those that are not. This monitoring may be used to infer personal data about the owner of the vehicle, resulting in a violation of privacy by disregarding data protection laws. This paper presents a study about the use of technologies to ensure privacy in the Internet of Things and proposes a model to protect data collected by LPR systems. Our solution uses private blockchains regulated by smart contracts to ensure that the storage of data complies with current data protection laws.**

***Index Terms*—IoT, Blockchain, Privacy, LPR.**

I. Introduction

In recent years, the terms Internet of Things (IoT) and Smart Cities have become popular around the world. IoT is defined as a broad view where things like everyday objects, places, and environments are interconnected with each other through the Internet [1]. This definition can serve as a basis for the design of Smart Cities. Smart cities are presented as structures for real-time data collection and integration based on the use of sensors, applications, personal devices, and other interconnected resources [2]. These features, when integrated into a computing platform, provide a set of smart services developed to solve urban problems and contribute to both the sustainable development of cities and the improvement of the quality of life of their citizens.

In this context, several areas of city infrastructure can benefit from smart services. In the field of public safety, one of the solutions that have been standing out in the major cities of the world in the last years is the monitoring of the movement of vehicles through license plate recognition (LPR) systems, also called license plate readers.

One of the major concerns raised by the use of LPR systems regards the privacy of the captured data [3]. These systems do not distinguish between vehicles under criminal investigation and those that are not. This feature violates the data protection laws in countries around the world that protect the storage of personal data of the population without consent or motivation provided by law. The image captured of a vehicle, its license plate, the surrounding environment, its location, and the time and date the image was captured are data that provide a basis for inferring personal characteristics about the driver. Things done by or for the individual driving the vehicle, in addition to the presence of that individual at a particular location at any given time, can be determined by analyzing the captured data.

This article proposes a storage architecture that uses blockchain technology to guarantee the privacy of data captured by LPR systems. The architecture proposed relies on the Ethereum platform, a decentralized network capable of executing smart contracts. In our model, each smart contract will match the privacy preferences of a license plate that will be anonymized through public encryption. Thus, the storage of data captured by the LPR system cannot be carried out if privacy protection is enabled in the smart contract associated with the license plate. A gateway will be responsible for controlling access to the blockchain smart contracts. In case of motivation foreseen by the legislation, the smart contract associated with a specific license plate can be changed by a competent user to allow the storage of the data captured by the LPR system.

The remainder of this paper is organized as follows. Section II presents the theoretical basis on which this work was based, discusses the operation of LPR systems and the current scenario of adoption of these systems around the world, and ponders about data protection legislation in different countries. In Section III, a state-of-the-art analysis is conducted regarding the use of blockchain to ensure privacy in IoT environments. Section IV describes the architecture proposed to solve the privacy problem of the chosen scenario. Section V discusses the results obtained from experiments performed to evaluate the proposed architecture. Finally, Section VI presents the conclusions obtained with the development of this work and discusses future work.

II. Background

LPR systems are image recognition systems that extract the license plates of vehicles that transit at a particular point in a traffic lane, recording the date and time of the passage. According to [4], LPR systems identify the vehicle license plate number from one or more camera-taken pictures that may be of the color, black-and-white, or infrared type. This

978-1-7281-3265-5/19/$31.00 ©2019 IEEE

identification works by combining various techniques, such as object detection, image processing, and pattern recognition. The identified license plate number can then be associated with data stored in databases, and this crossing of information allows for more complete analyzes and obtaining new information.

When a vehicle is detected by an LPR system, its code or license plate number is read and instantly compared to vehicle database records of interest in criminal investigations. A law enforcement officer can then intercept and stop a vehicle, check for evidence and, when necessary, make arrests. A record for all vehicles passing through a camera is stored, including those for vehicles that are not known to be of interest at the time of reading [5].

The storage of data regarding vehicles that are not of investigative interest has raised concerns about the privacy of citizens. On the one hand, police forces around the world claim that the use of LPR systems has increased the power of crime prevention and aided investigations. On the other hand, civic organizations and ordinary citizens have questioned whether LPR systems protect the personal data associated with the identified license plates. They fear that such data could be used for inappropriate purposes unrelated to public safety. In any case, the use of LPR systems by police forces has increased significantly around the world, under different reasons [6].

The expansion of LPR systems usage has raised questions about protecting the privacy of citizens. LPR systems allow the monitoring of any individual who has the license plate of his/her vehicle identified. Various organizations, groups, and news agencies have been concerned about this aspect of technology.

In the United States, the EFF (Electronic Frontier Foundation), a nonprofit organization working on digital rights advocacy, requested information regarding the use of LPRs in the country. By means of the transparency law, they assessed that, between 2016 and 2017, more than 2.5 billions of vehicular license plate recognitions were performed by LPR systems. Also, according to the reports analyzed, 99.5% of the monitored vehicles were not associated with criminal investigations [7].

Two of the most critical aspects of privacy regarding data captured by LPR systems are the time that the records are maintained by the LPR database and the control access to the database. The concern for privacy primarily focuses on readings kept in the LPRs database that were not associated with activities of interest to the police when they occurred. These records can be explored later, at the discretion of who owns the property and access to the database because, in most cases, LPR systems are sold to government institutions by private companies, and it is unclear to what kind of use such data may be susceptible.

However, each country has specific laws protecting the privacy of its citizens. In general, these laws protect the citizen against the storage of personal data captured without explicit consent. Some exceptions are allowed, such as the storage of personal data by public interest, public security, and others. The problem is compliance with these laws by the technologies employed in the LPR systems in use today. In the system proposed in this article, we will meet these requirements to offer a solution that fits most countries possible.

The General Data Protection Regulation (GDPR) is a European Union (EU) Law that was adopted by the European Parliament in April 2016. The law is an evolution of the 1995 European Directive (Directive 95/46/EC) and came into force after a transitional period of two years, on 25 May 2018 [8]. The law applies not only to organizations located within the EU but also to all companies which process and hold the personal data of data holders residing in the EU irrespective of the location of the company.

GDPR provides some exceptions to the need for consent to the use of personal data, such as in matters relating to national security, defense, and public safety [9]. In addition to these exceptions, we should highlight the provision in the law of excluding the need for consent in cases involving the prevention, investigation, detection or prosecution of criminal offenses, or the execution of criminal sanctions aimed at safeguarding against threats to public security and prevention.

Unlike the EU, the United States follows a sector-wide approach to data privacy protection. There is no comprehensive federal law that guarantees the privacy and protection of personal data. Instead, legislation at the national level primarily protects data within industry-specific contexts. Personal data protection in the country depends on a combination of federal and state laws, administrative regulations, and specific self-regulatory guidelines. The privacy protection guarantees are specific for each state and are located in a wide range of legislative instruments and jurisprudence [10].

GDPR refers to the term *pseudonymization* as a principle to protect personal data. GDPR defines the term as the processing of personal data in such a way that the data can no longer be assigned to a specific Data Subject without the use of additional information. In other words, pseudonymization is the treatment by which a data loses the possibility of an association, directly or indirectly, to an individual, but by the use of additional information maintained separately by the controller in a controlled and safe environment.

In this context, blockchain technology presents itself as a viable solution to the problem above, since its principles rely on the intensive use of cryptography, a key feature of blockchain networks, in addition to bringing reliability behind all the interactions in the network. Smart contracts – automatic execution scripts residing in the blockchain – integrate these concepts and allow distributed and highly automated workflows [11].

III. Related Work

In our bibliographic research, we did not find references that propose the use of blockchain for LPR systems. Thus, in our review about the related work, we sought to identify studies that define architectures for IoT scenarios similar to our proposal.

Yu et al. [12] report several security issues that should be considered in IoT environments such as data integrity, authentication, access control, and privacy. To ensure data integrity, the authors suggest the use of proof-of-space consensus algorithm, where miners need to prove that they have disk space to store data generated by IoT devices. According to the authors, the use of smart contracts with pre-defined access policies can guarantee authentication and access control. Finally, to ensure privacy, it is recommended to use private blockchains and consortium blockchains, allowing only those who participate in the blockchain to have access to stored information.

The work proposed by Cha et al. [13] uses a gateway in conjunction with the blockchain to ensure the privacy of IoT devices. The model proposed by the authors applies smart contracts to store the access policies of each IoT device. Thus, when a user requests access to the data of IoT device through the gateway, this user must accept the privacy policies of the IoT device, according to the smart contract. In the scenario described by the authors, the blockchain is used to store each user's privacy preferences, making them fraud-proof and also solving privacy conflicts between users and IoT service providers.

Different from the solutions presented by other authors, Ayoade et al. [14] used only one smart contract to manage data. This contract has functions for user registration, device registration, and data reading/writing policy. A gateway does the communication among the IoT devices, user and blockchain. The authors conclude that their architecture is valid for the resolution of the data management problem in IoT devices, but they point out that scalability problems will arise if the data of the IoT devices are stored in the blockchain.

Dorri et al. [15] propose an architecture that uses a centralized blockchain for each Smart Home present in the system, and each of these centralized blockchains is connected to a public blockchain to reduce network overhead. To ensure privacy, the scheme proposed by the authors uses different private keys to perform the communication between the centralized and the public blockchain.

Wang et al. [16] report the use of blockchain for crowdsensing applications. The proposed architecture consists of using smartphone devices to process sensing data. A cryptocurrency award is given to users to encourage data mining process. The management of the payment made to users is performed through a blockchain. To ensure data privacy and user identity, the authors chose to use the K-anonymity method that prevents attacks on group nodes.

Le et al. [17] describe an IoT scenario with various types of forensic data sensor. The authors propose the use of a blockchain allowed to store the data acquired by the sensors. To guarantee user privacy, the authors propose the employment of a Merkle signature on the public key of each user, thus avoiding the identification and mapping of the keys.

Table I summarizes the main features observed in each work discussed above. The SC column identifies if the authors chose to use smart contracts to provide privacy in their architectures.

In our architecture, we use all these techniques in order to obtain a high level of privacy for the users. Following, we describe the scenario adopted and the architecture used to guarantee the privacy requirements.

TABLE I
STATE-OF-THE-ART CHARACTERIZATION

Work	SC	E	A	BC
Yu et al. [12]	•			•
Cha et al. [13]	•			
Ayoade et al. [14]	•			
Dorri et al. [15]		•		•
Wang et al. [16]			•	•
Le et al. [17]		•	•	•
This work	•	•	•	•

Where: SC: uses Smart Contract, E: applies encryption, A: employs anonymization, BC: uses blockchain.

IV. ARCHITECTURE

We consider in our system that a user should not be monitored when he/she does not want to. Nevertheless, we take into account that the legislation of each country establishes conditions for a person to be monitored at a certain moment. Based on this premise, our architecture addresses three requirements:

- An individual who does not want to be monitored will not be monitored unless the government has a legal order for that;
- The government can monitor someone. However, at the end of the investigation process, the user should be alerted that he/she was being monitored;
- An individual can be monitored and be aware of that if it is necessary (e.g., monitoring of people on probation).

Fig. 1 illustrates the scenario described. First, the user requires the license plate for his/her car and defines the privacy preferences for monitoring, and this information is registered in a smart contract stored in the private blockchain (*i*). At this moment, the public and private keys of the user are also generated. When the license plate is captured by an LPR system, the captured image is sent to the gateway responsible for managing all communications (*ii*). This gateway connects to the database that has stored the public keys corresponding to each license plate and retrieves the corresponding public key. After retrieving the public key, the gateway connects to the blockchain and checks the privacy preferences of the captured license plate through the smart contract (*iii*). If the privacy preference of this license plate allows the image to be captured, the gateway stores its image in a storage service (*iv*).

To meet the second and third requirements of the proposed architecture, Fig. 2 illustrates how the system works if a user needs to be monitored. In (*i*), to initiate the monitoring process, the government user must obtain a court order authorizing the monitoring. The court order must be sent to the gateway. The order is encrypted (*ii*) using the license plate owner public

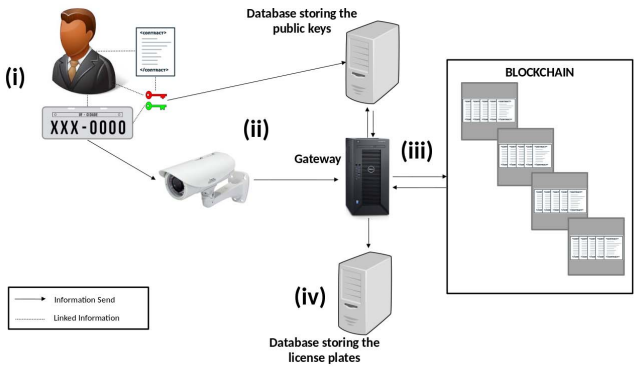

Fig. 1. Proposed Architecture.

key, and the smart contract privacy preferences are changed, allowing the license plate to be monitored for a specified period. Upon completion of the monitoring time (*iii*), the encrypted court order is sent to the user stating that he/she has been monitored. When the monitoring time ends, the contract privacy preferences are updated to the original settings.

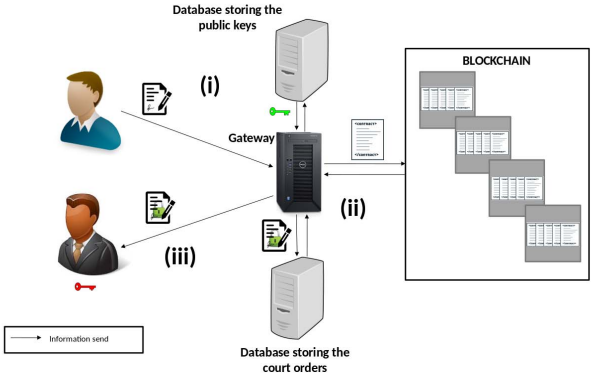

Fig. 2. Architecture for monitoring an user for a specified time.

V. Results

To verify the feasibility of the proposed architecture, we have conducted performance tests and measured the execution time in each of the tests, as follows.

A. Public Key Recovery

The first test aimed at identifying the time it takes to retrieve the public key of a user by connecting the gateway to the database, as shown in Fig. 2. In the experiments, we have used PostgreSQL 9.4 database running on a host computer with Debian 9.8 OS, 4 GB of RAM, and Intel Core I5 1.6 GHz processor. The gateway was executed on another computer running Ubuntu 18.04 LTS with 8 GB of RAM and Intel Core I5 2.3 GHz processor. We have measured the time to fetch 1, 10, 100, and 1,000 keys at a time, using three database sizes: 100,000, 1 million, and 10 million license plates. Table II summarizes the query execution time to retrieve a given number of keys for each database size. We can observe that the execution time of the queries varies according to the number of keys retrieved and also with the database size. It is worth noting that the maximum number of license plates stored in the database is greater than the number of vehicle plates licensed in cities like New York, which do not reach the amount of 10 million private cars [18].

TABLE II
Key recovery query execution time varying database size

Keys Recovered/DB Size	100K	1KK	10KK
10	1.68 ms	2.09 ms	1.91 ms
100	1.80 ms	7.13 ms	9.18 ms
1000	4.12 ms	10.03 ms	12.05 ms

B. Smart Contract Cost

The smart contract developed to manage privacy preferences of each user consumes a quantity of gas to be stored in the Ethereum blockchain. We have implemented a private blockchain in our system using Ganache in conjunction with Truffle to develop the contracts. The first point observed during the network building was the limit amount of gas for each block. By default, Ganache uses blocks of 6,721,975 gas, while Main-net currently uses blocks of 8,000,000 gas. To define the limit amount of gas per block in our blockchain, we have first verified the gas cost of the smart contract we developed.

In our tests, we have verified the gas cost of each contract by varying the number of addresses allowed to change privacy preferences of each user. Table III presents the results obtained. As we can observe, the gas cost increases according to the number of addresses mapped in the contract. The Ethereum network becomes expensive to store values, and so we have chosen to use ten addresses for each contract.

TABLE III
Gas cost varying the address quantity

Addresses	Gas Cost
1	173,833 gas
10	300,290 gas
100	1,324,958 gas

In our network, we have defined the limit quantity of 6,005,800 gas per block, allowing the storage of exactly 20 contracts per block. Since we employed a private blockchain, this value will not compromise the network performance.

C. Contract Registration

We have used the web3.js library to register and verify the contracts by making the connection with the blockchain. We have measured the execution time for the registration of 1, 10, and 100 contracts. Table IV summarizes the results and shows that the time spent to register transactions in blockchain varies

TABLE IV
EXECUTION TIME FOR CONTRACT REGISTRY IN BLOCKCHAIN

Contracts	Execution time
1	0.52 s
10	4.20 s
100	38.35 s

linearly with the number of contracts being registered at the same time.

Based on the tests developed, we observe that the architecture proposed in our work is feasible to be employed in real-world systems. However, it is necessary to mention that for using it in more than one city, the system may not remain scalable. In view of this, considering the tests developed in our work, we propose the use of sidechains for each city or city conglomerate, limiting the number of license plates to 10 million per conglomerate.

VI. CONCLUSIONS AND FUTURE WORK

This work presented a state-of-the-art architecture that guarantees privacy for users in LPR systems. The proposed model consisted of using a private blockchain in conjunction with smart contracts and anonymization through ECC cryptography.

For this work, we have conducted a review of the literature for solutions to provide privacy in IoT platforms. The study demonstrated the lack of proposals to guarantee privacy in LPR systems and characterized the solutions currently developed for other IoT scenarios. In view of this, our work seems to be the first one to propose a solution to provide privacy in LPR systems using blockchain.

The results obtained confirmed the feasibility of implementing the proposed architecture. However, according to the tests performed, it is necessary to use sidechains for each state/city to maintain network performance satisfactorily. Using only one blockchain to store all vehicle license plates of a country would make the system impracticable to be implemented because of performance and scalability issues.

The solution proposed in this work is not restricted to LPR systems. Anyone can adapt the proposed architecture to other environments that need privacy, security, and trust in IoT scenarios.

It is worth mentioning that all the modules of the proposed architecture were tested separately. As this research is a work in progress, the integration of all modules of our architecture is still under development.

As future work, we intend to develop the remaining part of the contract to perform the monitoring of users for a specified period. We will also develop optimizations for smart contracts to provide even more privacy for system users.

ACKNOWLEDGMENT

This work was financed by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001 and by Fundação de Amparo à Pesquisa de Santa Catarina – Brasil (FAPESC) – Grant 2019TR169.

REFERENCES

[1] T. L. Koreshoff, T. Robertson, and T. W. Leong, "Internet of things: a review of literature and products." *Proceedings of the 25th Australian Computer-Human Interaction Conference on Augmentation, Application, Innovation, Collaboration - OzCHI 13*, 2013.

[2] C. Harrison, B. Eckman, R. Hamilton, P. Hartswick, J. Kalagnanam, J. Paraszczak, and P. Williams, "Foundations for smarter cities," *IBM Journal of Research and Development*, vol. 54, no. 4, pp. 1–16, jul 2010. [Online]. Available: https://doi.org/10.1147/jrd.2010.2048257

[3] American Civil Liberties Unions, "You are being tracked: How license plate readers are being used to record americans' movements," 2019.

[4] S. Du, M. Ibrahim, M. Shehata, and W. Badawy, "Automatic license plate recognition (ALPR): A state-of-the-art review," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 23, no. 2, pp. 311–325, feb 2013. [Online]. Available: https://doi.org/10.1109/tcsvt.2012.2203741

[5] UK Home Office, "Automatic number plate recognition," 2019.

[6] CNN, "Policing advocates defend use of high-tech license plate readers," 2013.

[7] The Electronic Frontier Foundation, "Data driven: Explore how cops are collecting and sharing our travel patterns using automated license plate readers," 2018.

[8] European Parliament, "Regulation (eu) 2016/679 of the european parliament," 2016.

[9] EU, "Art. 23 GDPR - Restrictions," 2019.

[10] S. M. Boyne, "Data protection in the united states," *The American Journal of Comparative Law*, vol. 66, no. suppl_1, pp. 299–343, jul 2018. [Online]. Available: https://doi.org/10.1093/ajcl/avy016

[11] K. Christidis and M. Devetsikiotis, "Blockchains and smart contracts for the internet of things," *IEEE Access*, vol. 4, pp. 2292–2303, 2016, cited By 331. [Online]. Available: https://doi.org/10.1109/ACCESS.2016.2566339

[12] Y. Yu, Y. Li, J. Tian, and J. Liu, "Blockchain-based solutions to security and privacy issues in the internet of things," *IEEE Wireless Communications*, vol. 25, no. 6, pp. 12–18, December 2018.

[13] S. Cha, T. Tsai, W. Peng, T. Huang, and T. Hsu, "Privacy-aware and blockchain connected gateways for users to access legacy iot devices," in *2017 IEEE 6th Global Conference on Consumer Electronics (GCCE)*, Oct 2017, pp. 1–3.

[14] G. Ayoade, V. Karande, L. Khan, and K. Hamlen, "Decentralized iot data management using blockchain and trusted execution environment," in *2018 IEEE International Conference on Information Reuse and Integration (IRI)*, July 2018, pp. 15–22.

[15] A. Dorri, S. S. Kanhere, and R. Jurdak, "Towards an optimized blockchain for iot," in *2017 IEEE/ACM Second International Conference on Internet-of-Things Design and Implementation (IoTDI)*, April 2017, pp. 173–178.

[16] J. Wang, M. Li, Y. He, H. Li, K. Xiao, and C. Wang, "A blockchain based privacy-preserving incentive mechanism in crowdsensing applications," *IEEE Access*, vol. 6, pp. 17 545–17 556, 2018.

[17] D. Le, H. Meng, L. Su, S. L. Yeo, and V. Thing, "Biff: A blockchain-based iot forensics framework with identity privacy," in *TENCON 2018 - 2018 IEEE Region 10 Conference*, Oct 2018, pp. 2372–2377.

[18] "New York State Vehicle Registrations," https://dmv.ny.gov/statistic/2017reginforce-web.pdf, accessed in May 2019.