RESEARCH ON RELIABLE, SECURE AND PRIVACY PRESERVING TRANSMISSION IN THE VANETS

by

Zishan Liu

Dissertation submitted in fulfilment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

Department of Engineering
Faculty of Science and Engineering
Macquarie University
Sydney, Australia

December 2019

STATEMENT OF CANDIDATE

I certify that partial of the work in this thesis entitled "Research on Reliable, Secure and Privacy Preserving Transmission in the VANETs" has previously been submitted for a degree in the partner university, Beijing University of Posts and Telecommunications, according to the Cotutelle agreement with Macquarie University.

I also certify that the thesis is an original piece of research and it has been written by me.

In addition, I certify that all information sources and literature used are indicated in the thesis.

Zishan Liu

Dedicated to my superviors Iain B. Collings, Lin Zhang and Wei Ni

My family

And all of my friends.

ACKNOWLEDGMENTS

The past one half years at Macquarie have been an unforgettable and invaluable experience to me. I would not have been able to make this journey without the help and support of my superviors, my friends, my family... many, many people and I feel indebted to them. First and foremost, my most sincere and greated thanks go to my supervior Prof. Iain Collings and my co-supervior Dr. Wei Ni. I am grateful to my supervior, for his careful guidance, patience and support to me that I could no have asked for more. He always has a very proficient, insightful and high-level view about the wireless communication field. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my Ph.D study. During my study in Macquarie, my co-supervior, Dr. Wei Ni, who works at the Australian Commonwealth Scientific and Industrial Research Organization (CSIRO), has given me insightful comments and encouragement. His selfless, kindness, optimism, patience and rigorous spirit on research have deeply touched and motivated me to have fun during the research.

I thank my supervior Lin Zhang from Beijing University of Posts and Telecommunications. I am very grateful for his recommendation and encouragement to let me apply the Cotutelle Ph.D project in Macquarie. Prof. Zhang is a kind, caring and supportive advisor that I could not have asked for more. He is like an older friend and a family member of mine (if he doesn't mind me saying so). He always believes in me even though I am not always that confident about myself.

I am forever grateful to my superviors.

My sincere thanks also goes to Dr. Shihao Yan and Dr. Min Li, who helped me a lot when we stayed together in the EMC building, and worked together in the wireless communication group in Macquarie. They have helped a lot and given me insightful comments on my work.

I have been extremely lucky to be surrounded by many great teachers and friends when I am staying in Macquarie. Just to name a few (and forgive me for not being able to list all of them): Ms. Sheley, Jane Yang, Xiaoxia Yang, Xiaojing Chen, Yameng Zheng, Wenwen Zhang, Sicong Tian, Jinsong Hu, Lingmeng Li and ZhenXu Bai.

I thank my parents Jinglian fan, Tianshu Liu, and my little brother, Ziming Liu, my elder sister, Minmin Fan, and my elder brother Peng Fan for their support and love to me. I would like to thank Zhong Zheng for his love and support to me during we once stayed together.

ABSTRACT

The thesis combines the focus area and challenges of vehicular ad hoc network (VANET), and studies the MAC layer performance optimization and pseudonym management, so as to achieve the reliable, secure and privacy preserving transmission in the VANET. The main contributions of the thesis can be summarized as follows.

We propose a MAC layer attack resistant pseudonym (MARP) scheme to deal with a new MAC layer context-linking attack, guarantee the location privacy and fulfill the transmission requirements in the VANET. We present a three-layer software defined architecture for the Internet of Vehicles. Facilitated by the architecture, a MAC layer aware pseudonym (MAP) scheme is proposed to coordinate both the channel accessing and pseudonym change for the vehicles by the local RSU cloud. Security analysis and simulation results show that the MAP scheme can substantially outperform the compared existing methods in terms of the pseudonym distribution efficiency, location privacy and the reliability of safety message transmission. Lastly, a fully uncoordinated approach to change pseudonyms is proposed for distributed networks, where each node uses a pseudonym until its expiration and then changes after a random delay with no interaction with other nodes in the network. We use both theoretical analysis and simulation results to prove that, the k-anonymity can be achieved at a negligible throughput loss in the case of large networks.

Contents

List of Figures

List of Tables

List of Publications

Journal publications:

- **Z. Liu**, L. Zhang, W. Ni and I. Collings, "Uncoordinated Pseudonym Changes for Privacy Preserving in Distributed Networks," *IEEE Transactions on Mobile Computing*, early access, 2019.

- **Z. Liu**, Z. Liu, L. Zhang and X. Lin, "MARP: A Distributed MAC Layer Attack Resistant Pseudonym Scheme for VANET," *IEEE Transactions on Dependable and Secure Computing*, early access.

- **Z. Liu**, L. Zhu, J. Li, G. Luo and L. Zhang, "A MAC layer aware pseudonym (MAP) scheme for the software defined Internet of vehicles," *China Communications*, vol. 15, no. 9, pp. 200-214, Sept. 2018.

- L. Zhang, S. Jia, **Z. Liu**, et. al., "Bus-Ads: Bus Trajectory-Based Advertisement Distribution in VANETs Using Coalition Formation Games," *IEEE Systems Journal*, vol. 11, no. 3, pp. 1259-1268, Sept. 2017.

- **Z. Liu**, Z. Liu, Z. Meng, et. al. "Implementation and performance measurement of a V2X communication system for vehicle and pedestrian safety", *International Journal of Distributed Sensor Networks*, vol. 12, no. 9, pp, 2016.

- L. Zhang, **Z. Liu**, R. Zou, et al., "A Scalable CSMA and Self-Organizing TDMA

MAC for IEEE 802.11p/1609.x in VANETs", *Wireless Personal Communications, Springer*, Vol. 74, pp. 1197-1212, January, 2014.

Conference publications:

- **Z. Liu**, L. Zhang, W. Ni and I. B. Collings, "A Cross-Layer MAC Aware Pseudonym (MAP) Scheme for the VANET," in *2018 IEEE Global Communications Conference (GLOBECOM)*, Abu Dhabi, United Arab Emirates, 2018, pp. 1-6.

- G. Luo, S. Jia, **Z. Liu**, et. al., "sdnMAC: A software defined networking based MAC protocol in VANETs," *2016 IEEE/ACM 24th International Symposium on Quality of Service (IWQoS)*, Beijing, 2016, pp. 1-2.

- S. Jia, **Z. Liu**, K. Zhu, et. al., "Bus-Ads: Bus-based priced advertising in VANETs using coalition formation game," *2015 IEEE International Conference on Communications (ICC)*, London, 2015, pp. 3628-3633.

- R. Zou, **Z. Liu**, L. Zhang and M. Kamil, "A near collision free reservation based MAC protocol for VANETs," *2014 IEEE Wireless Communications and Networking Conference (WCNC)*, Istanbul, 2014, pp. 1538-1543.

- **Z. Liu**, R. Zou, H. Zhang and Lin Zhang, "An adaptive and reliable integrated MAC mechanism for VANETs," *2013 16th International Symposium on Wireless Personal Multimedia Communications (WPMC)*, Atlantic City, NJ, 2013, pp. 1-5.

- **Z. Liu**, T. Liang, J. Guo and L. Zhang, "Priority-Based Access for DSRC and 802.11p Vehicular Safety Communication," *2012 International Conference on Connected Vehicles and Expo (ICCVE)*, Beijing, 2012, pp. 103-107.

Chapter 1

Introduction

The main application scenario of the work presented in the thesis is the vehicular ad hoc nework (VANET). This chapter introduces the basic concept and the standardization progress of VANET, and the contribution and framework of the thesis.

1.1 Vehicular Ad Hoc Networks (VANETs)

To smooth the transportation system and reduce vehicle accidents, the Intelligent Transportation System (ITS) has been actively promoted in recent years. Vehicular Ad hoc network (VANET) has been designed to play a cornerstone role in the development of ITS. By enabling the vehicle-to-vehicle (V2V), vehicle-to-roadside unit (V2R), and vehicle-to-cloud (V2C), i.e., vehicle-to-everything (V2X) communications [1–3], the VANET can improve traffic safety, efficiency and traveling comfort by supporting traffic safety, transportation efficiency, Internet services and other related applications. As the premise of realizing various services, how to guarantee the transmission quality of different types of messages is one of the key challenges in the VANET.

VANETs have some unique network characteristics compared to the traditional mobile ad hoc networks (MANETs). First, vehicle nodes have high mobility, resulting in dramatic

changes in the network topology [4–6]. Second, the density of vehicles varies greatly in different spatial and time scenarios, resulting in the variation of the traffic load in the network, and high requirements for the scalability of the network. The nodes usually transmit the safety messages in the shared wireless channel in a broadcast manner, which brings hidden terminal and transmission collision problem [7]. The medium access control (MAC) layer is designed to meet the key requirements of data transmission quality of services (QoS), while the characteristics of the VANETs have brought great challenges and improvement potential in the design and optimization of the MAC layer protocols. At present, the design of efficient MAC layer protocol for vehicle networking has become a focused research topic.

The security and privacy protection of the data transmission in the VANETs is closely related to the personal safety and property of the drivers, and even the security of the society. The openness and high mobility of the vehicular environment makes the network vulnerable to the attacks and security threats [8–10]. For example, a malicious node or criminal group can send forged messages to the vehicles or traffic management infrastructures, which may subsequently cause the traffic safety or congestion problem. In addition, the vehicles need to periodically broadcast "beacons" or basic safety messages (BSMs) including their locations, speed and acceleration/deceleration every 100-300ms to the surrounding vehicles and RSUs, in order to create cooperative neighborhood awareness and facilitate the safety applications. Unfortunately, this would also provide information for the adversaries to track the vehicles and hence compromise their location privacy. Therefore, it is of practical significance to protect location privacy in the VANETs. Only under the premise of security and privacy protection for the communication entities, the development and deployment of VANETs can be recognized and accepted by the public.

1.2 Standardization of the VANETs

Motivated by the importance of vehicular communications, in 1999, the United States Federal Communication Commission (FCC) has allocated 75 MHz radio spectrum in the 5.9 GHz band for the dedicated short range communications (DSRC) to be exclusively used for the V2X communication. The DSRC spectrum is divided into seven 10 MHz channels, which include one control channel for the transmission of safety and control messages, and six service channels for safety and non-safety related applications. The European Telecommunications Standards Institute (ETSI) also has allocated 30 MHz in the 5.9 GHz band for the ITS applications.

Since 2004, in the United States, the DSRC standardization work has been carried out by the IEEE 802.11p and 1609 working groups. In the IEEE 802.11p, the specific physical layer operation, node transmission process and channel accessing mechanism are specified. The IEEE 1609 series protocols mainly focus on the upper layer operation. The IEEE 802.11p and 1609.x together form the wireless access in vehicular environment (WAVE) stack, as shown in Fig. 1.1. IEEE 1609.0 [11] describes the overall architecture of the WAVE protocol stack. IEEE 1609.1 [12] defines the format and data storage format of the control messages in the WAVE system, specifies the control flow of resource management, and defines standard interfaces for application registration and management. IEEE 1609.2 [13] defines the certificate management system for the security-related services and related operations including signature, encryption, and authentication. IEEE 1609.3 [14] specifies the network layer and transport layer of the communication protocol. IEEE 1609.4 [15] works on the upper layer of the IEEE 802.11p, which specifies the multi-channel operation for the MAC layer. The IEEE working group completed the basic standardization work of DSRC in 2010 and made a series of supplementary explanations in the following years. The DSRC/WAVE has become one of the most widely used standard for vehicular communication.

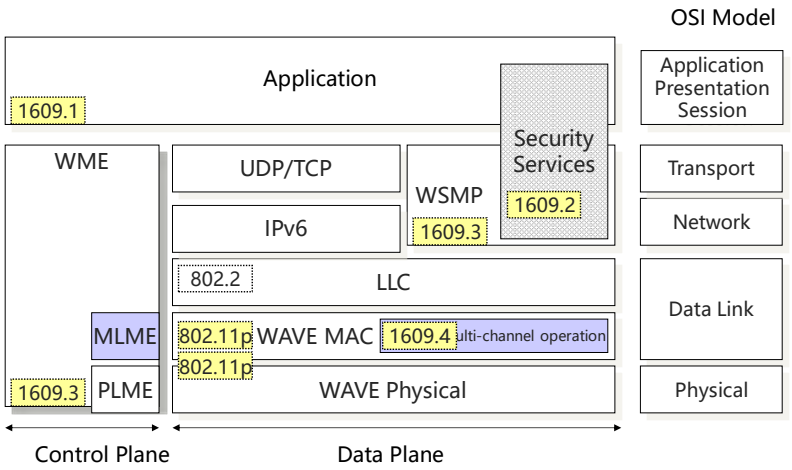


Figure 1.1: DSRC/WAVE protocol stack

In 2014, the ITS standardization group of ETSI, TC ITS, published the Collaborative Intelligent Transportation System (C-ITS) [16, 17]. The communication architecture of C-ITS is described in [18]. The underlying communication standard ITS-G5 [19] adopts the same protocol as IEEE 802.11p. In the network and transport layer, GeoNetworking [20, 21] defines the single-hop and multi-hop communication to deliver messages within a specific geographic area. For data security and privacy protection, C-ITS defines the similar public key infrastructure (PKI) system to the IEEE 1609.2 [22, 23]. The European Commission recognizes that C-ITS has the potential to improve road traffic safety and efficiency, and has been developing a number of projects to improve traffic safety by using the vehicular communication technologies for more than a decade. In addition to the development of the DSRC standard, with the advancement of device-to-device (D2D) and 5G communication technologies, the LTE-V has been vigorously promoted and developed in recent years to support the vehicular communication. In 2016, the 3GPP launched the completion of the first edition for the LTE-V2X communication standardization [24].

1.3 Thesis Objectives and Outline

Motivated by the urgent requirement of the reliable, low-latency and privacy preserving transmission in VANETs, this thesis has the following objectives:

1. To introduce a distributed pseudonym scheme that works consistently with the proposed hybrid MAC layer protocol, which can provide a reliable one hop broadcast service for the safety messages, and preserve the location privacy for the vehicles in VANETs.

2. To present a software defined network (SDN) based resource coordination scheme, which exploits the local RSU cloud to coordinate the vehicles to broadcast the safety messages in a contention-free manner, and change their pseudonyms dynamically to improve their location privacy in VANETs.

3. To develop a decentralized pseudonym change approach for the distributed networks, such as the VANETs, wireless sensor networks (WSNs) and Internet of Things (IoTs). The approach enables the nodes in the network to change pseudonyms in an uncoordinated manner. We prove that the identity privacy can be guaranteed with negligible throughput loss when the number of nodes is large in the network.

The rest of the thesis is organized as follows. Chapter 2 describes the related work under consideration. Chapter 3 introduces the MAC layer attack resistant pseudonym scheme, called MARP, and proposes an analytical model to quantify the location privacy achieved in the scheme. Chapter 4 firstly presents a three-layer software defined architecture for the Internet of Vehicles (SDIV). Facilitated by the SDIV architecture, a MAP scheme is proposed for the reliable and privacy preserving transmission. The RSU clouds coordinate the transmission slot reservation for the vehicles, according to the utilization status of the slots, and make the pseudonym change decision for the vehicles by deter-

mining whether the vehicles can achieve privacy gains through changing pseudonyms. Security analysis and simulation results show that the MAP scheme can substantially outperform the compared existing methods in terms of the pseudonym distribution efficiency, location privacy and the reliability of safety message transmission. Chapter 5 first explains the motivation for the design of the fully uncoordinated pseudonym change approach, and then develops an analytical model to evaluate the privacy performance and the throughput loss incurred by pseudonym change in the distributed networks. Corroborated by simulations, the accuracy of the analytical model improves, as the number of nodes increases. The simulation results also show that the uncoordinated approach can substantially outperform the compared existing methods in terms of privacy preservation, and incur negligible throughput loss in the case of large networks. Finally, Chapter 6 concludes the thesis and suggests some further research topics.

Chapter 2

Transmission Requirements and Key Technologies in VANETs

VANETs can improve traffic safety, efficiency and traveling comfort by supporting traffic safety, transportation efficiency, Internet services and other related applications. Traffic safety related applications are considered to be the most important in the vehicular networks. They have different transmission characteristics and requirements to the traffic flow in the conventional wireless networks, and they are also the most typical applications to represent the value of the vehicular networks. The safety related messages have stringent high requirements of the QoS, and location privacy preservation for the transmission. This chapter is intended to analyze the transmission requirements, especially for the safety related messages, the key technologies and the design challenges to realize these requirements in the VANETs.

The content of this chapter is organized as follows. First, the transmission requirements in the VANETs are outlined in section 2.1, and the related technologies to meet these requirements are described in section 2.2 and 2.3. The challenge of the existing work to meet the transmission requirements for the VANETs is discussed in section 2.4,

draws forth the motivation of our work that will be presented in the subsequent chapters.

2.1 Performance, Reliability and Security Issues in VANETs

In VANETs, the transmission requirements of the non-safety related messages are similar to that in the traditional mobile network, this section mainly considers the transmission requirements of the safety related messages. The transmission requirements for the non-safety messages and the management messages related to the security and privacy preserving are introduced briefly.

The typical service-oriented traffic flow, such as the text, voice, video and other categories generally have a large volume of data to transmit, and they have less requirement in the real-time transmission, or the transmission delay can be reduced by the caching approach. Their transmission has a higher tolerance to the packet loss and error rate. The transmission method is usually point-to-point, and the packets are encrypted to ensure the data integrity and confidentiality. However, the transmission characteristics and requirements between the non-safety and safety messages vary considerably in the vehicular networks. The main transmission requirements and features of the safety messages mainly include the following aspects.

1. One-to-many transmission method: the safety messages in the VANETs are generally broadcasted by each vehicles to all the surrounding vehicles for cooperative awareness. It defines that each vehicle needs to generate the BSMs or cooperative aware message (CAMs) at a frequency of up to 10Hz, and broadcast them to the nearby nodes within the lifetime of the messages. The emergency safety messages are generated by event-driven, and required to broadcast via multi-hop propagation to notify all the nodes within a specific geographic area.

2. Low-latency requirement: the safety messages are generated frequently and have a very short lifetime. In order to ensure the usefulness of the safety messages in the highly changing vehicular networks, the safety message has a very stringent requirement in the real-time transmission. Generally, each message needs to be transmitted before the new safety message delivered to the buffer in the MAC layer, otherwise, it would be discarded. For the periodic safety messages, the delay requirement is generally within 100ms.

3. Highly reliable transmission: to ensure the availability of the safety applications, the transmission of the safety messages is expected to be highly reliable. For the emergency safety messages, the packet delivery rate is usually required to be above 99%.

4. High priority requirement: the significant importance of the safety applications leads to the high priority of the safety messages. The safety messages are typically delivered in the CCH, and has the highest priority to contend to transmit within the queues of each vehicle.

The above four requirements are mainly related to the QoS aspect of data transmission. The MAC layer controls the channel accessing procedure of the vehicles to transmit, thus, the design of the MAC protocol plays a crucial role in supporting the reliable and real-time transmission of the safety messages in VANETs.

5, Data security and privacy preservation: the openness and broadcast nature of the VANETs brings security and privacy threats to the safety message transmission. Thus, the transmission is expected to fulfill the security and privacy requirements. There has been a consensus that the requirements mainly consist of the following aspects [8, 25–27] in VANETs.

- Confidentiality: the messages with confidentiality requirement are expected to be protected from being accessed by the unauthorized third parties.

- Authentication and identification: the mutual authentication must be included be-

tween the communication entities, including vehicles, RSUs and the central authorities, to verify the identities and legality of the transmitting nodes, and avoid multiple attacks, such as forgery attacks, replay attacks and Sybil attacks.

- Transmission integrity: the identity of the transmitting node must be verified, and the message should be signed, so as to ensure the authenticity and integrity of the messages.

- Location privacy: to protect the real identity of the drivers, and location privacy of the nodes, it is expected to use the pseudonyms for the nodes in VANETs. In order to ensure the location privacy, the new and old pseudonyms of the nodes cannot be associated with each other, and the real identity of the nodes and their location information cannot be correlated by the eavesdropper.

- Non-repudiation and accountability [28]: the authentication of an entity must be unique in the whole network, and the nodes cannot deny their responsibilities in the communication process.

- Conditional anonymity: since the vehicle node usually corresponds to its driver, whose identity should be resolved to the law authorities in case of the law enforcement. Therefore, conditional anonymity is usually expected for VANETs.

2.2 MAC protocols in VANETs

Each safety message should be successfully delivered to the surrounding vehicles and RSUs in real-time and with high reliability, which is one of the main functions of a MAC protocol in VANETs. There have been various MAC protocols proposed for VANETs. In this section, we presents a brief review of the MAC protocols and the multi-channel operation for the VANETs.

2.2.1 The Categories of the MAC Protocols in VANETs

As shown in Fig. 2.1, the MAC protocols in VANETs can be classified into the contention-based, contention-free and the hybrid schemes, according to the accessing mode of the vehicles. The classification of the MAC protocols in VANETs also considers the adaptive multi-channel switching [15, 29–34], the support of multi-priority transmission [33, 35–37], and the radio resource allocation approach [38–41].

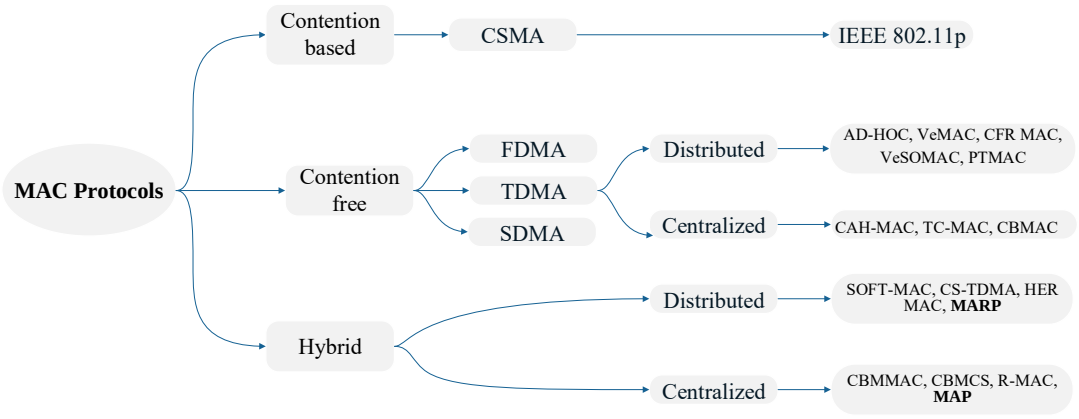


Figure 2.1: The categories of the MAC protocols in VANETs

- Contention based protocols: generally the CSMA approach is applied for the vehicles contending to access the channel in a distributed manner, such as the most applied IEEE 802.11p protocol. There have been some work to prove that the protocol cannot guarantee the QoS requirements of the safety messages. In order to improve the scalability of IEEE 802.11p, some work has made adaptive adjustment of the parameters, including physical layer carrier sensing threshold, contention window [42, 43] and the transmission power [43] in the VANETs. However, when the traffic density becomes high, the transmission collision problem becomes inevitable, thereby degrading the transmission performance.

- Contention-free protocols: various contention-free protocols have been proposed for

VANETs, based either on TDMA [39, 44–48], frequency division multiple access (FDMA) [49–51] or space division multiple access (SDMA) [52, 53]. These protocols allow each vehicle to access a channel in a specific time slot, frequency block, or geographically-divided radio resource. By allowing the vehicles to use disjoint resources, the transmission collisions and hidden terminal problem can be alleviated. The SDMA-based schemes was proposed by some researchers in the early years during the development of VANETs. The SDMA based scheme defines which time slots a vehicle is allowed to access based on its current location [46]. Thus, the main drawback of the SDMA approaches is the lack of flexibility and scalability to deal with the rapid change vehicle densities and speeds. FDMA-based MAC protocol works in the similar way to the SDMA approach, by assigning the frequency block the vehicles, so as to deal with transmission interference and noise. The TDMA-based protocols have been promising to provide bounded transmission and alleviated transmission collisions by allocating different transmission slots to the adjacent nodes for VANETs. The periodical frame structure in the MAC layer is adaptive to the periodic transmission of the safety messages.

- Hybrid protocols: the hybrid protocols combine multiple accessing methods together for the nodes to access the channel in VANETs. The work [54] provided a comprehensive review of the hybrid MAC mechanism in VANETs. Abdalla *et al.* proposed a MAC protocol combining CSMA, SDMA, OFDMA and TDMA in [49]. The scheme divides the street into different cells and allocates different subcarriers for each cell. The vehicle nodes reserve the corresponding subcarriers and transmission slots according to its real-time location, so as to avoid transmission interference between adjacent vehicles. In [29], we proposed a coordinated access multi-channel protocol combining TDMA and CSMA, called CS-TDMA. CS-TDMA effectively combines the flexibility of CSMA with the advantages of TDMA to avoid collisions

and provide bounded transmission delays. In this thesis, the MARP scheme proposed in Chapter 3 combines CSMA and distributed TDMA for the transmission of various types of messages in VANETs, and the MAP scheme proposed in Chapter 4 applies the centralized TDMA approach for the transmission in the control channel, and the coordinated CSMA for the transmission in the service channel.

2.2.2 Multi-channel Operation

IEEE 1609.4 [15] defines the multi-channel operation in VANETs. Each nodes needs to synchronize the time using the GPS signal or the control messages from the RSUs. The time is divided into synchronized intervals (SIs) with constant duration, which usually is 100ms. Each SI consists of a control channel interval (CCHI) and service channel interval (SCHI). During the CCHI, the vehicle needs to transmit and receive in the CCH, and then selectively switches to the SCH during the SCHI. A 4ms guard interval is also introduced at the start of each channel interval, so as to avoid inter-symbol interference (ISI). The CCH is reserved for the safety and control messages, while the SCH is used for the safety and non-safety applications.

IEEE 1609.4 defines four different channel switching modes for VANETs, as shown in Fig. 2.2. In the continuous access mode, the vehicle nodes will continuously tuning to the control channel, so that the control messages and most important safety messages can be exchanged as much as possible, at the expense of other services in the SCH. The alternating service channel access is also the default switching mode of IEEE 1609.4, with the duty cycle being the same between the CCHI and SCHI, i.e., the vehicle switches the channel every 50ms. In the immediate SCH access mode, the vehicle communication device can immediately switch to the SCH for data transmission when the control channel is idle, and then it switches to the CCH at the beginning of the next SI. The extended SCH access mode allows each vehicle to continuously transmit in the SCH without switching

to the control channel at each SI. In this mode, the vehicle obtains more stable data transmission and throughput with sacrificing the safety message transmission in the CCH.

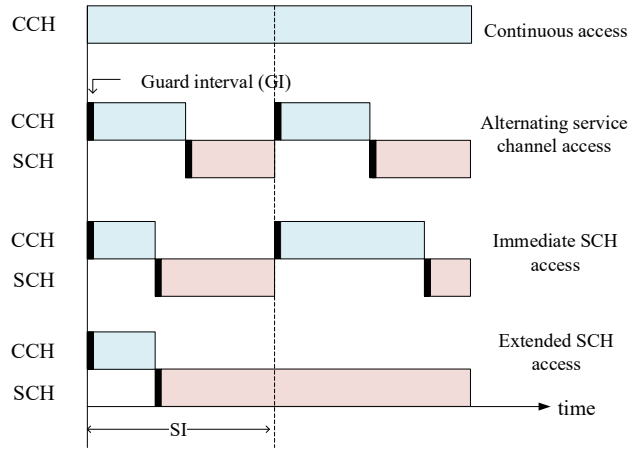


Figure 2.2: Multi-channel switching modes under IEEE 1609.4

2.3 Pseudonym Scheme

Pseudonym schemes have been generally applied, to ensure data security and privacy preservation in VANETs. The pseudonym schemes and pseudonym changing strategies have been comprehensively reviewed in [55, 56] for the VANETs. As shown in Fig. 2.3, the lifecycle of a pseudonym mainly includes pseudonym issuance, use, change, resolution and revocation [55].

2.3.1 Pseudonym Issuance

Before each vehicle participates in the secure and privacy preserving communication, some security related parameters need to be configured for the vehicle and related entities. Each vehicle typically uses the electronic license plate (ELP) obtained from the vehicle registration authority as its real identity, and a long-time certificate that allows

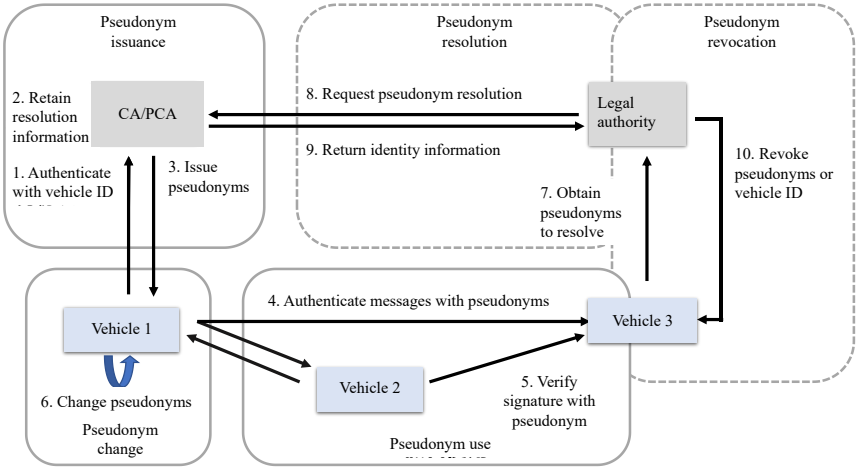


Figure 2.3: The pseudonym lifecycle in VANETs

unambiguously resolution of a vehicle. The pseudonym issuance is usually carried out by the third trusted parities (TTPs) for the nodes in the VANETs. The self-issuance method is also proposed [57, 58] to let the vehicles issue pseudonyms by themselves. The self-issuance approach can reduce the pseudonym management overhead, however, it is difficult to realize the pseudonym resolution and revocation. In this thesis, the proposed schemes are based on the PKI approach to let the TTPs to coordinate the pseudonym issuance in the VANETs.

When a pseudonym is issued, it is usually given an expiration date or validation duration for the use, so that the number of pseudonyms that each vehicle can access at the same time can be limited. By this way, the Sybil attack can be avoided. The pseudonym distribution is usually managed by multiple entities, including the central authority (CA), pseudonym certificate authority (PCA) and the RSUs. By splitting the responsibilities for pseudonym management between multiple entities, the inner-attack that may compromise the privacy of vehicles can be avoided.

2.3.2 Pseudonym Use

Once the vehicle has obtained a pseudonym, it can participate in the V2X communications. In order to ensure that the vehicle's pseudonym and related certificates cannot be accessed by other entities or malicious nodes, the relevant security parameters are generally stored in the vehicle's hardware security module (HSM) or the on-board unit's tamper proof device (TPD). The use of pseudonyms is generally associated with the signing of the transmitted messages, and the verification of the received messages. Typically, the signature is generated by the asymmetric signature algorithm such as the Elliptic Curve Digital Signature Algorithm (ECDSA) [59], or the message authentication code [60]. By using the pseudonyms, the source authentication and message integrity can be achieved.

2.3.3 Pseudonym Change

Each pseudonym has a minimum lifetime to ensure the stable transmission and keep a moderate cost for the pseudonym management. Meanwhile, the short-lived pseudonyms have to change frequently so as to break the temporal and spatial link-ability of the vehicles. The ETSI standard [61] recommends that pseudonyms should be changed every 5 minutes, while the SAE J2735 [62] suggests the value to be 120s or 1 km. Simply changing the pseudonym alone or in sparse scenarios with only few neighbors is still not sufficient to confuse an adversary. The pseudonym change should be carried out by multiple vehicles simultaneously to confuse the adversary. At the same time, the change of the pseudonym must include all communication layer identifiers, such as IP or MAC addresses, so as to ensure the consistency and validity of the pseudonym change. The most promising schemes in the related literature expect a set of vehicles in the same geographical area, usually called mix-zones, to perform pseudonym changing simultaneously. There have been various strategies proposed for the vehicles to construct mix-zones and changes

pseudonyms together. Based on the decision making approaches, the pseudonym change strategies can be categorized as follows.

- Specific geographic location based pseudonym change: Freudiger et al. first proposed mix-zone [63] based pseudonym change scheme. Before the vehicles change the pseudonyms in the mix-zone, a session key is obtained from the RSU to encrypt all the information they send during the mixed area, making the pseudonym changing process unobservable. Ref. [58] recommended using the social area such as intersections and parking lots as mix-zones. Similar work can be found in [64–66], which optimized the selection of the mix-zone to improve the location privacy of the vehicle nodes.

- Silent period based pseudonym change: when the vehicles decide to change their pseudonyms, a silent period is used to protect them from being continuously tracked. The work [58] proposed that the vehicle enters a silent period and changes the pseudonym when encounters the traffic lights at low speed. Similarly, the authors of [67] proposed a pseudonym change policy in an urban environment where vehicle nodes enter a silent period to change or exchange pseudonyms only when the traffic lights at the intersection turn red. The authors also propose to preserve the location privacy of the nodes by allowing them to transmit obfuscating location and speed information. To make the pseudonym changing process unobservable, the related literature also proposes a scheme for dynamically adjusting the length of the silent period according to the distance between the vehicles. When the distance between the vehicles is smaller, the silent period should be shorter, thereby ensuring the traffic safety in the VANETs.

- Context based pseudonym change: vehicles dynamically change pseudonyms based on certain contexts, such as the geographic location information, vehicle density,

traffic condition, driving state, the cooperation of nearby vehicles, and other spatial-time information. Ref. [68] proposed a traffic-aware pseudonym changing strategy, called TAPCS, for VANETs. In TAPCS, vehicles continuously sense the driving conditions, and change their pseudonyms when encounters traffic jams. However, the scheme requires a large amount of communication overhead to judge the traffic congestion, the start and end of the traffic jam. At the same time, the occurrence of a traffic jam is random, resulting in low pseudonym utilization or the long use time of the same pseudonym. In [69], the authors proposed a scheme for the vehicles to change their pseudonyms by observing the context of the nearby vehicles. When at lease one neighboring vehicle is within the silent period, the node changes its pseudonym after a random silent period.

- Cooperation based pseudonym change: in [70], the author proposes a strategy for the vehicles to perform pseudonym changing based on the cooperation status of the surrounding vehicles. Freudiger et al. pointed out that the pseudonym resource could be a kind of precious resource in the VANETs, so that the vehicles sometimes would be unwilling to cooperate with the surrounding nodes to change the pseudonym. Then they presented a pseudonym change strategy based on the non-cooperative game between the vehicles. In [71], each node is encouraged to change pseudonym by managing their reputation value in cooperating. When the vehicle assists other vehicles in changing pseudonyms, it can increase its reputation value.

2.3.4 Pseudonym Resolution and Revocation

The previous steps involve all participants while the resolution and revocation of pseudonyms are optional, and only relevant to nodes that are misbehaving in the network. The law authority may obtain the pseudonym of an illegal node through behavior monitoring and

detection. It sends a pseudonym request to the relevant trusted party, in order to obtain the real identity of the pseudonym holder. After the law authority verifies the node identity information, if the pseudonym scheme supports traceability, the law authority can perform pseudonym revocation or identity revocation. Although the pseudonym resolution can be carried out directly within one authority parity by searching the database, it could bring potential threats of the internal attacks in the network [72–75]. Schaub [73], Fischer [72], Bißmeyer [75] and other researchers have proposed to separate the responsibilities of the management entities in the role of pseudonym issuance, distribution, resolution, and revocation. The certificate revocation list (CRL) is usually transmitted to notify the relevant entities the revocation information of the misbehaving nodes in the VANETs.

2.3.5 Privacy Metrics

Privacy-preserving evaluation metrics are significant for the evaluation and measurement of location privacy in VANETs. In [76], the authors summarized the definition of the most popular privacy metrics in the VANETs. The anonymity set size [58, 66, 77], entropy [66, 78], adversary tracking probability [76, 79] and the time-to-confusion of the adversary [80] are commonly used to evaluate the performance of pseudonym schemes for the VANETs. Among them, the most commonly used location privacy metric, anonymity set size, measures the number of vehicles in the same geographical area that change pseudonyms at the same time. In [81], the authors developed an analytical model to evaluate the anonymity set size of the random pseudonym change strategy. Privacy entropy [82] is defined based on the Shannon's classical information theory, and used to measure the linking probability distribution of the vehicle' pseudonyms in an anonymity set. Privacy entropy is sometimes used as the privacy level [83] or the tracking uncertainty level [84] of the adversary.

Pseudonym age can be used to decide whether and when to change the pseudonym. To evaluate the age of pseudonyms, Freudiger et al. developed a framework upon the vehicle cooperation probability, traffic mobility, pseudonym cost and the aging rate [83]. As the quality of privacy or the degree of privacy risk strongly depends on the maximum duration adversaries can track a vehicle, time-to-confusion is thereby proposed [80]. It measures the duration an adversary could correctly follow the trace of a targeted vehicle until it could not determine the next sample of the vehicle with sufficient certainty. Intuitively, the time-to-confusion is limited by the pseudonym age when the mix-zones bring enough confusion to the adversary. However, it is not always the case since more intelligent attacks have surfaced. Wernke et al. [85] pointed out that an adversary could jeopardize the unlinkability of pseudonyms using the context linking attacks. Emara et al. [86] described a method to track vehicles based on beacon messages in VANETs. In [87], Bloessl et al. presented a scrambler attack on location privacy based on the physical characteristics of the vehicles. In [88], the authors evaluated the effectiveness and the efficiency of different pseudonym change strategies and observed that the pseudonym changing interval of about 45s were sufficient in the specific urban scenario, and no pseudonym strategies provided satisfactory privacy protection in the highway scenario.

2.3.6 Communication Overhead and Cross-layer Effects

Most of the existing pseudonym schemes ignored the difficulty of implementing the pseudonym schemes, and the impact of the generation, management, change and revocation of pseudonyms on the transmission performance in VANETs. In MixGroup [66], group signature is used to encrypt the negotiation messages during the pseudonym exchange process for the nodes. However, each time a pseudonym is exchanged, the vehicle needs to transmit the encrypted messages for four rounds, causing substantial communication overhead, and significant degradation of other messages' transmission. In TAPCS [68], the silent period

is adopted to change the pseudonyms during a traffic jam. It also requires heavy communication expenses in the discovery of the traffic jam, interaction for the initiation and the termination of the silent period.

Changing pseudonyms can also affect the transmission performance [89]. However, existing analyses have been based on centralized pseudonym changes, or focused on the effect to the application layer performance with little consideration on the QoS or throughput in the networks. K. Emara [90] studied the effect of pseudonym changes on the safety applications, such as forward collision warning and lane change warning in vehicular networks. In [91], the pseudonym management cost was analyzed in terms of communication delay. In [92], the impact of pseudonym change and silent period was evaluated on an intersection collision avoidance (ICA) system for driving assistance.

2.4 Challenges to Realize Reliable, Secure and Privacy Preserving Transmission

Efficient MAC protocol plays a key role in meeting the reliability and real-time transmission in the VANETs. There are still several key challenges need to be solved in the existing work [93, 94].

- It is difficult to achieve the flexibility and effectiveness of the MAC schemes in the VANETs with high mobility. The contention based scheme is flexible in channel accessing, however, it suffers from serious collision problem. The contention free scheme often stipulates the number of time slots in each frame, thus it lacks of adaptability and scalablity for the network with highly changing traffic densities. In the early stage of the deployment of VANETs, the MAC protocol must support the nodes to access the channel in a fairly distributed approach, due to the lack of infrastructures. At the same time, the distributed channel accessing must ensure

that communication overhead does not consume too much bandwidth and provides bounded transmission delay for the messages.

- The hidden terminal issue needs to be avoided in the MAC protocol. The problem arises not only in the CSMA based protocols, but also in the TDMA protocols, caused by the high mobility of the vehicles. The nodes that use the same time slot may be far away from each other at one moment, and become hidden terminals and incur merge collisions after driving for a short interval.

- The MAC protocol needs to support different QoS requirement for various applications. It needs to provide high-reliability and low-latency transmission services for the safety messages, while ensuring the throughput of other services. In general, the safety messages have the higher priority to access the channel, however, the pseudonym changing may interfere with the vehicles' transmission.

- Multi-channel operation needs to be defined by the MAC protocol. In general, safety messages and control messages that require high QoS need to be transmitted in the CCH. Non-safety messages that require a large throughput are transmitted in the SCH. Generally, static channel switching cannot meet the requirements of data transmission in the dynamic vehicular environment. For example, in a blocked section, a 50ms control channel interval may be not enough for all the nodes to complete their transmission fo the safety messages. Whereas in the case of sparse traffic density, the control channel resource may be wasted. Therefore, the adaptive channel switching is expected to be supported so as to improve the channel utilization.

The pseudonym scheme plays a key role in the privacy protection of nodes in VANETs. At present, there have been some effective pseudonym schemes proposed. By monitoring the environments and the privacy quality of vehicles, the pseudonym change is dynamically

performed. However, existing schemes still have the following challenges to solve, which are mainly relevant to the interaction between the pseudonym scheme and the MAC layer performance.

- Pseudonym scheme needs to achieve non-linkability. In order to keep the vehicles' pseudonym from linking by the adversary, pseudonym change needs to be synchronized at all the communication layers. However, it is a threat that may cause information confusion during the the routing process [89]. Therefore, the pseudonym change strategy needs to take both the location privacy and the performance of transmission into consideration, and the performance evaluation and measurement should not limit to the privacy related metrics.

- Pseudonym schemes need to consider the impact on other types of data transmission. The age of the pseudonym directly affects the privacy quality of the node. Generally, the highly frequent for the nodes to change pseudonyms, the higher privacy quality can be achieved. However, the frequently pseudonym changing cause the vehicles to consume more pseudonyms. It not only increases the amount of data for the CRL distribution, and also increase the communication expense for the pseudonym distribution and pseudonym change negotiation. Therefore, the management of pseudonyms must consider the communication overhead and their impact on the safety message transmission.

- Pseudonym schemes must be coordinated with the design of the MAC layer protocol. The existing work designs the pseudonym schemes and MAC protocols separately. The MAC protocol often ignores the support for the transmission of the pseudonym management related messages, while some pseudonym schemes cannot be adapted to the specific MAC protocols or ignore the impact on the safety message transmission. It is feasible for the adversary of eavesdropper to observe the activity of the nodes

from all the layers in the open vehicular environment. Thus, semantic information to link the pseudonyms such as the MAC layer transmission pattern of the vehicles, is easy to acquire by the adversary. The design of the pseudonym scheme must deal with the linking attacks from different layers that may taken by the adversary.

The QoS and quality of privacy for the data transmission requires efficient MAC protocol and pseudonym scheme to allocate the channel resources and pseudonym resources consistently. The pseudonym change process should be efficient, and it cannot occupy too much channel resources. The pseudonym change should be carried out in all the layers, and deal with the context linking attacks that could be carried out by observing the MAC layer activities. At the same time, when designing the MAC protocol, it is necessary to consider the transmission of the pseudonym management related messages, so as to ensure that the pseudonym scheme can work smoothly. Based on the above analysis, the thesis takes the joint design of the MAC layer and pseudonym schemes as the main research work. The thesis will attempt to solve the above research challenges in the following chapters.

Chapter 3

A Distributed MAC Layer Attack Resistant Pseudonym Scheme

3.1 Introduction

Vehicular ad hoc networks (VANETs) are expected to enhance the transportation safety and efficiency, as well as provide infotainment services. Based on the Dedicated Short Range Communication (DSRC), vehicles can periodically broadcast the BSMs including their locations, speed and acceleration/deceleration every 100-300ms to the surrounding vehicles and roadside units (RSUs). BSMs play a vital role in creating cooperative neighborhood awareness and facilitating the safety applications. However, due to the broadcast nature of the VANETs, V2V and V2R communications are vulnerable to the long-term and large scale tracking from adversaries. Thus, the identities and locations that reveal and link vehicles or drivers must be protected to achieve the location privacy.

Recently, many security and privacy-preserving schemes and protocols have been developed. For example, standardization bodies such as IEEE 1609.2 [13] and ETSI [61], have proposed the public key infrastructure (PKI) based cryptography to protect vehicu-

lar communications. To ensure the communication authenticity, integrity, identity privacy and non-repudiation, pseudonyms certificates are used with the corresponding short-lived private keys to sign the messages. Furthermore, the pseudonyms need to change frequently and suitably [55], so that the messages originating from the same vehicle are unlinkable, so as to protect the location privacy.

However, a new MAC layer linking attack can still be launched to compromise the unlinkability in the pseudonym schemes. The dilemma between location privacy and transmission efficiency arises in the VANETs. This is because previous work considers the design of pseudonym scheme for location privacy, and the MAC layer protocols for transmission efficiency separately. Adversaries could utilize the MAC layer activity time as the "fingerprint" of vehicles to infer their identities regardless of the changing pseudonyms. The design of the pseudonym schemes also lacks of consideration on the influence of pseudonym changing on the transmission performance.

Various MAC layer protocols have been proposed for VANETs. The DSRC has employed IEEE 802.11p [95] for the MAC layer operation, in which BSMs broadcasting is based on CSMA/CA. However, prior studies [96, 97] have shown that the CSMA based MAC protocol is only adaptable to low-to-moderate channel loads. Meanwhile, a context-linking attack is possible to launch using the transmission cyclicality of the BSMs from the CSMA based MAC operation. Many researchers prefer to Time Division Multiple Access (TDMA) based mechanisms for VANETs, such as VeMAC [46] and CFR MAC [98], etc. In TDMA, time is divided into synchronized frames that consist of a number of time slots. Each vehicle reserves one or more time slots to transmit BSMs periodically and keeps silent during other slots that have been reserved by its neighbors. Unfortunately, the more coordination of a MAC protocol, the more likely to facilitate the context linking attack. Adversaries could link the new pseudonym with the old ones by monitoring the slot utilization information of a targeted vehicle.

To illustrate the MAC layer context linking attack, an example is shown in Fig. 3.1. The vehicles access the wireless medium to broadcast the BSMs periodically. At the first CCHI of frame $\mathcal{F}_1$, the vehicles of Bob, Alice and Grey use the pseudonyms x_1, y_1 and z_1 to broadcast BSMs at time slot t_2, t_4 and t_6 (either the virtual time slots in CSMA or the synchronized time slots in TDMA), respectively. They switch to use the new pseudonyms x_2, y_2 and z_2 at frame $\mathcal{F}_2$ simultaneously. Thus, they become indistinguishable in the pseudonym changing at the upper layer, e.g., application layer. However, their driving trajectories can be monitored continuously based on the BSMs transmitted at the specific time instants from the MAC layer. The transmission cyclicality can be used to link the new pseudonyms with the old ones, and even correlate with the driving route and sensitive locations (home, working place, *etc.*) together to de-anonymize Alice, Bob and Grey.

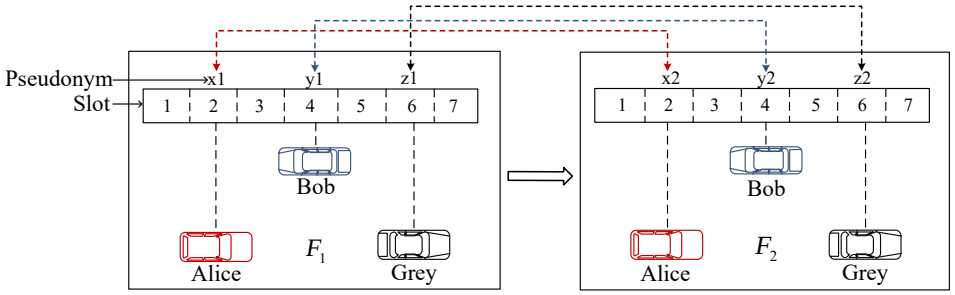


Figure 3.1: The MAC layer context-linking attack using vehicles' transmission pattern

The pseudonym schemes should be designed collaboratively with the message transmission as they influence each other. Lefèvre *et al.* [92] take the road intersection as the traffic scenario to analyze the impact of the privacy strategies on intersection collision avoidance (ICA) system. They draw a conclusion that the requirements of safety applications should be taken into account when designing pseudonym changing strategies. Fonseca *et al.* [99] suggest that the pseudonyms must change consistently across multiple layers, i.e., the IP and MAC address must be hidden or changed simultaneously with the pseudonyms to avoid trivial linking from other identifiers. However, it still lacks a

scheme to resolve the new MAC layer attack caused by the inconsistency between channel accessing and pseudonym changing. The separated design of pseudonym schemes rarely consider the communication overhead and impact on the MAC layer performance, and some are even unsuitable to work under the coordinated based MAC schemes. The problem of guaranteeing location privacy while achieving transmission efficiency needs to be addressed collaboratively for the VANETs. To overcome these challenges, this chapter proposes a MAC layer Attack Resistant Pseudonym (MARP) scheme for the reliable and privacy preserving transmission in the VANET. Specifically, the main contributions of the chapter are listed as follows.

- A new MAC layer context linking attack is identified, which is facilitated by the inconsistent operation of the pseudonym changing and periodical MAC layer activities. The attack behavior is analyzed under both CSMA and TDMA based MAC schemes.
- The MARP scheme is proposed to guarantee the location privacy and fulfill the BSM transmission requirements for vehicles in VANET. In MARP, vehicles change their pseudonyms and time slot by cooperatively negotiating with their neighbors to construct mix-zones based on the current pseudonym age. The cooperation only requires one bit in each message. The MARP scheme divides the synchronized frames (i.e., CCHI) to a dedicated transmission period and a flexible schedule period. The dedicated transmission period consists of separated time slots used by vehicles to broadcast BSMs. The vehicle nodes randomly shuffle their transmission slots when changing pseudonyms, so as to keep unlinkable to the adversaries. The schedule period is used for the communication between vehicles and RSUs or trusted authorities (TAs), e.g., to request pseudonyms or receive the certificate revocation list (CRL) from the TA. When a mix-zone is constructed, vehicles becomes indistinguishable by releasing their transmission slots and old pseudonyms simultaneously. By this

way, the unlinkability of the pseudonyms can be achieved to preserve the identity and location privacy of the vehicles.

- An analytical model is proposed to quantify the performance of the MARP scheme in terms of the pseudonym age, anonymity set size and time-to-confusion. Based on the mean field theory, the analytical model presents a general solution for the location privacy measurement regardless of the transmission pattern and vehicle distribution when the vehicle number is large in the network. Extensive simulation results corroborate the analytical model, and verify that MARP can resist the new MAC layer context linking attack. The performance of the MARP scheme is also compared with several existing schemes to show that the MARP scheme can preserve location privacy with sustainable sacrifice on the transmission reliability. The scheme fills the blank of the cross-layer design of pseudonym schemes.

The remainder of this chapter is organized as follows. Section 3.2 presents the preliminaries of the paper and the MAC layer context linking attack on the location privacy in VANETs. The MARP scheme is proposed in section 3.3. Subsequently, the analytical model to evaluate the performance of the MARP scheme is developed in 3.4. The analytical results are presented and verified by the simulation results in section 3.5. Finally, section 3.6 concludes the chapter.

3.2 Preliminaries

In this section, the assumptions made throughout the chapter is introduced. A list of notations used in this chapter is shown in Table 3.1.

Table 3.1: Notation and Description in Chapter 3

Notations	Description
V_i	A representative vehicle
R	The transmission range of a vehicle
N_{ts}	The number of time slots in a frame
U_i	The tracking uncertainty of V_i by an adversary
$Cert_i$	The certificate of V_i
PK_i, SK_i	The master key pair of V_i
PID_{ij}	The jth pseudonym of V_i
$Cert_{ij}$	The pseudonym certificate of V_i corresponding to PID_{ij}
$\|\|$	The symbol of the concatenation operation
ts_i	The transmission slot reserved by V_i in the DT
s_i	The slot status of ts_i
τ	The cooperation pseudonym age threshold
ε_p	pseudonym age
$f(z,t)$	PDF of pseudonym age equals to z at time t
$F(z,t)$	CDF of pseudonym age at time t
$\bar{M}(z,t)$	The occupancy measure of the vehicles' pseudonym age of z at time t
t_c	The time-to-confusion of an adversary
$p(c(t))$	Probability of at least one neighbor cooperates at time t
$c(t)$	Probability of any neighbor cooperates at time t
n_c	The average anonymity set size in a mix-zone
ts_{min}	The minimum silent period
ts_{max}	The maximum silent period
λ	The average arrival rate of vehicles
ρ	The average traffic density on the roads

3.2.1 Network Model

The MARP scheme is designed based on the certificate management architecture SCMS [74]. As shown in Fig. 3.2, the SCMS has defined the duties of multiple entities in the overall architecture, which has been applied by the IEEE 1609.2 protocol. In SCMS, the root/enrollment certificate authority (CA) is responsible to issue the enrollment certificates that act as passport for each vehicle to request further pseudonym certificates. For any vehicle, before entering the network, the bootstrap and enrollment is required for the OBUs to obtain the certificate of the CA, the enrollment certificate and the information to locate the registration authorities. Before a new trip, each vehicle needs to request new pseudonym certificates that would be active during the trip. The requests are validated, processed and then forwarded by the registration authority to the pseudonym certificate authority (PCA). The PCA is responsible to issue short term pseudonym certificates and the request coordination entity is to ensure that a node does not request more than one set of certificates during a given time period. More detailed description of the whole framework can be referred to the paper [74]. In this work, the certificate management related entities are collectively referred to as the TAs. The considered VANET system mainly consists of three entities, the OBUs equipped in the vehicles, the RSUs along the roads, and the TAs.

- Each vehicle is equipped with an OBU to broadcast the BSMs. The precise driving information and time synchronization are achieved by the global positioning system (GPS) in each vehicle. The transmission power level is assumed to be fixed and known to each other. All the physical channels are symmetric and the communication range is denoted as R.
- RSUs are responsible for communicating as the gateways to deliver some certificate materials such as the certificate revocation list (CRL) from the TA to OBUs and collecting the traffic information from the vehicles within its coverage. Assume

RSUs are not trusted and they access the channel and authenticate messages in the similar approach as vehicles.

- The TAs are responsible for the device bootstrap, pseudonym certificate provision and revocation of malicious entities. Before joining the system, each vehicle and RSU need to register with the TAs. As general, we assume the TAs have the highest security level and are infeasible to be compromised. The TAs may be curious to compromise the location privacy of vehicles, therefore their duties are unlinkable for the pseudonym provisioning.

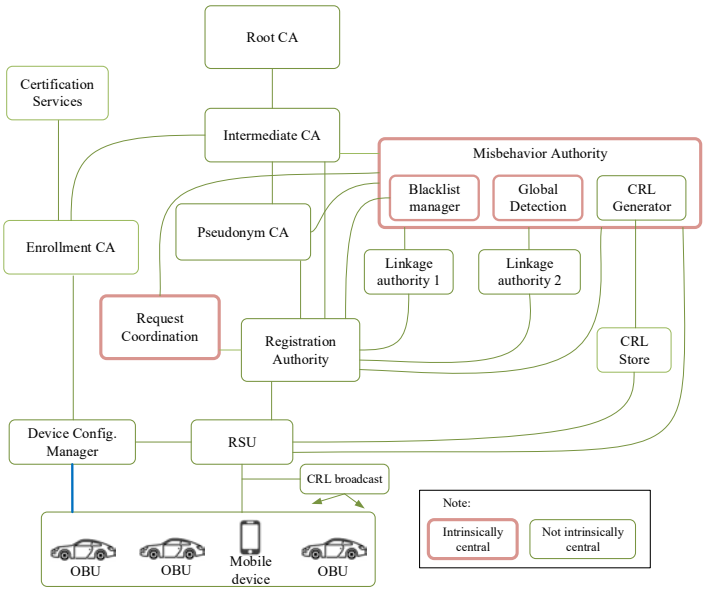


Figure 3.2: The network architecture of the MARP scheme

3.2.2 Threat Model

In this chapter, the Global Passive Adversary (GPA) model [55] is considered, which aims to track the vehicles by eavesdropping their packets, rather than compromising their transmissions. The GPA can carry out the linking attack in attempt to de-anonymize the vehicles by trying to establish the connections between the pseudonyms of each vehicle. In

this paper, the linking and tracking techniques released by the adversary mainly include: 1) collect the safety messages and context information (e.g., MAC address, coordination information or the pseudonyms contained in the messages) to increase the link-ability of the target vehicle. 2) reconstruct the driving paths from the location samples of the targeted vehicles. In this paper, we ignore the vision-based and the physical layer based attacks, and only consider the radio communication based GPA.

3.2.3 The MAC Layer Context Linking Attack

Previous work considers the MAC layer operation and pseudonym schemes separately. Unfortunately, their inconsistent operations facilitate the context-linking attack, by which the GPA can easily track a specific vehicle by analyzing its transmission periodicity of the messages. To formally describe the attack, we start by considering the tracking uncertainty of an adversary in a mix-zone.

The adversary $\mathcal{A}$ observes a set of n vehicles change pseudonyms simultaneously in a mix-zone at the time instant T. $\mathcal{A}$ compares the new set of pseudonyms with the old pseudonyms, and based on the mobility and other context information to predict the most probable linking. The adversary's uncertainty on the linking of a representative vehicle v_i is denoted as U_i and computed as follows [82].

$$U_i(T) = -\sum_{j=1}^{n} p_{j'|i} \log_2(p_{j'|i}), \tag{3.1}$$

where $\sum_{j=1}^{n} p_{j'|i} = 1$, and $p_{j'|i}$ measures the probability that the new pseudonym $PID_j^{'}$ of vehicle v_j is linked with the old pseudonym PID_i of v_i. U_i is upper-bounded by $\log_2 n$, when there is no side information to distinguish v_i from the other vehicles in the anonymity set, i.e., $p_{j'|i} = 1/n$. However, the upper bound of U_i can be acquired only in the case that the unlinkability is guaranteed for the vehicles from every layer, i.e., $\mathcal{A}$ cannot infer the linkability of the pseudonyms from the side-information of the vehicles,

e.g., the transmitted packets, the MAC layer context information and the identifiers. Thus, the information unlinkability of v_i can be represented as $U_i(M) = 1 - \Pr\{M(t_1) \leftrightarrow M(t_2)\}$. Since $\mathcal{A}$ could link v_i from any layer, the final linking uncertainty of $\mathcal{A}$ becomes $U_i = \min\{U(m), \forall m \in \mathcal{M}\}$, where $\mathcal{M}$ denotes the information set of v_i that $\mathcal{A}$ could obtain, and m denotes each information category of $\mathcal{M}$.

When vehicles exploit the CSMA based protocol for the MAC layer operation, we assume that the BSM generation pattern of vehicles is periodical and independent from each other. As shown in Fig. 3.3, BO_i^1, BO_j^1, BO_i^2 and BO_j^2 denotes the backoff duration of V_i and V_j in frames F_1 and F_2 respectively. Then the probability that linking the new pseudonym of V_i with its old pseudonym by the MAC layer attack is equal to the probability that the messages transmitted by V_i do not mix with the other vehicles in the same frame. In this way, the messages of V_i would be taken as broadcasted in a virtual time slot that starts from the time instant V_i contends to access the channel. In this way, the lower bounded probability is calculated as: $p_{i'|i} = (1 - \frac{T_F - T_{BO_i}}{T_F})^{k-1}$, where T_F is the duration of the synchronized frame and T_{BO} is the duration of the service time of the message. Based on the assumptions and analysis of the DSRC based BSM transmission in [97], when $k = 2$, the probability is approximately to be 1 and when $k = 60$ in one-hop range, the data rate is 12Mbps, packet size is 400 bytes and the contention window is 32, the probability $p_{i'|i}$ is still nearly 0.5616 conditioned on the message is successfully transmitted, which is much larger than $\frac{1}{60}$. Consider an extreme case when $k = 200$, the probability would be lower bounded by 0.0614, which is still larger than $\frac{1}{200}$. Therefore the linking uncertainty is much reduced by the MAC layer context information.

To deal with the context linking attack under the CSMA based MAC schemes, we propose that when vehicles generate the BSMs periodically in each frame, they distribute the messages transmission attempts during the frames uniformly. By this means, the transmission time of all the vehicles would be uniformly distributed in each frame, so that

the linking uncertainty would be maximized to the adversaries. However, when exploiting the uniform transmission strategy in the CSMA based protocol, the packet delivery ratio is only about 60% when $k = 60$. Therefore, the transmission efficiency would be further improved by adding some coordination functions, e.g., coordinated time slotted based strategy, for the MAC layer protocol.

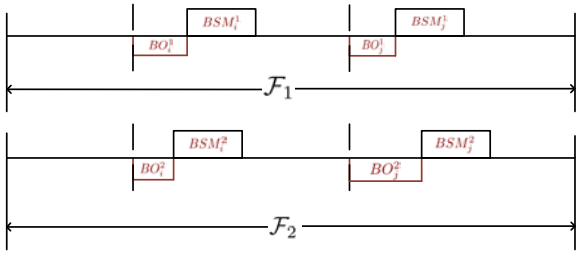


Figure 3.3: The transmission cyclicality of BSMs based on the CSMA in VANETs

When vehicles exploit the coordinated time slotted based protocols, the attack is still possible to launch by collecting the slot utilization pattern. The adversaries can map the slot index with the current pseudonym of each vehicle, denote as $M = \{(s_1, PID_1) \ldots, (s_k, PID_k)\}$, where s_i and PID_i are the slot index and pseudonym of vehicle i respectively. If an adversary finds the pseudonym assigned for the transmissions in a specific slot has been changed, the adversary takes it as the alternation of the pseudonym by the current vehicle, rather than the entrance of a new vehicle. This is because when vehicle changing pseudonym, the time slot is kept continuously in transmitting. While if there is a new entered vehicle, the time slot could be idle for at least one broadcast period for the vehicle to notify itself to its neighbors in TDMA. Meanwhile, the vehicles in the same mix-zone need to apply different slots to avoid transmission collisions. Thus, the new map becomes $M' = \{(s_1, PID_1'), (s_2, PID_2'), \ldots, (s_k, PID_k')\}$. Obviously, if the slot utilization and pseudonym changing operate inconsistently, the probability $p_{i'|i}$ becomes 1 to link PID_i with PID_i' via s_i. Therefore the tracking uncertainty of the targeted vehicle

V_i in the mix-zone becomes 0 and the adversary can track the vehicle based on the BSMs continuously.

The degree of location privacy not only depends on the location privacy achieved in mix-zones, but also on the maximum tracking time. Generally, the time-to-confusion is limited by the pseudonym age, if a targeted vehicle always changes the pseudonyms together with at least one neighbor in the mix-zone to confuse the adversary. However, when the MAC context linking attack is carried out, the time-to-confusion is much increased. Consider V_i reserves a time slot at the time t_0 and utilizes $PID_{i,1}$ to transmit. It changes the pseudonym at time t_1 and keeps the occupancy of the slot until t_2. Therefore, $\varepsilon_p = t_1 - t_0$. However, when t_2 is not in the same tracking step as t_1 of the adversary, the time-to-confusion is increased to $t_c = \min\{T_{trip}, t_{U>0} - t_0\}$, where T_{trip} is the time to finish the trip and $t_{U>0}$ is the time that the adversary becomes confused.

In other words, the location privacy of vehicles is achieved only if the following conditions are satisfied. 1) Each vehicle changes the pseudonym with enough cooperative vehicles in the mix-zones. 2) The MAC layer transmission pattern of each vehicle should be changed with the pseudonym simultaneously to resist the MAC layer context linking attack.

3.3 MAC Layer Attack Resistant Pseudonym (MARP) Scheme

MARP is a cross-layer scheme which coordinates vehicles to adaptively change pseudonyms and access the wireless channels in a distributed manner. It involves the time slots reservation to broadcast BSMs, mix-zone construction, pseudonym changing with new slot reservation. We assume that the system initialization, pseudonym provisioning, key generation, message signing, misbehavior report and revocation are operated according to the

standardized approach [74]. We select a representative vehicle, denoted by V_i, to describe the operation flow of each vehicle in the MARP scheme. Fig. 3.4 illustrates the operation flow of each vehicle in the MARP scheme. To prevent the context-linking attack, when the vehicles change their pseudonyms, they need to change the identifiers of each layer simultaneously, including the MAC address and IP address. In the MARP scheme, the Cryptographically Generated Address (CGA) protocol [100] is utilized to generate the new IP address based on its pseudonym of each vehicle, and the new MAC address from the new IP address accordingly.

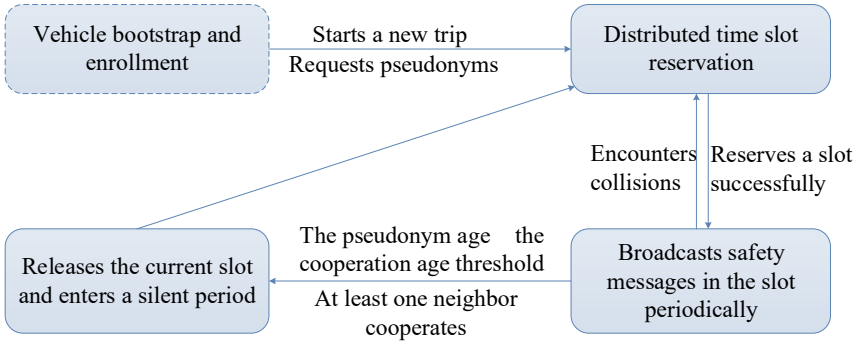


Figure 3.4: The operation diagram of every vehicle in MARP

3.3.1 Distributed Time Slotted Access

The proposed scheme is based on a slotted frame structure, as illustrated in Fig. 3.5. We design that every vehicle can occupy a dedicated slot in the distributed transmission (DT) period to broadcast its identity (including the pseudonym and certificate), safety information and the signature, and a list of its perceived status of all slots, which is referred to as the frame information (FI). Specifically, the kth time slot in the DT is denoted as ts_k and its status information is denoted as s_k, which consists of two bits: "00" indicates the slot is perceived to be unoccupied; "01" indicates that the slot is occupied by a vehicle within the one-hop transmission coverage of the sensing vehicle (including the sensing

vehicle itself); "10" indicates that the corresponding slot is occupied by a vehicle outside the one-hop transmission coverage of the designated vehicle but within that of a one-hop neighbor of the sensing vehicle, i.e., a hidden node to the sensing vehicle; and "11" indicates a transmission collision, i.e., multiple concurrent transmissions are received at the vehicle. Moreover, a contiguous number of slots towards the end of a frame can be released from being occupied by specific vehicles, and can be combined for contention-based access, e.g., by exploiting DIFS.

In addition to the periodical broadcast of safety messages, vehicles dynamically transmit the encrypted pseudonym requests to the roadside TAs based on their locations and the residual trip duration. Vehicles may also need to receive the CRL and report misbehavior by communicating with the roadside TAs. To separate the transmission of these occasional messages to the periodical safety messages, a contention based flexible schedule (FS) period has been defined for the vehicles to transmit the event-driven messages. By this way, the pseudonym request timing information does not harm the user privacy and the occasional transmission does not harm the safety messages transmission.

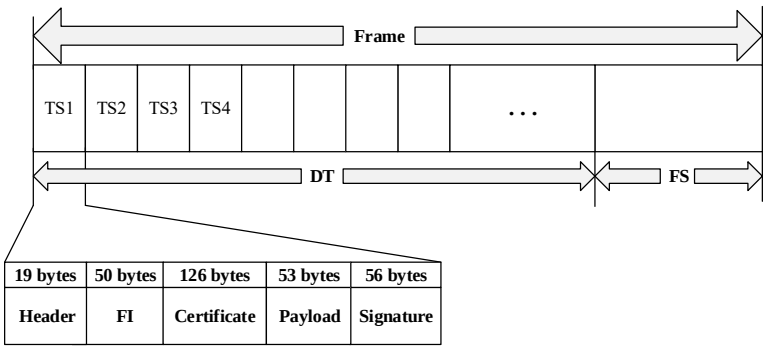


Figure 3.5: The frame structure in the MARP scheme

When V_i starts to acquire a new slot or shuffle the old slots, it initiates a random access procedure to reserve a dedicated slot that has not been occupied by all its two-hop range neighbors, and hence the hidden terminal issue is prevented. To achieve this, vehicle V_i

needs to carry out a procedure to update the FI by listening to the channel for a complete transmission period to receive the messages from its one-hop neighbors, extract their FI, and identify unoccupied slots within the transmission range of its one-hop neighbors and their one-hop neighbors (which can contain the hidden nodes to V_i). After updated the unoccupied slot set, vehicle V_i randomly selects an available time slot indexed ts_i and transmits its own messages as illustrated in the bottom of Fig. 4. The format of the message generated by the tagged vehicle V_i is $msg = (Header||FI||Cert_{PID_{ij}}||M||Sign_{k_{ij}}(M))$, where M is the safety payload, $Cert_{ij}$ is the pseudonymous certificate of vehicle V_i corresponding to the jth pseudonym and the signature $Sign_{k_{ij}}(M)$ is the signature of message M generated using the private key k_{ij} based on the Elliptic Curve Digital Signature Algorithm (ECDSA).

The vehicle V_i can be notified if the slot is successfully reserved by listening to the channel for the next $N_{ts} - 1$ transmission slots. Specifically, if all the one-hop neighbours of V_i have updated that status of ts_i is "01", the reservation of the slot dedicated to V_i is successful. There is a possibility that there is at least another vehicle, for example, V_j $(j \neq i)$, attempts to access the same slot as V_i. As a consequence, a transmission collision occurs. The vehicles V_i and V_j can be either within the one-hop transmission range of each other, or hidden nodes towards each other. In the case that they are within the one-hop transmission range of each other, the collision can be clearly indicated in the slot occupancy information as "11"(*collide*) from the following transmissions of other vehicles. In the case that V_i and V_j are hidden nodes to each other, some of their one-hop neighbors may update the slot occupancy to indicate that the slot is successfully reserved (by either of these two vehicles), while others indicate a collision in the slot. In both of the cases, the "collide" indicator of the slot can be perceived. The reservation of the slot is unsuccessful for both vehicles. The vehicles restart their reservation process again, and repeat until they successfully reserve the slots. At the end of each time slot, each

vehicle should update its neighbor set and the status of the time slots, and broadcast the updated FI in its own messages. The pseudo-code of the time slot reservation is described by Algorithm 1 in Appendix A.1.

It is also possible that the more than one vehicles, that used the same slot, previously beyond the two-hop coverage of each other, move towards each other and become hidden nodes, and cause transmission collisions (referred to as "merging collision") at the slot. By monitoring the lists of slot occupancy from surrounding vehicles, the vehicles of interest could be aware of the collisions, release the occupancy of the slot, and start the reservation process again, as described above. Each vehicle utilizes the slot in each frame to broadcast the safety messages periodically and release the time slot until it changes pseudonym or detects a merging collision. The pseudo-code for the time slot shuffle procedure is illustrated by Algorithm 2 in Appendix A.1.

As discussed, the distributed slotted access is able to eliminate the hidden issue, and embrace the mobility of vehicles. Moreover, it is able to guarantee collision-free delivery of critical safety information in a timely fashion, which is of practical value to many mission critical applications and scenarios, such as columns of military vehicles in war zones.

3.3.2 Mix-zone Construction

Both the reserved slots and pseudonyms of the vehicles need to be changed regularly. This is because they can be correlated and associated with particular vehicles, and the use of them for an extensive period of time would provide adversaries opportunities to identify each individual vehicle and track its trajectory. It is particularly important to change the slots and pseudonyms of multiple vehicles together to confuse an adversary. For this reason, we assume that every pseudonym preloaded at a vehicle is assigned with a predefined lifetime for the effective use of the pseudonym.

The MARP scheme employs an fully distributed strategy for that each vehicle makes

the pseudonym changing decision independently. Each vehicle listens to the BSMs broadcasted by its neighbors and determines whether to construct the mix-zone based on its own pseudonym age and the cooperation condition of its neighbors. Before each transmission, each vehicle checks whether its pseudonym age has reached the cooperation age threshold, i.e. $\varepsilon_p \geq \tau$, and the number of the cooperative vehicles. The vehicles can set different cooperation age thresholds, so as to achieve their preferred location privacy levels. If it is, V_i enters a semi-silent period by transmitting the BSMs with null location information and sets the cooperation indicator $ChangeReq$ continuously until it meets at least k cooperative vehicle. The vehicles can set different k requirement, to achieve their preferred anonymity set size in the mix-zone. By setting the cooperation indicator in the messages, a consensus is implicitly made between the cooperative vehicles to construct the mix-zone. After the mix-zone is constructed, the cooperative vehicles randomly select a silent period from $[ts_{min}, ts_{max}]$ to keep silent. During the silent period, there would be more vehicles driving into the mix-zone to mix together. As a consequence, the anonymity set size and the available slot set size are both increased. After the silent period expires, the vehicle performs the slot reservation process using a new pseudonym. Algorithm 3 in Appendix A.1 describes the mix-zone construction procedure for each vehicle in the MAPR scheme.

3.3.3 Dynamical operation of the MARP scheme

At the end of the section, the adaption of the MARP scheme to extreme traffic scenarios and different safety, security and privacy preferences is discussed. Firstly, the RSUs can cooperate with the vehicles to adjust the duration of the transmission period and the schedule period, so as to adjust the frequency of the broadcast of the safety messages and the security related messages, i.e., new pseudonyms, CRLs and misbehavior reports.

Secondly, even without the RSUs, vehicles can detect the traffic density accurately by

updating the status of the transmission slots. When the traffic is congested, decentralized congestion control schemes are easily to apply, e.g., vehicles can decrease the transmission frequency by taking turns with their neighbors in the same time slots to broadcast. On the other side, when the vehicles detect the traffic density is too sparse, vehicles can choose a longer silent period when construct the mix-zone to gather the cooperative neighbors so as to protect the location privacy with moderate sacrifice on the traffic safety, or a lower anonymity set size in the mix-zone.

In other words, the propose scheme is adaptable to different traffic scenarios and support different levels of location privacy and transmission requirements. While in this paper, we consider the free-flow traffic scenarios and the cooperation age threshold of all vehicles is assumed to be the same.

3.4 Analytical Evaluation

3.4.1 Performance Metrics

In this section, we analytically evaluate the location privacy quality that the vehicles could obtain by the MAPR scheme. The metrics include the pseudonym age, average anonymity set size and time-to-confusion. In addition to the three privacy metrics, the packet overhead and the frame duration are also taken into consideration to evaluate the transmission efficiency of the vehicles in the MARP scheme.

- The *pseudonym age* ε_p is defined as the interval during which a pseudonym is used.
- The *average anonymity set size* when vehicles change pseudonyms indistinguishably in the mix-zone is derived, which depends on the probability that at least two vehicles cooperate to construct a mix-zone, denoted by $p(c)$.
- The *time-to-confusion* , denoted by t_c, is defined as the successive duration that an adversary can distinguish and trace a target vehicle. It depends on the pseudonym

age, the anonymity set size, the silent duration before changing the pseudonym and the safety message transmission rate of vehicles in the mix-zone.

3.4.2 The Age Fluid Model

A novel analytical model, named by the *age fluid* model, is proposed to describe the evolution of the pseudonym age of the vehicles in the network. By the Kurtz' theorem [101], we note that, as the number of vehicles becomes large, even though their individual pseudonym age is time-varying, the pseudonym age distribution in the whole network is asymptotically deterministic.

Define the pseudonym age set of the system at time t is $\vec{\varepsilon_p}(t) = (\varepsilon_{p_i}(t)_{i=1}^{N})$, whereas $\varepsilon_{p_i}(t)$ characterizes the most recent pseudonym age of vehicle v_i, N is the total number of vehicles within the same geographical area. The pseudonym age for a vehicle v_i is calculated as $\varepsilon_{p_i}(t) = t - t^l$, where t is the current time, and t^l is the time of the last pseudonym change. The dynamics of age is characterized by the *aging* and *jump* process. When at least k vehicles decide to construct the mix-zone together, they change the pseudonyms and the age jumps to 0. The rate of the jump process depends on the number of vehicles within the transmission range, and the cooperation probability of the vehicles. Otherwise, the pseudonym age of a vehicle experience the aging process, in which it increases in a rate of the BSM broadcasting. The pseudonym aging in the drift process depends on the transmission interval of the BSMs transmission. Therefore, the higher transmission frequency, the larger aging rate, to characterize the effect of the transmission frequency to the location privacy. The rate of the jump process depends on the number of vehicles, and the cooperation probability of the vehicles. If v_i decides to change its pseudonym at time t, it needs to wait for the cooperative neighbors. Denote the cooperative neighbor set of v_i as $\mathcal{C}_i$. The pseudonym changing cost of v_i is denoted by $\gamma(v_i)$, which is defined as the sum of the semi-silent period to hear from at least one

cooperative neighbor and the silent period before changing the pseudonym. The cost is expressed in the aging rate which models the gap between the jump process and a new round aging process of the pseudonyms of each vehicle. Thus, it can be used to evaluate the effect of pseudonym change on the BSM transmission. The new pseudonym starts its lifetime from $t+\gamma$.

$$\begin{cases} \varepsilon_{p_i}(t+\gamma) = 0, t^l = t, \ when \ |\mathcal{C}_i| > 0 \\ \varepsilon_{p_i}(t) = \varepsilon_{p_i}(t^-), \ when \ |\mathcal{C}_i| = 0 \end{cases} \tag{3.2}$$

As the pseudonym age distribution of each vehicle is tight and homomorphic. In other words, the elements in $\vec{\varepsilon_p}(t)$ are labelless. Thus, the pseudonym changing process of all vehicles in the network is a density-dependent Markov process. Thus, the pseudonym evolving process of all the nodes in the network can be characterized by the evolving process of the pseudonym age distribution in the network. Define the occupancy measure of vehicles in the network with the pseudonym age of z by $M^N(z,t) = \frac{1}{N}\sum_{i=1}^{N}\delta_{\varepsilon_i(z,t)}$, which also is the density of vehicles with the pseudonym age of z in the network, i.e. $f(z,t)$. Therefore, the cumulative distribution function (CDF) over $\bar{M}(z,t)$ is:

$$F(z,t) = \bar{M}^N(t)([0:z]) = \int_0^z \bar{M}(\varepsilon_p, t)d\varepsilon_p. \tag{3.3}$$

$F(z,t)$ denotes the proportion of N with pseudonym age less than or equal to z. Based on the mean field limits, when N is large enough and the mobility of the vehicles is independent from each other, the evolving-stationary model indicates that the collection of the occupancy measure of the pseudonym aging of vehicles $\bar{M}(z,t)$ converges in distribution of deterministic process $\{\bar{m}(t)|t \geq 0\}$. Therefore, the evolving flow of the pseudonyms can reach a stationary rate, i.e. $\frac{\partial F(z,t)}{\partial t} = 0, t \to \infty$. Then we work out the distribution of $\{\bar{m}(t)\}$ by characterizing the dynamical changes of the occupancy measure for the age of pseudonyms less than or equal to $z(z>0)$ in the network in a small interval ∂t.

A fraction of pseudonyms experience the aging process and their age grows older in the interval ∂t. However, only the fraction of vehicles' pseudonyms within $(z-\partial t, z]$ grow

older and become older than z, and need to be removed from $F(z,t)$. While the others do not need to be subtracted, as they are not in $F(z,t)$ before the interval ∂t. Therefore, the rate of change of $F(z,t)$ caused by the aging process is calculated as:

$$\lim_{\partial t \to 0} \frac{|F(z-\partial t,t) - F(z,t)|}{\partial t} = \frac{\partial F(z,t)}{\partial z} = \bar{M}^N(z,t) \tag{3.4}$$

Meanwhile, the rest vehicles' pseudonyms experience the jump process, in which the vehicles meet with each other and decide to change the pseudonyms in the mix-zone. If the vehicle's old pseudonym age at time t is less than z, then after the pseudonym changing within ∂t, the pseudonym age is reset to 0, which is still less than z. Therefore, it causes no change of the $F(z,t)$. If the vehicle's old pseudonym age at time t is equal to or large than z, and it decides to change the pseudonym age within ∂t, therefore $z \geq \tau$. Then after the pseudonym changing within ∂t, the pseudonym age is reset to 0, which causes an increase of $F(z,t)$. Consider the time cost by the semi-silent period to gather the cooperative neighbours is upper bounded by γ_s, the changing rate of $F(z,t)$ caused by the jump process is calculated as:

$$\int_z^{\infty} p(c(t)) f(\varepsilon_p, t) d\varepsilon_p = \int_{\tau}^{\tau+\gamma_s} p(c(t)) f(\varepsilon_p, t) d\varepsilon_p, \tag{3.5}$$

where $p(c(t))$ is the probability that at least one of the encountered vehicles chooses to cooperate. It can be calculated as the combination of the probability of having n vehicles in the surrounding area, and the probability that at least one of them cooperates:

$$p(c(t)) = 1 - \sum_{n \geq 0} p_n (1 - c(t))^n, \tag{3.6}$$

where p_n is the probability of having n neighbors within the one-hop set of a vehicle v_i, $c(t)$ is the cooperation probability of a neighbor. Assume the network is homogeneous that all the vehicles apply the same age threshold, therefore the cooperation probability at any time t is calculated as

$$c(t) = \int_0^{\infty} c(z) f(z,t) dz = \int_{\tau}^{\tau+\gamma_s} f(z,t) dz, \tag{3.7}$$

where

$$c(z) = \begin{cases} 0, z < \tau \\ 1, z \geq \tau. \end{cases} \tag{3.8}$$

Thus, for the evolving stationary system, unique solution of the differential equation $\frac{\partial F}{\partial t}$ here is

$$\begin{cases} -\frac{\partial F(z,t)}{\partial z} + \int_{\tau}^{\infty} p(c) f(\varepsilon_p, t) d\varepsilon_p = 0, \\ F(\infty, t) = 1, \forall t \geq 0, \end{cases} \tag{3.9}$$

Based on (3.7), (3.8) and $\frac{\partial F}{\partial z}(z,t) = f(z,t)$, (3.9)can be transformed to

$$\begin{cases} \frac{\partial f}{\partial z} + p(c) \cdot f(z) \cdot c(z) = 0 \\ \int_0^{\infty} f(z) dz = 1 \end{cases} \tag{3.10}$$

When $z < \tau$, (3.10) becomes $\frac{\partial f}{\partial z} = 0$. Thus, $f(z) = f(0)$. When $z \geq \tau$, (3.10) becomes $\frac{\partial f}{\partial z} + p(c) f(z) = 0$. Thus, $f(z) = f(0) e^{-p(c)(z-\tau)}$. The final solution of (3.10) is given by

$$f(z) = \begin{cases} \frac{1}{\tau + \frac{1}{p(c)}}, 0 \leq z < \tau \\ \frac{e^{-p(c)(z-\tau)}}{\tau + \frac{1}{p(c)}}, z \geq \tau \end{cases} \tag{3.11}$$

Based on the cooperation function of $c(z)$ and $f(z,t)$, $c(t)$ is calculated as: $c(t) = \frac{1}{1+\tau p(c)}$. Thus, when the cooperation probability $p(c)$ is calculated, the final result of $f(z,t)$, and the average anonymity set size of a mix-zone can be achieved.

3.4.3 Derivation of the Cooperation Probability

To calculate $p(c)$, the probability p_n needs to be determined. Consider the traffic is under a balanced steady flow condition that the average arrival rate is λ that follows the Poisson distribution. The transmission range R of each vehicle is assumed to be much larger than the width of the roads. According to [102], the speed per vehicle is identically

and independently distributed yielding a truncated normal distribution. The probability density function (pdf) is given by

$$g(s) = \frac{\xi}{\sigma\sqrt{2\pi}} e^{-(\frac{s-\bar{s}}{\sigma\sqrt{2}})}, \tag{3.12}$$

where $\bar{s}$, s_{min}, s_{max} and σ are the average, minimum, maximum and the standard deviation of vehicle speeds respectively. Based on the well established traffic flow theory that, over a specific road segment, the average traffic density is given by

$$\rho = \lambda \bar{s}^{-1}. \tag{3.13}$$

Given vehicles distributes according to the spatial Poisson process as: $Pr(N = n) = \frac{(2R\rho)^n}{n!} e^{-2R\rho}$, with the expected number $\lambda_{2R} = 2R\rho$. The probability of v_i having n neighbors is equal to the probability that the total number of vehicles within the specific area is $N = n + 1$. Thus the probability mass function (pmf) of $|N(v_i)|$ is given by

$$Pr(|C_{v_i}| = n) = Pr(N = n + 1) = \frac{\lambda_{2R}^{(n)} e^{-\lambda_{2R}}}{(n)!}, n = 1, 2, ... \tag{3.14}$$

Subsequently, (3.6) is calculate as

$$\begin{aligned} p(c) &= 1 - \sum_{n=0}^{\infty} \frac{\lambda_{2R}^n}{(n)!} e^{-\lambda_{2R}} (1 - c(t))^{n-1} \\ &= 1 - \frac{1}{1 - c(t)} \left\{ \textstyle\sum_{n=1}^{\infty} \frac{\lambda_{2R}^n}{n!} e^{-\lambda_{2R}} (1 - c(t))^n \right\} \\ &= 1 - \frac{1}{1 - c(t)} \left\{ e^{-\lambda_{2R} \cdot c(t)} - e^{-\lambda_{2R}} \right\} \end{aligned} \tag{3.15}$$

By replacing $c(t) = \frac{1}{1+\tau p(c)}$ by $p(c)$ in equation (3.15), we obtain

$$p(c) = 1 - \frac{1 + \tau p(c)}{\tau p(c)} \{ e^{-\lambda_{2R} \frac{1}{1+\tau p(c)}} - e^{-\lambda_{2R}} \}. \tag{3.16}$$

Since the value of $p(c)$ must in the range of $[0, 1]$, the final value can be solved by the iterative method for (3.16). Assume each vehicle maintains a constant speed and moves with negligible interaction with its neighbors within an observation period. Thus it is

practical to assume that the one-hop neighbor set of v_i stays constant when constructing a mix-zone, and can be calculated based on vehicle density distribution in the network. If the number of the one-hop neighbors is known, correspondingly, (3.16) is approximately calculated as $p(c) = 1 - (1 - \frac{1}{1+\tau p(c)})^n, n \geq 1$.

3.4.4 Time-to-confusion

Based on the derivation of the distribution of the pseudonym age, we can calculate the average pseudonym age based on the assumption that the system is stationary and uniform. The upper bound of the pseudonym age can be calculated out from equation (3.11) and (3.16) as

$$\bar{\varepsilon}_p = min\{z|f(z) = 0, z \geq \tau\}. \tag{3.17}$$

According to the evaluation, the average anonymity set size within a specific area depends on the vehicle cooperation rate, the traffic density on the roads, the pseudonym age threshold and the silent period ts. Give the average vehicle arrival rate λ, the average number within one hop range is given by $\frac{\lambda_{2R}}{1-e^{-\lambda_{2R}}}$. Since the cooperation decision of each vehicle is independent to each other, the average anonymity set size is calculated as: $E[C(v_i)] = \frac{\lambda_{2R}}{1-e^{-\lambda_{2R}}} \cdot \min(1, \frac{ts_{min}+ts_{max}}{2\cdot(1+\tau p(c))})$.

As the silent period duration is uniformly distributed, the expected silent period of each vehicle is $\frac{ts_{min}+ts_{max}}{2}$. The cost for pseudonym changing is approximately calculated as $\gamma \cong \bar{\varepsilon}_p - \tau + \frac{ts_{min}+t_{max}}{2}$. Based on the analysis in subsection 3.2.3, the traceability of a targeted vehicle v_i becomes zero when the target vehicle changes the pseudonym and releases the transmission slots simultaneously with at least one cooperative neighbor. The time-to-confusion can be bounded by the pseudonym age, i.e., $t_c = min\{\bar{\varepsilon}_p, t_{trip}\}$. Therefore, by adjusting the value of τ, different privacy levels can be achieved in the network.

Otherwise, if the time-to-confusion of an adversary is not bounded by the pseudonym age in a pseudonym change scheme, the pseudonym change scheme is taken to fail to guarantee the location privacy. As the influential factors include the cooperation age threshold, traffic density and the silent period, more effective pseudonym schemes can be designed based on the framework of the MARP scheme and the analytical model by designing the cooperation age threshold and silent period for vehicles to achieve the preferred bounded time-to-confusion.

3.4.5 Packet Overhead Evaluation

As described in subsection 3.3.1, the format of each BSM consists of the header, payload, certificate with the corresponding pseudonym and the signature. The format of the safety message is shown in Fig.3.5. The size of each BSM is calculated as

$$L_m = B_{N_{ts}} + B_{header} + B_{payload} + B_{sign} + B_{cert}, \tag{3.18}$$

where $B_{N_{ts}}$ is the number of bytes to denote the status of each slot respectively in the FI, B_{header} is the number of bytes of the header specified for the WAVE safety message, $B_{payload}$ is the number of bytes of the safety data, B_{sign} is the number of bytes required for the signature, B_{cert} is the number of bytes required for the certificate. Assume $N_{ts} = 160$, corresponding to the most congested scenario in which the number of vehicles in one hop set is 80. It requires 50 bytes to represent the slot status. $B_{sign} = 56$, $B_{header} = 19$, $B_{payload} = 53$, $B_{cert} = 126$. Therefore, the total length a BSM is 304 bytes. The setting of the duration of each time slot depends on the transmitting delay of the safety message. Considering a mandatory supported transmission rate 12 Mbps by IEEE 802.11p, the transmission delay takes less than 0.22 ms. Assume the duration of each time slot is 0.25 ms, which is rational to support the transmission of the safety message. The transmission period is thereby being 40 ms and the flexible schedule period is 6 ms. The synchronized

frame duration is compliant with the default multi-channel operation defined by the IEEE 1609.4 [15].

3.5 Simulation

In this section, the accuracy of the analytical model is firstly verified, and then extensive simulations using a city scenario are carried out to validate the performance of the MARP scheme with comparison to other pseudonym schemes.

3.5.1 Analytical Analysis and Model Validation

In this subsection, we evaluate the analytical results presented in the previous subsection. In order to verify the accuracy of the analytical model, we presents the simulation results compared with the numerical results.

We consider a 2-km highway scenario with two opposite directions as described in the analytical model. The average traffic density is assumed to be from 0.0125 to 0.1 with the increment step as 0.0125 vehicles per meter per lane. Each vehicles moves with a constant speed drawn from the normal distribution and the number of vehicles on the system is kept constant during the simulation. To analyze the performance of the stable system, the vehicles start to move when they reserve the transmit slots successfully. When a vehicle reaches one end of the road, it re-enters the same lane from the entrance point, and starts to increase the age of its current pseudonym when transmitting the new safety message. Assume the physical channel is ideal so that each vehicle can communicate with all the vehicles within its transmission range with no obstacles. The transmission range of each vehicle is assumed to be 200m, therefore the expected number of vehicles within the one-hop range is from 10 to 80. The other simulation parameters as summarized in Table 3.2.

Table 3.2: Simulation Parameters for the MARP Scheme

Parameter	Values
Mean vehicle speed	40km/h
Vehicle speed standard deviation	10km/h
The width of the street	5m
τ: cooperation pseudonym age threshold	[1, . . . , 5] minutes
Safety message broadcast frequency	10Hz
ts_{min}: the minimum silent period	2 seconds
ts_{max}: the maximum silent period	8 seconds
t_{trip}: driving duration	20 minutes
The number of lanes on the road	2

First, the probability $p(c)$ of at least one vehicle cooperating in one frame interval is observed in Fig. 3.6. As we can see that the cooperation probability is rather low in a synchronized broadcast interval. The reason is that the pseudonym age of vehicles is approximately uniformly distributed, so that if one vehicle decides to change the pseudonym in one frame, it needs to wait and spends a silent period to gather more neighbors with their pseudonym age evolved for the cooperation. We evaluate the influence of pseudonym age threshold and the traffic density. As the number of neighbours in the targeted area (i.e. one hop transmission range) increases, the cooperation probability $p(c)$ increases logarithmical. When the pseudonym age threshold increases, we observe that $p(c)$ decreases for any value of τ. The reason is that for larger value of τ, a larger fraction of vehicles would have an age of pseudonym below τ at any constant. For this reason, for a high age threshold, fewer vehicles cooperate at a specific instant, and consequently $p(c)$ decreases.

Fig. 3.7 compares the average pseudonym age (Fig. 3.7a) and the average anonymity set size (Fig. 3.7b) calculated in the analytical model (denoted as Ana) to the results obtained by simulations (denoted as Sim) with five different cooperation age thresholds. The value of the average pseudonym age obtained from the analytical model presents a pretty accurate match with the simulation results. Although the cooperation probability

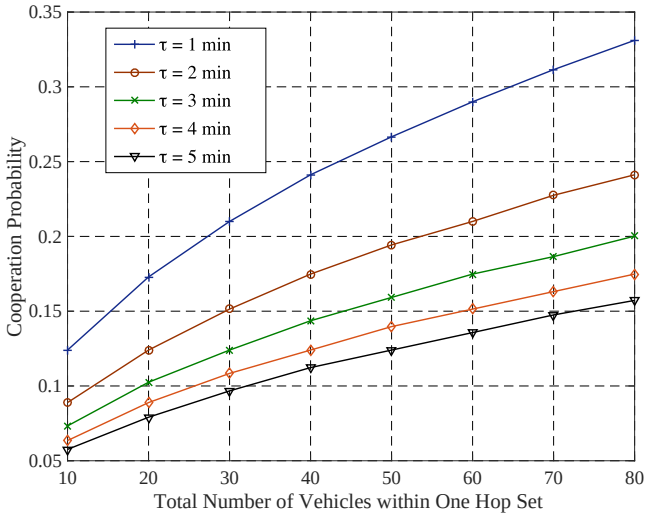
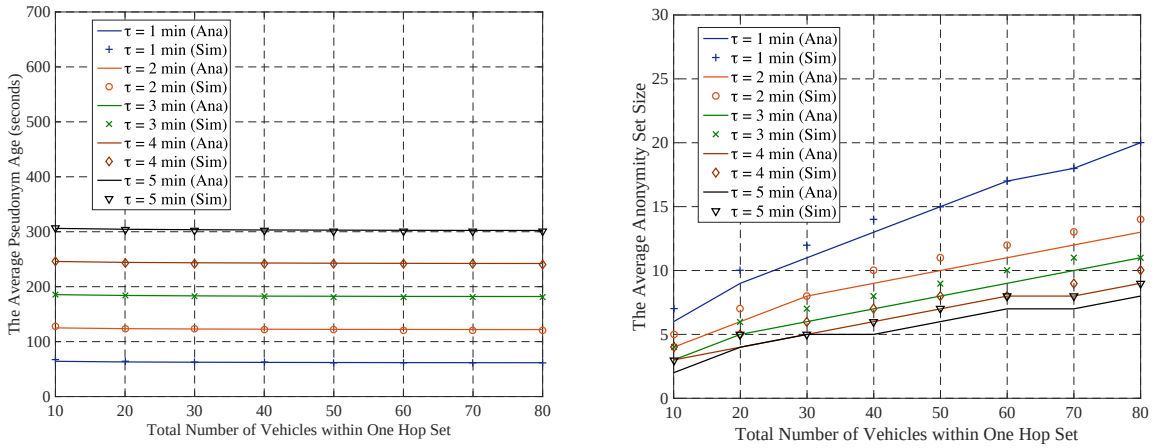


Figure 3.6: Influence of vehicle number and pseudonym age threshold: Probability of at least one neighbor cooperates

is rather low in a frame interval, the silent period in the mix-zone facilitate the increasing of the anonymity set size. As τ increases, we observe that the average pseudonym age increases accordingly. The pseudonym age is mainly dominated by the cooperation pseudonym threshold, as its value is close to the threshold with no larger than 10 seconds. As the number of vehicles in the targeted area increases, the average pseudonym age decreases slightly as more vehicles are cooperative, making the jump process in the system occur more frequently. Fig. 3.7b shows the anonymity set size in the mix-zones obtained in MARP. The effectiveness of MARP is validated even under sparse traffic density scenarios. The anonymity set size decreases with the increase of τ. The results are rational and obvious, as with a large τ, $p(c)$ decreases and fewer vehicles choose to cooperate. In addition, a dense traffic has a positive impact on the anonymity set size. It is also the reason that some researchers propose to change pseudonyms when vehicles detecting traffic congestion [68] or driving into large intersections [58]. The simulation results of the anonymity set size are a little higher than the analytical results. The reason is that during the simulation, the number of vehicles is limited so that their pseudonym

age is not independently distributed and becomes similar after several rounds of mix-zone construction.



(a) Average pseudonym age

(b) Average anonymity set size

Figure 3.7: Influence of vehicle number and pseudonym lifetime threshold.

3.5.2 City Scenario and Simulation Schemes

In addition to the verification of the MARP scheme based on the same assumption of the analytical model, we further study the performance of MARP using a city grid layout scenario. Specially, the scenario is with the size of $10km^2$, and consists of 16 square city blocks and 9 large intersections. The initial locations and driving routes are randomly generated for the vehicles in the simulation. All the vehicles drive continuously until the simulation time ends, and never drive out of the simulation area. Each street has two lanes with opposite driving directions. Vehicles with larger speed can go ahead of the vehicles with lower speeds. Assume when a vehicle reaches a junction area, it chooses on of all possible directions randomly. The traffic light is deployed at each intersection, with the duration of 1 minute.

At the beginning of the simulation, vehicles remain stationary and randomly select

a time instant within a interval as long as the cooperation age threshold to acquire a time slot by using the MARP scheme. Once the vehicle reserves successfully, it starts to transmit the safety messages, and its pseudonym starts to count. When all the nodes reserves successfully, the vehicles begins to move, and the simulation timer begins to count [46]. The adversary in the simulation would observe all the transmissions of the vehicles, and try to link their pseudonyms. When the information of the nodes of all layers becomes indistinguishable, the adversary loses the linking ability. On the other hand, if the adversary could link the vehicles by launching the context linking attack, it could track the vehicles continuously. In the simulation, we measure the performance of the MARP in terms of the average pseudonym age, anonymity set size and the time-to-confusion of the nodes, under different traffic densities and two different pseudonym cooperation thresholds, $\tau = 2$ and $\tau = 5$ minutes,as the recommended pseudonym age of the SAE J2735 [62] and European standard [61] respectively.

Considering the cross-layer performance, the packet delivery ratio (PDR) is used to verify the transmission reliability for MARP. The PDR of a vehicle is calculated as the total number of BSM messages that have been successfully transmitted within the lifetime to the total number of the BSMs generated by the vehicle. During a vehicle releases the old transmission slot and keeps silent before the new transmission slot reservation, the messages are assumed to be undelivered. When a vehicle encounters merging collisions or does not reserve a time slot, the safety messages are also assumed to be undelivered. The transmission delay is an important metric in the measurement of transmission efficiency. Using the TDMA based MAC layer operation, safety messages are always broadcasted in the specific transmission slots, thus the transmission delay in the MAC layer is bounded within 100ms.

The performance of the MARP scheme is compared with other three different cross-layer schemes.

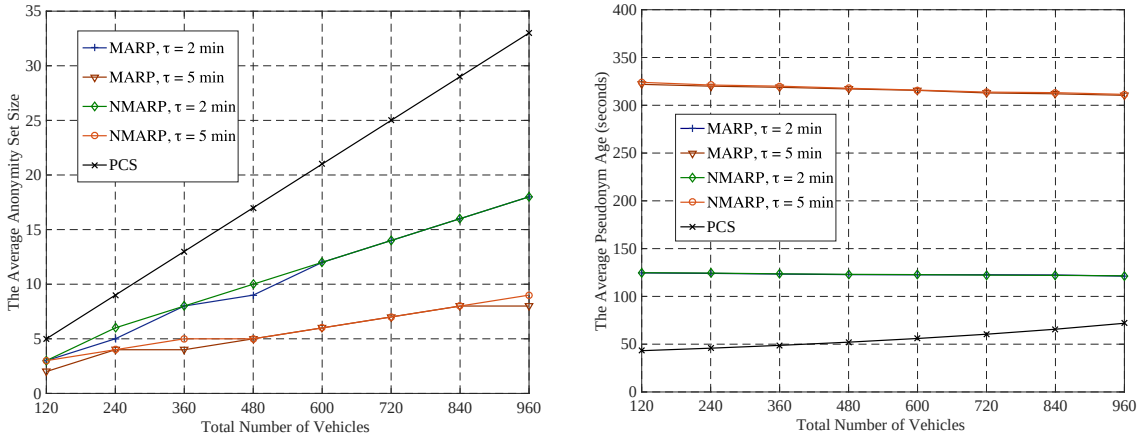
- NMARP: vehicles apply the same pseudonym changing and channel accessing strategies as defined in MARP. However, the transmission slots are not shuffled when changing pseudonyms, until collision happen or the trips are finished.
- PCS [58]: in the effective pseudonym changing at social spots (PCS) strategy, vehicles stop transmitting the safety messages when they drive across the intersection, and begin to construct the mix-zones. After vehicles leave the mix-zone, they start to utilize the new pseudonyms to transmit in a distributed manner. In the MAC layer, the effective TDMA based VeMAC [46] protocol is applied for the MAC layer operation while the pseudonym changing is periodic for each vehicle.

3.5.3 Simulation Results

The average anonymity set size in the mix-zones achieved in the three schemes is presented in Fig. 3.8a. It shows that the PCS strategy achieves the largest anonymity set size. In PCS, all the vehicles gathering in the intersections change the pseudonyms together without the consideration of the pseudonym age. The total number of vehicles at road intersections is much larger than the number of vehicles with the pseudonym lifetime expired within the one hop range in the MARP and NMARP schemes. All the three schemes can achieve a guaranteed anonymity set size level in the mix-zones. As the pseudonym changing strategies of MARP and NMARP are both based on the cooperation age threshold, they show almost the same performance in terms of the anonymity set size. In the figure, it is demonstrated that the anonymity set size decreases in MARP and NMARP as the increase of the cooperation pseudonym age threshold, which also correlates with the analytical results.

In Fig. 3.8b, the pseudonym age in the three schemes are investigated. As the figure shows, the pseudonym age in MARP and NMARP is mainly determined by the cooperation age threshold. In contrast, the pseudonym age in PCS is influenced by the frequency

of encountering intersections. In the simulated scenario, the average traveling duration between two intersections is about 1 minute. Therefore, vehicles consumes more pseudonyms in PCS in the simulated scenario. While in the real traffic environments, the utilization of pseudonym relies on the traffic typologies and traveling routes. If vehicles encounter intersection too frequently, it causes a waste of pseudonyms. On the other hand, if the intersection distribution is too sparse, it causes a long tracking duration of vehicles by adversaries and even compromises the location privacy.



(a) Average anonymity set size

(b) Average pseudonym age

Figure 3.8: The average anonymity set size and pseudonym age of vehicles in different schemes.

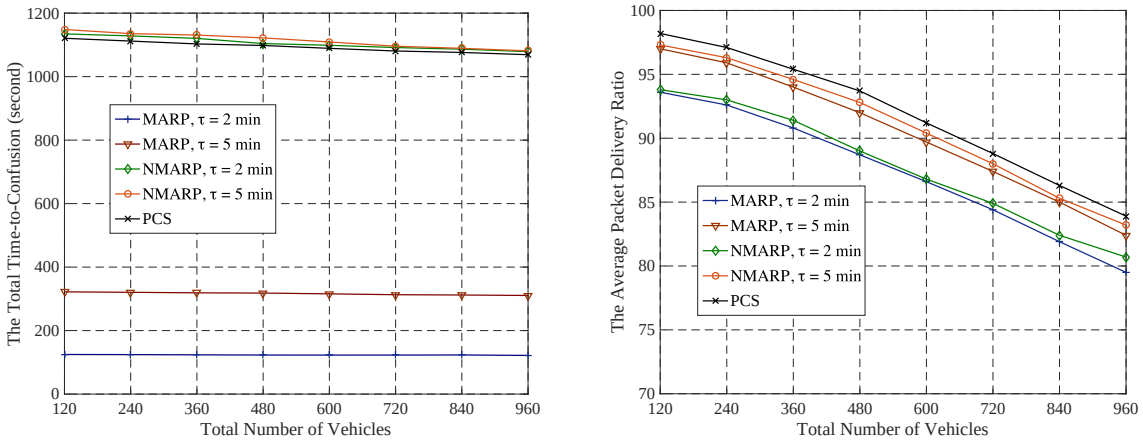
As Fig. 3.9a illustrates, the maximal time-to-confusion is unbounded by the pseudonym age in NMARP and PCS regardless of the pseudonym changing. In NMARP, vehicles change pseudonyms without shuffling the transmission slots. Vehicles usually acquire separately slots within two-hop range, the adversary can distinguish each vehicle based on its transmission slot index. Thus, the time-to-confusion is much larger in NMARP and PCS than in the MARP scheme. The time-to-confusion in PCS shows a slight decrease when compared with NMARP. The reason is that in PCS, vehicles change pseudonyms

at intersections, where the merging collisions happen more frequently. Thus, there is a higher probability that the pseudonym changing and slot releasing happen simultaneously among several vehicles in PCS, causing the adversaries lose the linking of the pseudonyms. However, the adversaries are still able to track a targeted vehicle continuously at most cases in PCS. In the MARP scheme, vehicles change the pseudonyms with the transmission slots during the silent period, which cuts the linking clue of the new pseudonym and old pseudonyms of vehicles in the network.

Fig. 3.9b compares the average PDR of the vehicles. Both the transmission collisions and pseudonym changing degrade the transmission performance. In VeMAC, transmission slots can be utilized until collisions happen or trip finished for the vehicles as the transmission slots are not changed with the pseudonyms. As the number of vehicles increases, the merging collisions dominate the packet loss in the VeMAC. In MARP and NMARP schemes, vehicles explore the silent-period when change pseudonyms, thus the PDR is slightly decreased compared to the VeMAC scheme. However, the influence of shuffling the transmission slot consistently with the pseudonym changing is negligible as the performance of MARP is close to NMARP. Based on the simulation results, we conclude that the MARP scheme provides superior performance to the compared schemes as it not only guarantee the location privacy preservation but also provide delay-bounded and reliable transmission for the vehicles in the VANETs.

3.6 Conclusion

In this chapter, we have proposed a new MAC layer context linking attack that could compromise the location privacy, and pointed out that the pseudonym changing schemes must be designed in a collaborative manner with the other layers' operations in VANETs. To deal with the MAC layer attack, a cross-layer pseudonym scheme, MARP, is presented.



(a) Average time-to-confusion

(b) Average PDR

Figure 3.9: The average anonymity set size and pseudonym age of vehicles in different schemes.

In particular, we have developed an analytical model to formally analyze the performance of the MARP scheme, in terms of the pseudonym age, cooperation probability and the anonymity set size. The analytical model concludes that, different privacy level could be achieved by adjusting the pseudonym change threshold of the vehicles. To prevent the continuous tracking from an adversary, the vehicles need to change pseudonyms as well as their transmission patterns/resources to make them indistinguishable. The analytical model is general to be applied to both CSMA and TDMA based networks and can be utilized to analyze other pseudonym schemes. The simulation results calibrate the analytical accuracy, and verify the effectiveness of the MARP scheme. To the best of our knowledge, most previous work considered the MAC layer protocols and pseudonym schemes separately. This paper sheds light on the cross layer protocol design in the VANETs, as the security and privacy preservation must be provided for all the layers. Our future work is to explore more specific privacy preservation techniques under the framework of MARP by considering various traffic scenarios and location privacy preferences of different users.

Chapter 4

A MAC Layer Aware Pseudonym (MAP) Scheme for the Software Defined Internet of Vehicles

4.1 Introduction

With the emergence of the software defined network (SDN) [103–105], fog/edge computing [106, 107], the development and framework of the vehicular network has envisioned from the VANET to the development of the Internet of Vehicles (IoV) [1, 2]. The IoV is expected to support V2V, V2I, V2P and V2C communications, for the exchange and distribution of various categories of messages. In the IoV, the resource utilization efficiency could be much improved and various applications are to be supported [1]. The numerous messages provide rich information about the vehicles' locations, social behaviors and enable the tracking of vehicles, it is of practical significance to protect the location privacy in IoV.

Recently, standardization bodies such as IEEE 1609.2 [13] and ETSI [61], have applied the public key infrastructure (PKI) based pseudonyms to represent the short-lived public

keys of each vehicle. To prevent the Sybil attack, the pseudonymous certificate is assigned by the CA. The pseudonyms are expected to change frequently and appropriately [55, 58, 66, 108], so that the messages originating from the same vehicle are unlinkable to prevent the continuous tracking from the adversary. Each vehicle needs to request the new pseudonyms from the CA continuously during its trip. Therefore, the pseudonym is also very important resource to be distributed efficiently in the network.

The applications and services in the IoV impose certain quality of service requirements, the other challenge could be the design of an efficient medium access control (MAC) protocol for the provision of the transmission requirements for the various applications. Prior studies [109, 110] have shown that the CSMA operation is only suitable for low-to-moderate traffic scenarios. As the safety messages are required to broadcast periodically by each vehicle, the time division multiple access (TDMA) based schemes have become one of the promising candidates for IoV. By letting each vehicle report their sense on the status of all the time slots [46], [98], [111], the TDMA approach can alleviate the hidden terminal problem. In these distributed TDMA MAC protocols, the slot reservation decision of each vehicle is made based on the its own and the neighbors' transmissions. This is typically the coupling of the control and data plane. Therefore, the transmission failure could degrade the slot utilization efficiency in the TDMA protocols.

In the IoV, a collection of RSUs is assumed to make distributed decisions to coordinate the vehicles. Due to the large scale of the vehicular network, and the limited coverage size of the RSUs, the coordinated decisions have to be made across several adjacent RSUs. The distributed coordination algorithms usually explore long latency to adjust the resource allocation at a large scale, and becomes more arduous due to the varying traffic densities, high mobility and various QoS requirements. With the development of SDN and cloud computing, ultra-dense deployment of the RSUs can form local clouds to cooperatively coordinate the channel resources and manage the pseudonyms for the vehicles [112]. In

ref. [113], we propose the TDMA based sdnMAC protocol in the software defined vehicular network to apply the local RSU cloud to assign the channel resources for the vehicles to transmit, so that the transmission reliability is significantly improved. In [112], an SDN based pseudonym allocation system is proposed for the vehicular network. However, as pointed in chapter 3, the adversaries could link the pseudonyms of a vehicle based on its transmission and resource utilization pattern. The more coordinated of the resource allocation, the more likely to facilitate the MAC linking attack by taking the transmission and resource allocation context as the "fingerprint" to link the vehicles' old and new pseudonyms. For example, in VeMAC, the disjoint sets of the transmission slots for vehicles moving in opposite directions reveal their driving directions. Meanwhile, the pseudonym changing introduces negotiation and cooperation overhead, which degrades the safety message transmissions [90]. To the best of our knowledge, it still lacks a protocol to resolve the operation inconsistency between the MAC layer and pseudonym schemes in the IoV.

In this paper, we present a hierarchical architecture named as the software defined Internet of Vehicles (SDIV) where the multi-channel transmission and pseudonym management are coordinated by the RSU clouds. Facilitated by the architecture, an efficient MAC layer Aware Pseudonym (MAP) scheme is proposed to deal with the new MAC layer semantic linking attack. Specifically, the main contributions of the paper are three-fold.

- We present a three-layer SDIV framework, which could be elastic to be deployed in various transmission scenarios to coordinate and manage the resources efficiently in the IoV.

- Facilitated by the SDIV framework, the MAP scheme is designed. In the MAP scheme, the local RSU cloud assigns the channel resources to the nodes based on the vehicles' transmission pattern, the traffic density and vehicle mobility, so as to alleviate the transmission collisions and ensure the QoS requirements of the vehicles.

The pseudonym changing decision is formulated by the RSU cloud according to a composite location privacy metric that includes the privacy level of each vehicle, transmission slot age and group confusion to deal with the MAC layer linking attack, and guarantee the location privacy of the vehicles.

- Security analysis and extensive simulations are conducted to show that the MAP scheme significantly enhances the location privacy, and improves the transmission efficiency and reliability in the SDIV.

The rest of the chapter is organized as follows. In section 4.2, we present the preliminaries of the chapter, and the hierarchical SDIV system architecture. Section 4.3 introduces the MAP scheme in the SDIV. Subsequently, the performance evaluation, and simulation results of the scheme are given in section 4.4 and 4.5, respectively. Finally, section 4.6 concludes the chapter.

4.2 System Model

4.2.1 System Assumption

We assume that the RSUs are connected by wired communication, and cover the network seamlessly. In other words, each vehicle could communicate with at least one RSU in the IoV. The DSRC and LTE resources both can be utilized for the on-board communication. For illustration, the distance between two adjacent RSU is denoted by $2R$, where R is also the transmission range of the RSUs. In addition, we assume that the communication link between the vehicles is symmetrical, which means that if the node v_i is within the communication range of node v_j, v_j must also be in the communication range of v_i. The physical layer wireless channel is assume to be ideal. If there is no collision during the transmission, all the nodes within the transmission range of the transmitting node could

receive the message. The threat model is assumed to be the same as that in section 3.2.2.

4.2.2 Software Defined Internet of Vehicles (SDIV) Architecture

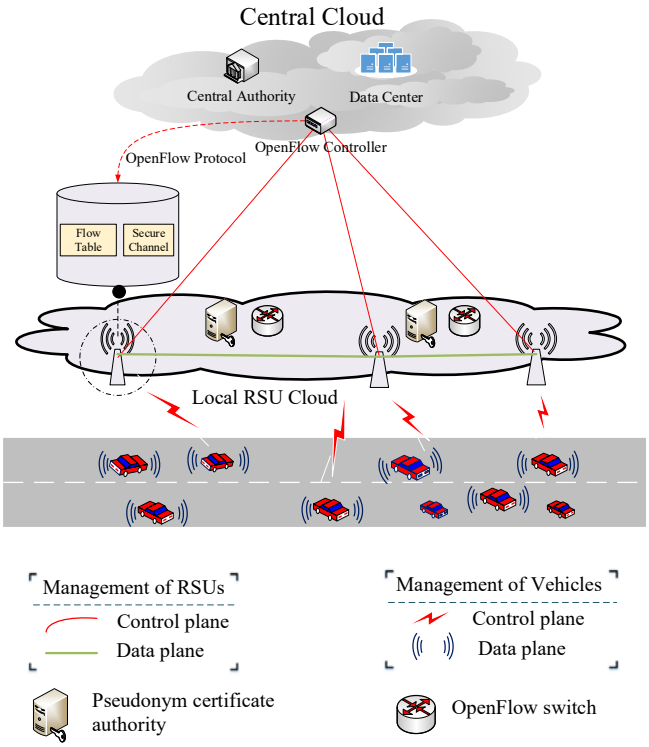


Figure 4.1: The hierarchical SDN based IoV architecture

SDN has been envisioned to control the network in a centralized and dynamical reconfigured approach with decoupling the control and data plane. The de facto protocol, OpenFlow is applied as the communication protocol between the control plane and the data plane in the proposed SDIV. The OpenFlow switch in the data plane is equipped with wireless interface to communicate with both the vehicles and the controller in the control plane. The SDN can be a significant advantage to improve the resource allocation and pseudonym management in SDIV.

Fig. 4.1 shows the hierarchical SDIV architecture. It consists of three layers: central cloud layer, RSU cloud layer and vehicle layer. The management of the SDIV is two-tiers: the RSUs connect together according to the flow tables to manage the vehicles, and they are managed by the controller in the central cloud layer. The de-facto protocol, OpenFlow is applied as the communication protocol between the control plane and the data plane in the proposed SDIV. The OpenFlow switch in the data plane is equipped with wireless interface to communicate with both the vehicles and the Controller in the control plane. The details about the three layers are described as follows.

- The vehicle layer is in the data plane of the SDIV. Each vehicle has at least one on-board unit (OBU) for the V2X communication, a global positioning system, and a temper-resistant device to store the cryptographic keys. The transmission range R is assumed to be the same for all the vehicles. For any vehicle, before entering the network, the bootstrap and enrollment are required for the OBUs to obtain the certificates of the CA, enrollment certificate and master key pairs. When starts a new trip or lack new pseudonyms, each vehicle needs to request more pseudonyms to preserve privacy. The requests are validated and processed by the nearby local cloud. Pseudonym changing decisions are made by the RSUs to notify the vehicles to change their pseudonyms together by constructing the mix-zones.

- The adjacent RSUs can communicate with each other via wired network, and connect with the local PCA to form the local RSU cloud. It takes the role as the OpenFlow Switch (OFS) to communicate with the vehicles via the wireless communication interface. The interconnection between RSU clouds and the communication between the RSUs and the central cloud layer is encrypted. The RSU cloud is controlled by the central cloud layer. By listening to the transmission of the surrounding vehicles, the RSU cloud can accurately perceive the utilization of the channel resources, the usage of the pseudonyms, and the vehicle densities. As the

control server of the vehicles, the RSU cloud is responsible to coordinate the time slot reservation procedure, and manage the pseudonym usage and changing for the vehicle. Thus, the data plane and the control plane can be decoupled.

- The central cloud layer is consisted of the CA, data center and the OpenFlow controller. The CA is responsible for managing the certificates, monitoring the behavior of all the entities including the vehicles, RSUs and PCAs, and generate the device certification revocation messages in the network, so as to ensure the network security and authentication. The data center collects and stores the information of all the local clouds, the information flows between the network entities. The OpenFlow controller is responsible for managing the RSU cloud in the control plane, and communicating with the RSU cloud via the encrypted channel in the data plane.

4.3 The MAC Layer Aware Pseudonym (MAP) Scheme in the SDIV

This section introduces the cross-layer MAP scheme, in which the RSU clouds coordinate vehicles to change pseudonyms and access the wireless channels in a contention free manner. The central cloud interacts with the RSU cloud to realize the allocation of the pseudonym resources, and the time slot resources across the geographical area. The interactive operation flow of a vehicle in the data plane and the controller in the control plane in the MAP scheme is illustrated in Fig. 4.2. We denote a representative vehicle by v_i and its local RSU by RSU_j to describe the MAP scheme. For clear exposition, the primary notations used in the chapter are described in Table 4.1.

Table 4.1: Notation and Description in Chapter 4

Notations	Description
v_i	A representative vehicle
$\mathbf{L}_i$	The location of v_i
$\mathbf{V}_i$	The driving speed of v_i
RSU_j	The local RSU cloud of v_i
RSU_m	The cross-region RSU cloud of v_i
R_i	The transmission range of v_i
N_{ts}	The number of slots per frame on DT
$\mathbf{N}_{ts}$	The set of slots on DT, $\mathbf{N}_{ts} = \{1, 2, \cdots, N_{ts}\}$
$\nu(j)$	The set of nodes within R to RSU_j
$\Phi(j)$	The set that includes RSU_j and its adjacent RSUs
$\psi(i, r)$	The set of nodes within the range of r to v_i,
	$\psi(i, r) = \{v_k \mid \|\mathbf{L}_k - \mathbf{L}_i\| \leq r\}$
R_p	The guard distance to enable pre-warning for
	the transmission collisions
U_i	The tracking uncertainty of v_i
$Cert_i$	The root certificate of v_i
PK_i, SK_i	The key pair of v_i
PID_{ik}	The kth pseudonym of v_i
(PID_{ik}, SK_{ik})	The pseudonymous key pair corresponding to PID_{ik}
$Cert_{PID_{ik}}$	The pseudonym certificate corresponding to PID_{ik}
$Sign_{ik}(M)$	The signature of message M signed by v_i
	using the private key SK_{ik}
GID_j	The group constructed by RSU_j for
	vehicles wanting to change pseudonyms
$SK_{GID_{ji}}$	The secret key of v_i in the group of GID_j
$Cert_{G_{ji}}$	The certificate for v_i in the group of GID_j
$E_{PK_{ik}}(M)$	The encrypted message M with v_i's public key PID_{ik}
$E_{SK_{ik}}(M)$	The encrypted message M with v_i's secret key SK_{ik}
$x \rightarrow y$	Node x sends a message to y
$\|\|$	The concatenation operation
$x \rightarrow y$	The node x transmits to y
$\gamma(u)$	The set of vehicles with privay level no larger than u
$N_\gamma(u)$	The number of vehicles with privacy level
	no larger than u
t_p^l	The time of the last pseudonym change by V_i
t_s^l	The time when reserves the current slot successfully

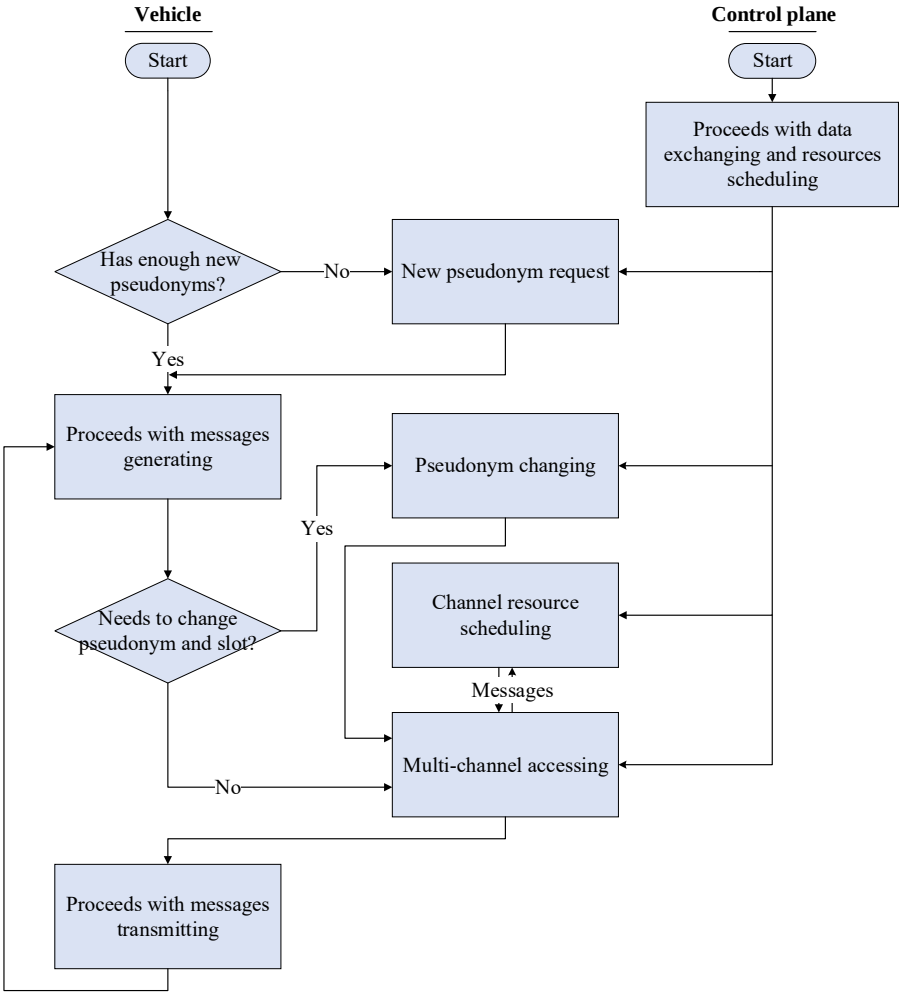


Figure 4.2: The interactive operation flow of each vehicle and the control plane in the MAP scheme

4.3.1 System Initialization and Basic Key Operation

In MAP scheme, efficient cryptographic schemes are generally compatible with the SDIV architecture. In this paper, we apply the Boneh-Boyen short group signature [114] for the key operation. Before v_i enters the network, bootstrap and enrollment are required to obtain PK_i, SK_i and $Cert_i$, and the public key and certificate of RSU_j, which are denoted by PK_j and $Cert_j$, respectively. The CA stores the information $\{RID_i, PK_i, SK_i, Cert_i\}$

of vehicle v_i in the database, and distributes to RSU_j for the tracking and revocation of v_i when dispute happens. After registration, v_i sends pseudonym request to RSU_j to receive a set of pseudonyms, and the corresponding private keys and the certificates $\{PID_{i,k}, SK_{PID_{i,k}}, Cert_{PID_{i,k}}\}_{k=1}^{w}$.

Every time v_i starts a new trip, it checks whether it still has new pseudonyms, if not, it requests a set of new pseudonyms from RSU_j. RSU_j verifies the request, sends the pseudonyms to v_i, and records them in the tracking list of v_i. In order to prevent the semantic linking attack, the pseudonym changing should be accompanied by the identifiers of the other layers and the transmission slots. In MAP, the CGA protocol is used to generate the new IP addresses from the new pseudonyms and to generate the MAC addresses from the generated IP addresses [100].

4.3.2 Channel Resource Scheduling

In the MAP scheme, time is divided into frames with duration of 100ms, which is compatible with the synchronous interval specified by the DSRC. The time synchronization is realized using the GPS signal of the OBU [46]. The CCH is divided into a number of time slots, which is denoted by N_{ts}. The local RSU is responsible to coordinate the time slots for the vehicles, while the time slots used for the RSUs are coordinated by the central cloud. Besides the safety and control messages, the pseudonym request message and the CRL message are exchanged in the SCHs.

The channel accessing coordination of RSU_j interacts with the channel coordination of the RSUs within $\Phi(j)$. The interactive coordination procedure considers both the transmission range and the vehicle density, to keep the number of vehicles within every two-hop set being equal to or less than N_{ts}.

Assume the transmission range of the RSUs is constant, which is equal to R. The RSUs transmit the traffic density to the central controller, so that the central controller

can control the transmission range of the vehicles under different RSUs with diverse traffic densities. When the central cloud decides the transmission range for $v_i \in \nu(j)$ under RSU_j, it needs to consider two aspects. Firstly, the vehicle needs to exchange message within at least one RSU, which gives $R_i \geq R$. Secondly, the vehicles within two hop set of each other needs to reserve different time slots, so as to avoid the hidden terminal problem. Assume the transmission range of vehicles is the same, thus R_i is given by

$$R_i = \max(R, \arg\min_{R'} |\psi(i, 2R')| = N_{ts}). \tag{4.1}$$

Eq. 4.1 is efficient when the number of vehicles is uniformly distributed. However, when traffic density is not uniform in the network, the transmission range of vehicles covered by different RSUs would be various, thus, $R_{V_i} = R_{V_{i2}}, V_i, V_{i2} \in \nu(j)$, and $R_{V_i} \neq R_{V_{i2}}, V_i \in \nu(j), V_{i2} \in \nu(k), RSU_j \neq RSU_k$. Considering the above two aspects, the transmission range of v_i is given by

$$R_i = \max(R, \arg\min_{R'} \arg\max_{V_k \in \psi(i,R')} \{\psi(i, |\mathbf{P_k} - \mathbf{P_i}| + R_k) = N_{ts}\}), \tag{4.2}$$

where each v_k is within the transmission range of v_i, and they determine the two-hop range of v_i, which is the largest distance between v_i and v_k plus the transmission range of v_k, i.e., $\arg\max\{|\mathbf{P_k} - \mathbf{P_i}| + R_k\}$. In this range, the number of the nodes should be no larger than N_{ts}, so as to guarantee each node can reserve at least one time slot.

In addition to the periodical broadcast of safety messages, vehicles dynamically transmit the encrypted pseudonym request or data service request to the RSUs. The CRL is also broadcasted by each RSU to the surrounding vehicles. It is assumed that the RSUs would broadcast the No. of the SCH in the CCH and then switch to the SCH to provide data services and broadcast CRL. When the vehicle needs to request a new pseudonym, it listens to the service broadcast of the RSU in the CCH, and then switches to the specific SCH to send a pseudonym request to the RSU based on CSMA. The encrypted security

related information does not harm the user privacy, and their separation transmission to the safety messages would not degrade the safety message transmission reliability.

In the MAP scheme, the transmission time slot of each RSU in the CCH is controlled by the central cloud. Each RSU has the priority to reserve the anterior time slots. In order to avoid transmission collisions between the adjacent RSUs, the same time slot cannot be used by the RSUs within two hop range of each other. For the description convenience, consider a highway scenario, in which the RSUs are located on the same straight line. As shown in Fig. 4.3, where the RSUs $R_1, R_2, R_3, R_4, R_5, R_6$ use the $1, 2, 3, 3, 2, 1$ time slots, respectively. Thus, the transmissions between the RSUs would not conflict. The update of the slot status is performed by each RSU in each frame, and transmitted to the vehicles and the adjacent RSUs.

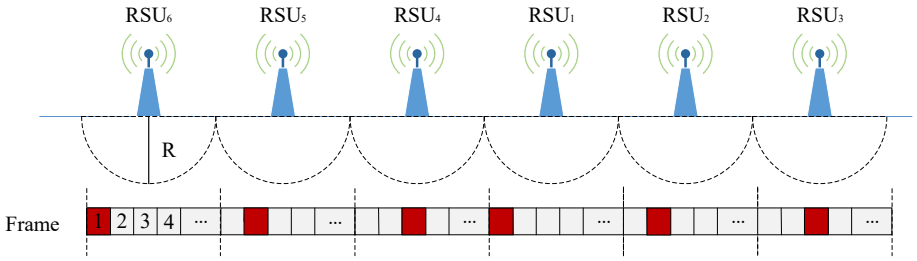


Figure 4.3: An example to illustrate the slot allocation to the RSUs in the MAP scheme

Each vehicle's transmission slot in the CCH is coordinated by the RSU. The management of the vehicles is independent and distributed in the network, thus, we select a representative vehicle, denoted by v_i, and its local RSU, denoted by RSU_j to describe their operation. For the description convenience, we assume that the transmission range of the vehicle is the same as the RSU's. Fig. 4.4 shows the packet format of the vehicle and RSU. The SI field in the header defines the updated slot status by the RSU. Each SI_i consists of two bits, defining the status of each slot $ts_i, \forall ts_i \in \mathbb{N}_{ts}$. The Service Info and Channel Info fields are used by the RSUs to notify the services provided by them

in the corresponding channels, and the header field is assumed to be consistent with the WAVE specification.

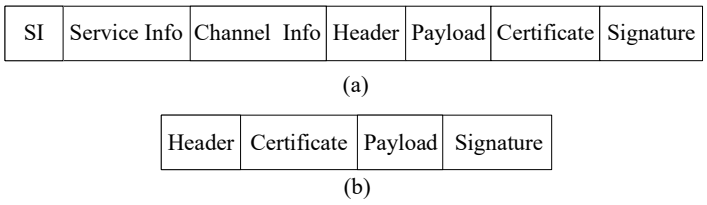


Figure 4.4: (a) The packet format transmitted by the RSU in the CCH (b)The packet format transmitted by the vehicle in the CCH

Specifically, to RSU_j's perspective, there are four cases of each slot' status, which are listed as follows.

- If a time slot, denoted by ts_i, is only occupied by one node in $\nu(j)$, then SI_i is set to "01".
- If ts_i is occupied by both v_k, where $v_k \in \nu(j)$, and v_m, where $v_m \in \psi(m, 2R)$, it indicates that a transmission collision has occurred, then SI_i is set to "11", to indicate that the slot has encountered a transmission conflict.
- If ts_i is occupied by both v_k, where $v_k \in \nu(j)$, and v_m, where $v_m \in \psi(m, 2R + R_p)$, v_k and v_m are moving toward each each, i.e., $\prec \mathbf{P}_k - \mathbf{P}_m, \mathbf{V}_k \succ < 0, \prec \mathbf{P}_m - \mathbf{P}_k, \mathbf{V}_m \succ < 0$, where R_p is the guard-distance to enable pre-warning of transmission collisions, and $\prec \mathbf{a}, \mathbf{b} \succ$ represents the inner product of $\mathbf{a}$ and $\mathbf{b}$. According to the analysis in [115], the merging collisions mainly happen between vehicles driving from opposite directions, or with high relative speeds. Thus, v_k and v_m have the largest opportunity to encounter each other, and a merging collision would happen. Then, SI_i is set to "10".

- In other cases, SI_i is set to "00", which indicates that the slot is unallocated, and it is available to the vehicles that need to reserve new time slots.

The Algorithm 4 in Appendix A.2 describes the channel resource coordination procedure of each RSU.

4.3.3 The Channel Accessing Procedure

When v_i starts a new trip, it needs to switch to the CCH to listen to the messages broadcasted by the RSU_j and the surrounding vehicles. It receives the SI from the RSU_j, determines the currently available time slots and its transmission power, and randomly selects a time slot with the status set as "00", denoted by ts_i, and then sends the safety message $msg = (Header||Cert_{ik}||M||Sig_k(M))$.

After v_i has attempted to reserve the time slot, it listens to the RSU_j for the updated SI, and uses the information to determine whether it has successfully reserved the time slot. If RSU_j has updated SI_i to "01", v_i determines that it has successfully reserved the time slot. If SI_i is still "00", v_i takes that RSU_j failed to receive its reservation message, and it needs to increase its transmission power appropriately to reserve the slot again. If SI_i has been updated to "11", which indicates that a transmission collision has occurred in ts_i, v_i then re-selects another available time slot to reserve, until it successfully reserves a time slot.

When v_i has become activate in the reserved time slot ts_i, it broadcasts the safety message, and receives the SI from the RSU_j in each frame. If v_i notifies that SI_i becomes "10", which indicates a pre-warning collision, and there still exists available time slots, v_i would starts to reserve a new time slot while still using the current for transmission. When v_i reserves the new time slot successfully, it would release the previous occupied time slot, to avoid the transmission collisions. If there is no idle time slot available, then v_i would keep using the current time slot to transmit. If v_i senses that SI_i becomes "11",

which means that a transmission collision has occurred, v_i would release the current time slot immediately, and begins to reserve a new time slot to transmit.

In addition to broadcast the safety messages in CCH, v_i receives the service advertisement information transmitted by the RSU_j, and it needs to switch to the specific SCH to send the pseudonym request, receive new pseudonyms and the CRL from the RSU, so as to realize both the safety message transmission reliability, and the location privacy preserving. The channel accessing procedure is described by Algorithm 5 in Appendix A.2.

4.3.4 The New Pseudonym Request Procedure

When the vehicle v_i lacks sufficient pseudonyms to use during the trip, it needs to request a set of new pseudonyms from the local RSU. The pseudonym request message is denoted by $req = RSU_j||E_{SK_{ik}}(PseudoReq||PK_i||Cert_i||Cert_{ik})$, where $PseudoReq$ contains the number of pseudonyms requested by the vehicle and the specific privacy requirements. When RSU_j receives the request from v_i, it verifies the message based on the data record of v_i, and determines whether to assign new pseudonyms to v_i. When v_i is driving cross the local region, assume that its local RSU becomes RSU_m, which does not record the identity information of v_i. RSU_m needs to request the data record of v_i, i.e., $\{Cert_i, PID_i, SK_{PID_i}, Cert_{PID_i}\}$ from RSU_j. If the local RSU decides to assign new pseudonyms to the vehicle, it transmits the new pseudonyms to v_i. The message is denoted by $rep = E_{PK_{PIDik}}(\{PID_{ik}, SK_{PID_{ik}}, Cert_{PID_{ik}}\}_{k=1}^{w})$. When v_i receives the rep message, it verifies the message, and transmits the acknowledgment, $ack = RSU_m||E_{PK_m}(Cert_{PID_{ik}}||timestamp)$ to the local RSU. The pseudonym request procedure is described by Algorithm 6 in Appendix A.2.

4.3.5 The Pseudonym Change Coordination Algorithm

The most applied metrics to evaluate the location privacy in the IoVs include the anonymity set size, the pseudonym age and driving distance [76], and the privacy entropy [82]. Based on the utilization duration of the pseudonym and the time slot, the anonymity set size, and the indistinguishability of vehicles in the mix-zone, the MAP scheme exploits a dynamical pseudonym changing strategy for the RSUs to coordinate the surrounding vehicles to change their pseudonyms. The related privacy metrics used in the MAP scheme are defined as follows.

- *Pseudonym age*: the pseudonym age at the time instant t, denoted by $A_p(t)$ is defined as the interval from the last pseudonym change time to the current time instant, i.e., $A_p(t) = t - t_p^l$. Similarly, the slot occupancy duration is given by $A_s(t) = t - t_s^l$,, where t_s^l is the time that reserved the current slot successfully. To balance the requirements between location privacy preservation and communication overhead, the utilization of a pseudonym should be longer than the changing threshold T_{th}, but limited to the expiry time T_{exp}. The slot can be used as the context information to facilitate linking the transmitting vehicle, therefore, it should be changed simultaneously when the vehicle changes pseudonym, so as to prevent the MAC layer semantic linking attack.

- *Anonymity set size*: the number of vehicles that change the pseudonyms simultaneously in a mix-zone, which is denoted by AS.

- *Group confusion*: we propose the metric for each vehicle v_i, denoted by G_i, to the geo-indistinguishability of V_i and the surrounding vehicles. The location of the vehicle will provide clues for the adversary to connect its pseudonym, and the location of the surrounding vehicle can be used to confuse the adversary. Based on the geo-indistinguishability concept [116], it is suggested that the geo-indistinguishability

of each vehicle when updating its pseudonym needs to satisfy $G_i \geq G_{th}$, which is define as, given $Varepsilon$, there should exist at least one neighbor vehicle v_k to satisfy:

$$|\mathbf{P}_k - \mathbf{P}_i| \leq -\frac{1}{\varepsilon}(W_{-1}(\frac{p_{th}-1}{e}+1)), \tag{4.3}$$

where $|\mathbf{P}_k - \mathbf{P}_i|$ is the Euclidean distance between v_k and v_i, $p_{th} \in [0,1]$ is set based on the tracking capability of the adversary. W_{-1} is the Lambert-W function, which is used to generate artificial noise [117]. An example is given for the value of d_{th}, $\varepsilon = 0.01, p_{th} = 0.05$, then $d_{th} = 122m$.

- Privacy entropy (privacy level): according to [82], the privacy level that the vehicle can achieve after changing its pseudonym in the mix-zone is given by

$$U_i(t_p^l) = -\sum_{j=1}^{AS} p_{j'|i} \log_2(p_{j'|i}), \tag{4.4}$$

where $\sum_{j=1}^{n} p_{j'|i} = 1$, $p_{j'|i}$ measures the probability that the new pseudonym $PID_j^{'}$ of vehicle v_j is linked with the old pseudonym PID_i of v_i. U_i is upper-bounded by $\log_2 AS$.

Besides, the vehicle's location privacy loss function $\beta_i(t, t_p^l)$ is defined as:

$$\beta_i(t, t_p^l) = \begin{cases} \lambda \cdot (t - t_p^l), t_p^l \leq t < T_i^f \\ U_i(t_p^l), t \geq T_i^f \end{cases} \tag{4.5}$$

where $T_i^f = \frac{U_i(t_p^l)}{\lambda} + t_p^l$ is the time when v_i reaches the maximum privacy loss after the last pseudonym changing. Based on (4.5), the privacy level of v_i at the time instant t is given by

$$U_i(t) = U_i(t_p^l) - \beta_i(t, t_p^l), t \geq t_p^l. \tag{4.6}$$

Based on each $U_i(t), \forall i \in \nu(j)$, RSU_j determines whether it needs to change the pseudonym for each vehicle $v_i, \forall i \in \nu(j)$. Suppose the pseudonym change cost of v_i is denoted by c_i,

which may include the cost of communication and pseudonym management. For all the vehicles under the coverage of the same RSU, assume $c_i = c(t), \forall v_i \in \nu(j)$. Based on (4.5) and (4.6), RSU_j can calculate the privacy level of each vehicle node $U_i(t)$ in real time. When v_i changes its pseudonym at time t, the privacy level of v_i is recorded as $U_i(c, AS) = \log_2 AS_i - c$ by RSU_j.

For each vehicle v_i, when $U_i(c, AS) \geq U_i(t)$, i.e., $AS \geq AS^i_{th}$, $AS^i_{th} = 2^{U_i(t)+c}$, the pseudonym changing procedure would increase the privacy level. Otherwise, the pseudonym changing is useless. In the MAP schemes, each RSU makes the pseudonym changing decision for each vehicle based on AS^i_{th} and AS_i according to the privacy level of the vehicle. RSU_j sorts the privacy level of each vehicle in $\nu(j)$. For the vehicle v_i, if $N_\gamma(U_i(t)) \geq AS^i_{th}$ and $G_i \geq G_{th}, \exists V_k \in \gamma(U_i(t))$, then the all the vehicles in $\gamma(U_i(t))$ need to change pseudonyms simultaneously. If the two conditions are not met at the same time, v_i does not need to change its pseudonym, and then RSU_j carries out the same decision for the next vehicle in $\nu(j)$. The specific pseudonym change coordination algorithm is described by Algorithm 7 in Appendix A.2.

4.3.6 The Pseudonym Change Procedure

In the MAP scheme, the group signature is exploited for the transmission during the pseudonym changing. The group identity GID_j is assigned by the RSU_j to the surrounding vehicles. The pseudonym change procedure includes the invitation from the RSU cloud, the construction of the mix-zones, the release of the transmission slots and pseudonyms of the vehicles, the silent period taken by the vehicles, and use of the new pseudonyms and transmission slots. The algorithm 8 in Appendix A.2 describes the pseudonym change procedure.

When RSU_j determines that the surrounding vehicles need to change pseudonyms, it sets the time window, denoted by T_m, for the vehicles to construct the mix-zone,

and broadcasts GID_j and AS to $\nu(j)$. When v_i receives the pseudonym changing coordination message, it checks whether it needs to change pseudonym. If it decides to change the pseudonym, it sets the *join_ack* field in the header to notify the RSU_j. In the next frame, v_i would keep silent, and the RSU_j uses the time slot ts_i to transmit $res = E_{PK_{ik}}(GID||SK_{G_{ji}}||Cert_{G_{ji}}||RSU_j||Cert_{RSU_j})$. RSU_j shuffles the time slots for the vehicles that change their pseudonyms, and allocates a new time slot to v_i by sending the message, $slot\_alloc = E_{PK_{GID_i}}(ts_i^{'}||Cert_j||timestamp)$, where $ts_i^{'}$ is the new time slot for v_i. When v_i receives the message, it uses the new time slot to transmit the acknowledgement message, $success = RSU_j||E_{PK_{RSU_j}}(ack||SafetyData||timestamp||Cert_{G_{ji}})$, in the next frame. If v_i has not changed its pseudonym at time t when $U_i(t)$ reaches 0, v_i would releases the current time slot, and randomly selects a silent period from $[ts_{min}, ts_{max}]$ to keep silent. When the silent period expires, v_i starts to reserve a new time slot based on the SI received from the RSU_j.

The following corollaries can be obtained for the pseudonym change coordination algorithm in the MAP scheme.

Corollary 1. *The above pseudonym change coordination algorithm can guarantee that the privacy level of each vehicle is always above 0.*

Proof. Firstly, if v_i has not changed its pseudonym at time t when $U_i(t)$ reaches 0, its privacy level would be increased by taking the silent period to mix together with other newly entering vehicles. Secondly, consider the pseudonym changing procedure for multiple rounds. Suppose that anonymity set size during the first round of pseudonym changing is denoted by n_0. In other words, the privacy level of the n_0 vehicles satisfies that $U \geq \log_2(n_0) - c(t)$. After the first round of pseudonym changing, the privacy level of the n_0 vehicles increases, and the difference of the privacy level between the vehicles in $\nu(j)$ decreases. Thus, after the privacy loss during the traveling, the anonymity set size during the pseudonym changing in the next round, denoted by n_1, would increases, i.e., $n_1 \geq n_0$.

Consequently, we conclude that the privacy level of the vehicles keeps to be larger than 0. ■

Corollary 2. *Given the network is stable, the anonymity set size level of the vehicles can converge, and the privacy of the vehicles reaches to the stable level during the pseudonym changing.*

Proof. Based on corollary 1, the privacy level of the vehicles becomes closer to each other after the pseudonym changing procedure. The pseudonym changing decision making is carried out by the RSU, however, the number of vehicles participating in the simultaneously pseudonym changing is controlled by the following equations with respect to k,

$$\begin{cases} 1 \leq k \leq N_\nu, \\ \nu[k] \leq u, \end{cases} \tag{4.7}$$

where N_ν denotes the number of elements in $\nu(j)$. Denote the descending sorted set of the privacy level of all the vehicles in $\nu(j)$ by $\mathbb{N}_\nu$, and $\nu[k]$ indicates the k_{th} element in $\mathbb{N}_\nu$. The solution set to (4.7) is denoted by U_s. And then, the convergence anonymity set size, denoted by N^*, is calculated as $N^*_{\gamma(u)} = \max(U_s)$, when U_s is not empty. It indicates that the efficiency of pseudonym changing procedure increases, and finally reaches to a largest level when all the surrounding vehicles can achieve a larger level of privacy. ■

4.4 Performance Evaluation

This section analyses the efficiency of the channel accessing, security and privacy preserving in the MAP scheme. The calculation of the packet overhead is similar to 3.4.5 for the MAP scheme, thus, the detailed calculation procedure is omitted for the MAP scheme. The time interval of each time slot in the CCH is set to 0.25ms, which enables the transmission of up to 180 nodes within a frame. The transmission range of each vehicle is

controlled by the local RSU, thereby, so that the number of nodes within two hop set can be adjusted, to guarantee that each node can reserve at least one slot to transmit with no transmission collisions.

4.4.1 The Default Channel Switching Approach

Dynamic channel switching approach can improve the channel resource utilization efficiency, however, in the TDMA based MAP scheme, the unbalanced traffic density causes the channel resource allocation conflict in the dynamical channel switching approach, since the vehicle density within the communication range of different RSUs could be diverse. This problem is difficult to solve through the coordination between the RSUs. The scenario shown in Fig. 4.5 illustrates the problem. In the case of unbalanced traffic flow, assume that there are more vehicles traveling from left to right than the vehicles traveling from right to left. Suppose there are 17 cars in the two-hop range of RSU_2, and the red-marked vehicle, occupies the $16th$ time slot. While there are only 10 nodes in the two-hop range of RSU_4, if utilizing the dynamical channel switching, the CCH interval controlled by RSU_4 would be shorter than that controlled by RSU_2. When the vehicles travels to the coverage of RSU_4, its transmission could not be heard by RSU_4, since RSU_4 may have switched to the SCH. Thus, considering the stable time slot reservation for the vehicles, the default channel switching (static channel switching) approach is utilized in the MAP scheme. The RSUs can adjust the transmission range of the vehicles based on the traffic density. As the vehicle density increases, the coverage of the RSU and the transmission range of each vehicle should be reduced, so as to provide a high quality of service for the transmission of the safety messages.

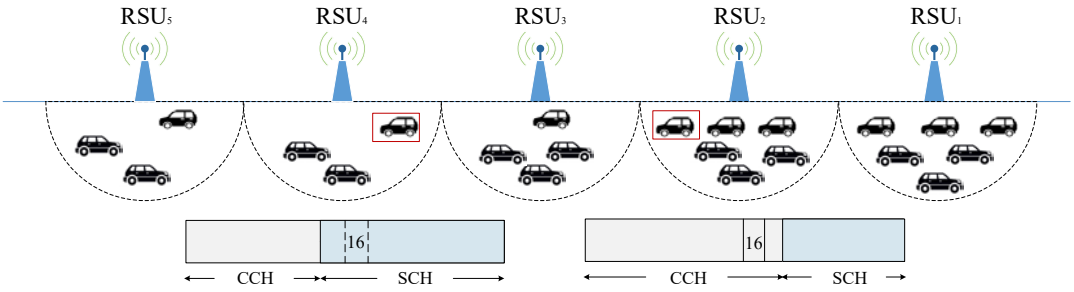


Figure 4.5: The channel coordination conflict caused by the diverse traffic density under different RSUs

4.4.2 Security Analysis

The MAP scheme provides the message integrity, authentication, traceability and defend against various attacks through the standard cryptographic primitives including the asymmetric key-based encryption. With the help of digital signatures, all entities verify the digital signature to ensure the message integrity. With the certificate in the messages, the sender is authenticated and verified. The encryption mechanisms ensure that the adversary cannot open the encrypted messages by launching brute-force attacks. The replay attacks would not be successful due to the usage of timestamp. Meanwhile, the adversary cannot simulate an RSU or forge the RSU messages as the RSUs are connected and controlled by the OpenFlow Controller by the central cloud. In order to defend against Sybil attacks, we consider that each pseudonym has a expiry time and the new pseudonyms are generated and distribute to the vehicles by the RSU cloud.

Based on the statement of corollaries 1 and 2, the RSUs can coordinate the vehicles to change their pseudonyms with sufficient neighbors and increase their privacy level. The privacy level of each vehicle is always larger than 0, so that the adversary always needs time to link the new pseudonym to the old pseudonyms of the vehicle. Furthermore, the group-confusion privacy metric is taken into consideration for the pseudonym changing of each vehicle, when constructing the mix-zones, it is difficult for the GPA to continuously

track a vehicle as it can not improve its posterior knowledge based on the vehicle's location information. All these vehicles are similar to the target vehicles, if the vehicles with indistinguishable location information change their pseudonyms and transmission slots simultaneously in the mix-zones. Thus, the time-to-confusion [80] of the adversary to track a vehicle is constrained by the vehicle's pseudonym age and its transmission slot usage duration. The time-to-confusion measures the duration that an adversary could correctly follow the trace until it could not determine the next sample of a vehicle with sufficient certainty. In contrast, if the MAC layer transmission is not considered when changing pseudonyms, the time-to-confusion could be unbounded by the pseudonym age as the MAC layer semantic linking attack facilitates the linking of the new and old pseudonyms of the targeted vehicles.

4.5 Simulation

4.5.1 Simulation Scenario

In order to evaluate the performances of the proposed scheme, extensive simulations are conducted using Matlab by considering a city grid layout scenario. As shown in Fig. 4.6, it consists of 16 square city blocks in a 10 km^2 region. There are 9 large intersections deployed with RSUs that are connected together. The central cloud is simulated to verify the identities of all the nodes. The adversary can observe the transmission of all the nodes in the network continuously, and try to track a randomly selected vehicle by collecting its transmitted messages, analyzing its transmission patterns and linking its pseudonyms. The transmission range of the RSUs and the vehicles is assumed to 250m. Each street has two lanes with opposite driving directions. Vehicles with larger speed can go ahead of the vehicles with lower speeds. Assume when a vehicle reaches a junction area, it chooses on of all possible directions randomly. The traffic light is deployed at each intersection,

with the duration of 1 minute. These vehicles always move within the simulation area and never drive out. The vehicle located at an intersection can communicate with the vehicles at both directions near the intersection. Each vehicle moves with a constant speed drawn from the truncated normal distribution.

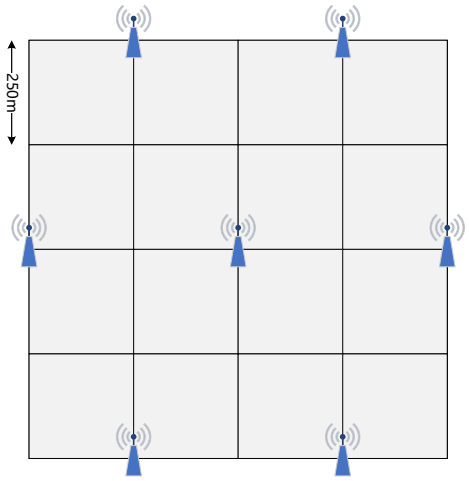


Figure 4.6: The simulation scenario for the MAP scheme

During the simulation, the vehicles are divided into two group, 80% start at the beginning of simulation to listen to the transmission of the RSUs and reserve the transmission slots, and the others are randomly activated during the following simulation. Each vehicle is preloaded with 5 pseudonyms. Other simulation parameters are showed in Table 4.2.

We measure both the location privacy in terms of the average pseudonym age, anonymity set size and the time-to-confusion, and the QoS of the safety message transmission in terms of the collision rate and the PDR of the MAP scheme under different traffic densities. The collision rate is defined as the number of collisions happened per vehicle per frame. The pseudonym age starts to count when the pseudonym is utilized to transmit the first message.

The performance of the MAP scheme is compared with other three different cross-layer schemes.

Table 4.2: Simulation Parameters in the MAP Scheme

Parameters	Value
Lane length	500m
Lane width	5m
The number of lanes	2
The total number of vehicles	[200:50:800]
The duration of each slot	0.25 ms
The cost of pseudonym changing	0.5
The privacy loss rate	0.2
ts_{min}	0.5 second
ts_{max}	1 second
The driving duration	20 minute
R_p	20m

- NMAP: vehicles apply the same pseudonym changing and channel accessing strategies as defined by the MAP scheme. However, the transmission slots are not shuffled by the RSUs when the vehicles change pseudonyms. Each vehicle re-selects new time slot to reserve when it receives pre-warning of collision or detects a transmission collision.
- PCS [58]: in the effective pseudonym changing at social spots (PCS) strategy, vehicles stop transmitting the safety messages when they drive across the intersection, and begin to construct the mix-zones. After vehicles leave the mix-zone, they start to utilize the new pseudonyms to transmit in a distributed manner. In the MAC layer, the standard IEEE 802.11p is applied.
- VeMAC [46]: the effective TDMA based VeMAC protocol is applied for the MAC layer operation while the pseudonym changing is periodic for each vehicle. In the VeMAC, three disjoint sets of time slots are assigned to the vehicles moving in opposite directions, and RSUs.
- CFR-MAC [98]: in the effective TDMA based CFR-MAC protocol, the time slots

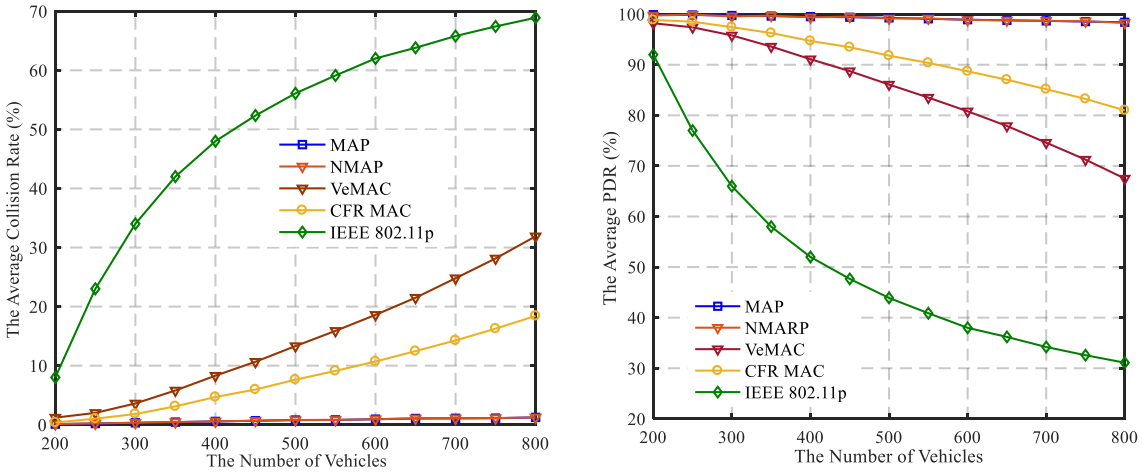
are divided into two disjoint sets for the vehicles moving in opposite directions, and each set is further divided into three subsets for the vehicles with different speed levels.

4.5.2 Simulation Results

Fig. 4.7a shows the collision rate in schemes by using the different MAC protocols. The simulation result shows that the transmission collision rate in the MAP scheme is close to 0, the collision rate of the CFR MAC protocol is slightly lower than that in VeMAC, and the collision rate in the IEEE 802.11p protocol is the highest. In the MAP scheme, the local RSU cloud can predict the transmission collision between vehicles, and the available time slots can always be allocated by the RSUs to the vehicles that need to reserve time slots. It is worth noting that the number of vehicles that need to reserve the new time slots per frame is relatively small, so that the accessing collision rate in MAP is close to zero. In VeMAC and CFR MAC, the merging collision that caused by the vehicle mobility cannot be avoided. Vehicles that collide with each other need to re-select the new time slots to reserve, which may further cause an accessing collision by choosing the same slot. As the number of vehicles increases, the collision rate increases. At the same time, the number of the available time slots is reduced, thereby causing the accessing collision problem become more serious.

Fig. 4.7b shows the average PDR obtained in the different schemes. The PDR is mainly affected by the packet loss rate that caused by three main reasons, the transmission collisions, the accessing collisions when the vehicles reserve new transmission slots, and the coordination and silent period during the pseudonym changing procedure taken by the vehicles. In MAP scheme, the vehicle nodes nearly do not have transmission collisions with the coordination from the RSUs. However, when the vehicles changing pseudonyms, they need to suspend their transmission by receiving the coordination message from the RSUs,

and keeping silent during the silent period to confuse the adversary. In VeMAC and CFR MAC, as the number of vehicles increases, the transmission collisions caused by merging collisions become the main reason for the degradation of the PDR in the IoV. In the IEEE 802.11p, the pseudonym changing has little interference on the PDR of the safety messages.



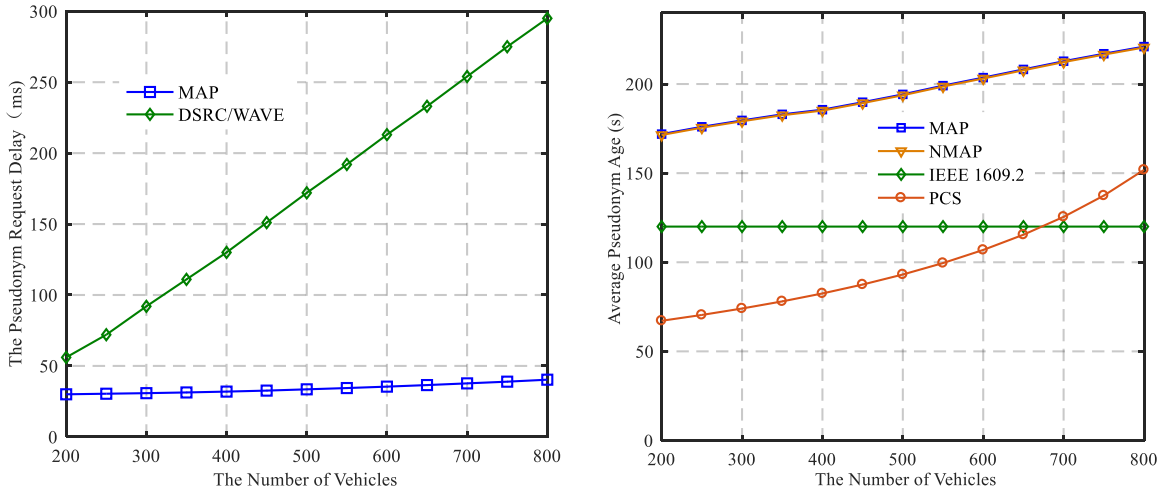
(a) Collision Rate (b) Average PDR

Figure 4.7: The average transmission collision rate and PDR in different schemes.

In the simulation, the average delay from the vehicles transmit the pseudonym request messages to receive the set of new pseudonyms is evaluated, and compared with other schemes, as shown in Fig. 4.8a. It shows that in the MAP scheme, the pseudonym request delay can always be kept within the interval of a frame, which indicates that the MAP scheme can ensure that the vehicles always have enough pseudonyms for the safety message transmission, and achieve the privacy preservation. In the MAP scheme, the efficient channel coordination ensures the transmission reliability of both the safety messages and the pseudonym request messages. Whereas in the DSRC/WAVE protocol, it shows that when the traffic density is high, the pseudonym request delay is increased. It takes more than one frame for the vehicles to obtain the new pseudonyms, which is a risk

to preserve the location privacy of the vehicles, and may affect the real-time transmission of the safety messages.

In Fig. 4.8b, the pseudonym age in the different schemes are investigated. The pseudonym age in the MAP and NMAP scheme is almost the same, and influenced by the pseudonym change decision made by the RSU cloud. The IEEE 1609.2 protocol applies the periodical pseudonym changing approach, thereby the pseudonym age staying constant. In the PCS scheme, the frequency of the pseudonym changing is decided by the frequency that vehicles encounters the road intersections. The scheme may be not efficient to preserve the location privacy for the vehicles, since it much depends on the traffic topology and the traveling route of the vehicles in the IoV. If the vehicles encounters the intersections too frequently, it leads to the waste of pseudonyms. On the other hand, if the distribution of intersections is sparse, it leads to a longer duration for the adversary to track the vehicle, and even compromise the location privacy of the vehicles.



(a) Pseudonym Request Delay

(b) Average Pseudonym Age

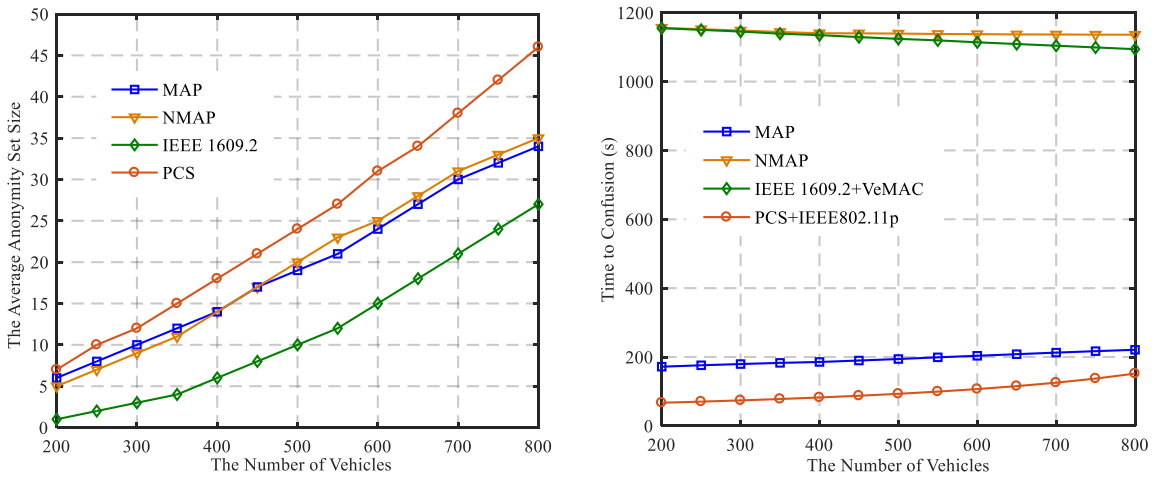
Figure 4.8: The average pseudonym request delay and pseudonym age in different schemes.

Fig. 4.9a shows the average anonymity set size when changing pseudonyms in the different schemes. The PCS scheme obtains the largest anonymity set size, since all the

vehicles driving across the intersections change their pseudonyms simultaneously. The MAP and NMAP schemes achieves almost the same results. In the IEEE 1609.2, when the traffic density is low, the vehicles may even change the pseudonyms with no cooperative neighbors, i.e., the anonymity set size is 1. It cannot guarantee the unlinkability of the pseudonyms, thereby compromising the location privacy of the vehicles.

Generally, when the pseudonym changing procedure could confuse the adversary, the time-to-confusion of the adversary is constrained by the pseudonym age of the vehicle. However, as we pointed out, the MAC layer context linking attack provides the clue for the adversary to track the vehicle continuously by linking the traffic pattern of the same vehicle. Thus, the design of the pseudonym schemes should guarantee the unlinkability of the pseudonyms, and make the time-to-confusion of the adversary be bounded by the pseudonym changing procedure. It is crucial to evaluate the time-to-confusion of the pseudonym schemes. The simulation result shows the effect of the MAC layer operation on the location privacy performance of different schemes. As Fig. 4.9b illustrates, the time-to-confusion is limited in only the MAP and PCS schemes. In the MAP scheme, vehicles change the pseudonyms with the transmission slots during the silent period, which cuts the linking of the new pseudonym and old pseudonyms of vehicles in the network. In NMAP, vehicles change pseudonyms without shuffling the transmission slots. Vehicles usually acquire separately slots under the coverage of the RSU, and the mix-zone construction of each vehicle is carried out by the same RSU. Thus, the adversary can distinguish each vehicle based on its transmission slot index. Therefore, the time-to-confusion is limited by the pseudonym age. In the NMAP and IEEE 1609.2 schemes, even though the vehicles change pseudonyms, the adversary can still track the targeted vehicle continuously.

The simulation results conclude that the MAP scheme can guarantee the reliable, secure and privacy preserving of the nodes, by the efficient coordination of both the channel resources and pseudonyms in the IoV.



(a) Average Anonymity Set Size

(b) Average Time-to-Confusion

Figure 4.9: The average anonymity set size and time-to-confusion of the adversary in different schemes.

4.6 Conclusion

In this chapter, we have presented the SDN based three-layer system architecture, namely SDIV for the IoV. In the SDIV, to manage the allocation of both the channel resources and pseudonyms, a cross-layer scheme, MAP, is proposed in the SDIV to provide reliable and privacy preserving transmissions for the nodes in the network. There are two tiers of the SDIV for the MAP scheme to make the coordination decisions efficient. The controller in the central cloud tier schedules the time slot information sharing between the RSUs, and decide the transmission range for the vehicles. The RSUs act as the local controller for the vehicles to coordinates vehicles to adaptively change pseudonyms and access the wireless channels in a contention-free manner. The MAP scheme can provide near-collision free transmission of the safety messages for the vehicles in the IoV. The multi-channel transmission of the safety messages and security related messages ensures that the safety message transmission is not affected by the security management, and also

does not affect the privacy preserving of the vehicles in the network. By measuring the privacy level of each vehicle in real time, the RSU cloud makes the pseudonym changing decision for the surrounding vehicles, and leverages the group signature for vehicles to change their pseudonyms synchronously. The security analysis and simulations results verify the effectiveness of the MAP scheme. To the best of our knowledge, most previous work considered the MAC layer protocol and pseudonym schemes separately. This work shed light on the cross-layer protocol design for the SDN-enabled IoV as the security and privacy preservation must be provided for all the layers.

The future work has planed to be carried out in two aspects. Firstly, we will consider how to coordinate the transmission in the SCHs and optimize the channel resource allocation between the RSUs with diverse traffic densities by the central controller. Secondly, how to manage and share the pseudonym resources between the RSUs with diverse pseudonym demand and privacy requirements by the central controller will be solved.

Chapter 5

Uncoordinated Pseudonym Changes in Distributed Networks

5.1 Introduction

Distributed networks have rapidly gained popularity in practice, such as VANET, wireless sensor network (WSN), mobile social network [27] and blockchains [118], where all network participants disseminate messages in a fully distributed and open fashion. Privacy, which includes identity privacy, unlinkability, non-repudiation, and accountability [119] is a critical concern in distributed networks.

One practical and efficient approach for protecting privacy is to utilize pseudonyms to represent participants without disclosure of their real identities [27]. Typically, pseudonymous communication is achieved by public key infrastructure (PKI) based cryptography that assigns each individual a set of public key certificates and corresponding key pairs, thereby providing authentication and data integrity while protecting identity privacy. Moreover, the PKI based pseudonym scheme is expected to resist a Sybil attack, wherein a malicious node pretend to be other nodes with multiple identities [56].

Frequent and unlinkable pseudonym changes are the key to prevent de-anonymization, as adversaries could continuously observe the pseudonyms and correlate the observed pseudonyms [55]. Simple pseudonym changes may not be enough to guarantee the unlinkability of pseudonyms, as the syntactic linking attack can be performed to link the new and old pseudonyms when a node changes the pseudonym alone. When a node needs to change its pseudonym, it needs to mix together with at least $k-1$ peers, so as to achieve the k-anonymity [120]. The set of the total k nodes is referred to as an anonymity set, and k is the anonymity set size. The anonymity set size, defining the number of nodes which change pseudonyms together, is one of the most widely adopted metrics to measure the privacy of nodes. The larger the anonymity set is, the better identity privacy can be achieved. Other metrics, such as entropy [66, 82], traceability [90] and time-to-confusion [80], can be evaluated based on the anonymity set size, the similarity of users' context and the adversary's capability.

There have been a number of schemes developed in an attempt to achieve efficient and unlinkable changes of pseudonyms in distributed networks. In VANET, mix-zones [121] are exploited by vehicles in proximity to change their pseudonyms concurrently. In mobile crowdsensing networks, pseudonym and mix-zone based privacy preserving framework [122] is proposed to protect the users' identities and trajectories to the application servers. To protect the transaction privacy in Bitcoin [118], approaches such as Zerocash [123] and CoinShuffle [124] are developed to mix pseudonymous transactions so as to prevent the linkability of transactions made by the same user. Most of existing solutions achieve privacy, either with the assistance of third trusted parties (TTPs) [66, 121, 125], or based on pre-determined time or locations [58, 126], which could compromise flexibility, and be susceptible to insider attacks.

Distributed pseudonym change strategies have been advocated to allow the nodes to change pseudonyms by negotiating with each other. In [127], distributed pseudonym

change is achieved by each vehicle via broadcasting cooperation request to its neighbors to construct mix-zones in vehicular networks. In [128], a distributed location privacy preserving strategy was proposed for delay tolerant networks, where each node can generates multiple virtual nodes to increase the anonymity set size. Florian et. al [129] designed an approach for the Bitcoin, namely BitNym, where pseudonyms were generated only dependent on the transmission volumes of the nodes to prevent Sybil attacks, wherein a malicious node pretend to be other nodes with multiple identities [56].

As pointed out by [129], pseudonyms are precious resources due to non-negligible costs of generation, management and storage. For this reason, nodes may decline the cooperation requests of pseudonym changes from the others. Incentive and game theoretic techniques were adopted to motivate nodes to cooperate [130–132]. Unfortunately, the cooperation or negotiation among the nodes could leak side information, and semantic linking attacks [133] could be carried out to link pseudonyms of a node based on the context of the node. In [81], the quality of privacy achieved in a distributed pseudonym change strategy was analytically studied for vehicular networks, where a vehicle did not change its pseudonym until it encountered at least one other vehicle to change together. In [83], the cooperation probability of nodes was evaluated for a distributed pseudonym change strategy, where each node still needed to broadcast pseudonym change invitations to achieve distributed cooperation.

The changes of pseudonyms are associated with non-negligible costs. The generation and management of pseudonyms can be costly to achieve the non-repudiation and accountability. Petit et. al [55] point out that the lifecycle of an efficient pseudonym scheme consists of the pseudonym issuance, use, change, resolution and revocation. Changing pseudonyms could also affect the transmission performance [89]. However, existing analyses of the impact have been based on centralized pseudonym changes, or focused on application (layer) performance with little consideration on the implementation of pseudonym

changes. K. Emara [90] studied the effect of pseudonym changes on the safety applications, such as forward collision warning and lane change warning in vehicular networks. In [91], the pseudonym management cost was analyzed in terms of communication delay. In [92], the impact of pseudonym change and silent period was evaluated on an intersection collision avoidance (ICA) system for driving assistance.

To the best of our knowledge, only a few studies, e.g., [61, 81], have to date designed fully uncoordinated pseudonym change protocols for implementations in distributed networks. Neither of the studies have analyzed the impact of uncoordinated pseudonym change protocols on the transmission performance. In [61], the periodic pseudonym change approach is applied for the VANET. In [81], an analytical model was proposed to calculate the anonymity set size, given the age of pseudonyms and broadcast rate, where the nodes changed pseudonyms randomly without coordination or negotiation. The analytical model studied the expected anonymity set size of each vehicle, and the conditional probability of a node changing its pseudonym together with its neighbors. The model did not consider any silent period. The impact of the pseudonym change on the network capacity was not studied.

Distinctively different from these existing researches, e.g., [61, 81], this chapter presents a fully uncoordinated pseudonym change approach in distributed networks, where multiple nodes can suspend their transmissions and then randomly change pseudonyms to confuse potential adversaries, and hence prevent the adversaries from establishing connections between the pseudonyms. Given the network size, pseudonym lifetime, and pseudonym change strategies, the silent period is analytically configured, under which the k-anonymity is achieved without any cooperation between the nodes. Specifically, we develop a new continuous-time Markov model to analyze the time-varying population of changing pseudonyms by using the Kurtz's theorem [101]. In turn, the silent period can be minimized, and so can the capacity loss resulting from the pseudonym changes. The

key contributions of this chapter are summarized as follows.

- In this chapter, a new analytical model is developed to analyze the uncoordinated changes of pseudonyms. The time-varying population of changing pseudonyms is proved to converge to a stationary distribution in the case of a large number of nodes in a network.
- Given a privacy requirement of k-anonymity, critical conditions are analytically established, under which the nodes can independently change their pseudonyms, while still preserving privacy.
- With new proposed pseudonym change policies, the proposed model is demonstrated to help configure policy parameters to minimize the change delay and the throughput loss resulting from the delay, while preserving k-anonymity privacy.
- Validated by simulations, the accuracy of the proposed model improves, as the number of nodes, N increases in the network. It is also shown when N is large, the k-anonymity can be achieved in the uncoordinated fashion at negligible throughput loss.

The rest of this chapter is organized as follows. In section 5.2, the system model is described. In section 5.3, the proposed model of the pseudonym aging and changing is elaborated on, based on which the conditions of signaling-free pseudonym changes and the optimal pseudonym change strategies are established. In section 5.4, numerical results are provided, followed by conclusions in section 5.5.

5.2 System Model

5.2.1 Attacker Model

In this chapter, we consider a GPA, denoted by $\mathcal{A}$, which is assumed to be powerful with visibility to all data packets disseminated in the network. The GPA can carry out the linking attack in an attempt to de-anonymize the nodes by trying to establish the connections between the pseudonyms of every node. Consider that messages are broadcast in distributed networks with little coordination between peers. The GPA intends to establish connections between the different pseudonyms of a node.

The tracking uncertainty of the GPA is measured by the privacy level of a node in the network during its pseudonym change. Suppose that a node, e.g., node i, changes its pseudonym at time T with the anonymity set size of k. U_i denotes the uncertainty level of the GPA on the linking result of node i. By using the concept of entropy, U_i is given by [82]

$$U_i(T) = -\sum_{j=1}^{k} p_{j|i} \log_2(p_{j|i}), \tag{5.1}$$

where $p_{j|i}$ is the probability that the current pseudonym of node j is linked to the previous pseudonym of node i, and $\sum_{j=1}^{n} p_{j|i} = 1, \forall i$. The larger the value of U_i is, the more effort is required for the GPA to perform an effective linking attack. Given the anonymity set size k, we have $U_i \leq \log_2 k$, where $\log_2 k$ is obtained when $p_{j|i} = 1/k, \forall j$. The capability of the GPA is measured by the privacy loss rate of the nodes, denoted by β. The more capable of the GPA, the higher privacy loss of the nodes in the network. Assume the tracking uncertainty of the GPA against node i is $U_i(t)$ at time t, and after time $\Delta t = \frac{U_i(t)}{\beta}$, the privacy level of the node i becomes 0. It means that, given the privacy loss rate of the nodes, the pseudonym age is bounded by the duration from the last pseudonym changing to the privacy level becomes 0.

5.2.2 Uncoordinated Pseudonym Change

As illustrated in Fig. 5.1, we consider a distributed network, where there are N nodes. When one of the N nodes transmit a message, all other nodes can receive the message. We assume there is a certificate authority (CA) at the TTP. The CA is responsible to assign pseudonyms and corresponding certificates to all participating nodes in the network. When a node, denoted by u_i, joins the network, it registers at the CA to obtain its master certificate, denoted by $CERT_i$, and the corresponding key pair $\{PK_i, SK_i\}$ from the CA. After confirming the eligibility of u_i, the CA generates a set of pseudonyms for u_i. To thwart Sybil attacks, each pseudonym is set up with a predefined lifetime and only one pseudonym can be utilized at a time by the node. We define each pseudonym as $\{PID_i, t_i^1, t_i^2\}$, where PID_i is the pseudonym of node u_i, t_i^1 is the activation time of the pseudonym, and t_i^2 is the expiry time of the pseudonym. The node can transmit messages under a pseudonym continuously until the pseudonym expires, and then suspends its transmission and replaces the pseudonym with a new one. The CA maintains a local database to manage the identity information $\{RID_i, PK_i, SK_i, CERT_i, PID_i\}$ for each node u_i. When u_i runs out of pseudonyms, it uses $\{PK_i, SK_i\}$ and $CERT_i$ to communicate with the CA securely and request new pseudonyms.

We assume the pseudonym change strategy is specified by the TTP, and adopted by all the N nodes in a fully distributed, uncoordinated manner. The CA can estimate the capability of the GPA, and determine the appropriate pseudonym lifetime τ and corresponding k-anonymity requirement. The specific pseudonym change strategy is defined by a function/policy that maps the pseudonym age to the pseudonym changing probability. Let $p_i(z)$ denote the probability density function (PDF) under which node i, denoted by u_i, changes its pseudonym at age z. Define $t_u < \infty$ as the maximum value of the pseudonym age, $\int_0^{t_u} p(z) = 1$. Distinctively different from any existing approach, we design that each node tracks its own pseudonym age and decides to change its current pseudonym based

on $p(z)$ without further coordination of the network or negotiation with the other nodes, thus ensuring that the pseudonym change process does not leak any information.

We note that the CA and TTP have been the bottleneck for distributed networks. Nevertheless, they are crucial to achieve non-repudiation, accountability and certificate revocation in the networks. We propose to apply the approach that the nodes carry out the pseudonym change procedures in a distributed manner, without coordination from the CA or TTP. Whereas the nodes still need to be authenticated by the CA and the pseudonyms still need to be assigned by the TTP. Moreover, a number of pseudonyms can be preassigned to each node, and each pseudonym can be utilized effectively towards the end of its lifetime. By this means, the interactions between the nodes and TTP are dramatically reduced, while the non-repudiation and accountability can still be provided.

A potential application for the uncoordinated pseudonym change is future Internet-of-Things (IoT), where thousands of, or even tens of thousands of sensors can collect sensitive data and upload the data in an asynchronous manner with no coordination between the sensors. The identity privacy and the location privacy of the sensors is important to protect source location privacy. Another potential application is to prevent adversaries from tracking vehicles and personnel. Continuously growing data exchange between vehicles/personnel and the Internet has been increasingly facilitating eavesdropping and linking attacks on the vehicles/personnel. In both applications, non-repudiation and accountability of the sensors or vehicles/personnel are important, so that the sensors, vehicles and personnel are expected to be traceable at the TTP.

5.3 Proposed Analytical Model

In this section, a new analytical model is proposed to analyze uncoordinated changes of pseudonyms in the network of interest. We are particularly interested in the inter-

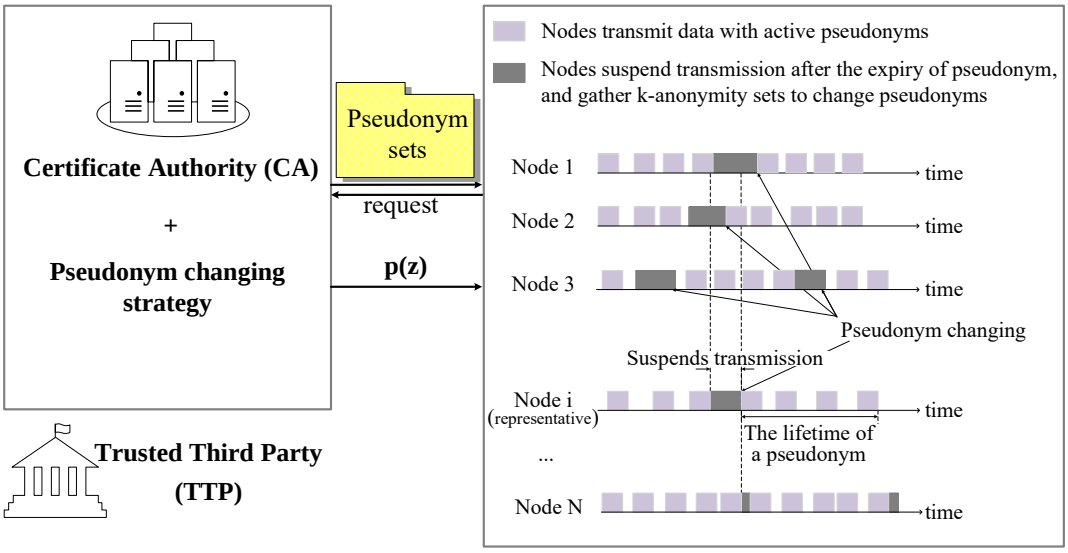


Figure 5.1: An illustration of the system of interest, where the pseudonyms of the nodes change in an uncoordinated, distributed fashion.

action between pseudonym lifetime, pseudonym change strategy, the number of nodes, and the privacy requirement of the nodes. In the proposed model, the evolution of users' pseudonym age is captured. This facilitates configuring pseudonym change policies with which privacy can be protected through independent changes of pseudonyms without coordination or negotiation. The impact of the pseudonym change and privacy requirement on the throughput is also analyzed. Notations used in this chapter are listed in Table 5.1.

5.3.1 Modeling of Pseudonym Aging and Changing

Let τ_i denote the pseudonym lifetime of node u_i, and $Z_i(t)$ denote the age of the current pseudonym, $Z_i(t) = t - t_i^c$, where t is the current time, and $t_i^c \leq t$ is the time of the last pseudonym changed. When the current pseudonym expires, i.e., $Z_i(t) = \tau_i$, the node suspends its transmission and changes the pseudonym with the probability $p_i(z)$ at age $Z_i(t) = z \geq \tau_i$. For illustration convenience, we consider a homogeneous system with the same pseudonym lifetime τ and pseudonym change strategy $p(z)$ for every node. The

subscript "$_i$" is suppressed for brevity.

Table 5.1: Notation and Description in Chapter 5

Notations	Description
N	The number of nodes in the system
k	The anonymity set size required for privacy preservation
u_i	A representative node i
$p(z)$	The PDF of changing pseudonym at age z
$P(z)$	The CDF corresponding to $p(z)$
$q(z)$	The PDF of the pseudonym has not changed before its age reaches z
$f(z\|q(z))$	The conditional PDF that a pseudonym changes at age z under the condition that it has not changed before
τ	The lifetime of each pseudonym
t_u	The maximum value of pseudonym age
n_i	The anonymity set size achieved by u_i
$Z_i(t)$	The pseudonym age of u_i at time t
$\mathbf{Z}^{(N)}(t) = \{Z_i(t)\}_{i=1}^{N}$	The pseudonym age of all nodes in the system at time t
$R_i(t)$	The instantaneous transmission rate of u_i at time t
$\gamma(z)$	The throughput loss ratio caused by pseudonym changing at age z
$M(z,t)$	The occupancy measure of pseudonyms with age z at time t in the system
$F_M(z,t)$	The CDF of the pseudonym age distribution in the system at time t

As illustrated in Fig. 5.2, the evolution of each node's pseudonym age follows two processes: a linearly pseudonym aging process, and a jump process. In the pseudonym aging process, the pseudonym age is shorter than τ, and it increases as the time elapses. When the pseudonym age is equal to or older than τ, the pseudonym changes at age z

with PDF $p(z)$, and the pseudonym age jumps to 0. After changing the pseudonym, the evolution of the pseudonym age transits to the aging process. Therefore, $Z_i(t), \forall i \in [1, N]$ is a continuous time, recurrent and irreducible Markov process.

Our approach for the theoretical development is first to prove that, given an arbitrary initial distribution of the ratio of the pseudonyms with age z at time 0, denoted by $M(z, 0)$, the evolution of the ratio over time, $M(z, t), \forall z \in [0, t_u]$ asymptotically converges to a density dependent jump Markov process with Lipschitz property. In other words, $M(z, t)$ asymptotically converges to a stationary process, denoted by $m(z)$, as the number of nodes in the network rises. Then, we derive the stationary distribution, $m(z)$, and prove its uniqueness. We also establish the dependence of the stationary distribution of the pseudonym age under the pseudonym change strategy $p(z)$ and the pseudonym lifetime τ, which is key to evaluate the anonymity set size of a node in the network, as will be articulated in Section 5.3.2.

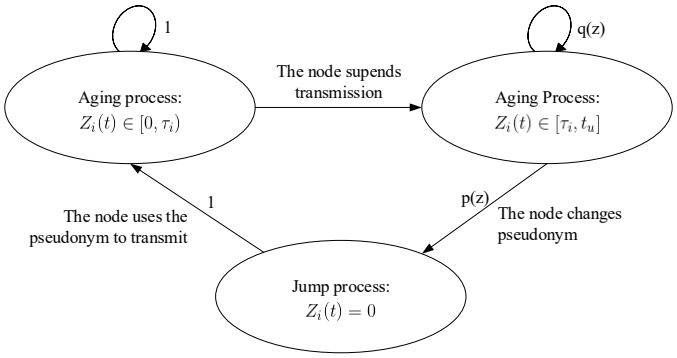


Figure 5.2: The pseudonym evolution process of each node

Denote the pseudonym ages of all nodes at any time t as a label-free, orderless set, $\mathbf{Z}^{(N)}(t) = \{Z_1(t), Z_2(t), \cdots, Z_N(t)\}$. Therefore, for any permutation $\sigma(\cdot)$ of the indexes $\{1, 2, \cdots, N\}$, we have

$$\begin{aligned} \mathbf{Z}^{(N)}(t) &= \{Z_1(t), Z_2(t), \cdots, Z_N(t)\} \\ &= \{Z_{\sigma(1)}(t), Z_{\sigma(2)}(t), \cdots, Z_{\sigma(N)}(t)\}. \end{aligned} \tag{5.2}$$

Based on (5.2), $\mathbf{Z}^{(N)}(t)$ collects the ages of all pseudonym, and it is an N-dimensional continuous time Markov process. The evolution of $\mathbf{Z}^{(N)}(t)$ can be interpreted as the transition rate of the ratio (or population) of the pseudonyms with age $z, \forall z \in [0, t_u]$ at any given time t. The ratio of the pseudonyms with age z at time t, $M(z,t)$, is given by

$$M(z,t) = \frac{1}{N}\sum_{i=1}^{N}\delta(z), \tag{5.3}$$

where $\delta(\cdot)$ stands for the Dirac measure, and $M(z,t) \in \{0, \frac{1}{N}, \frac{2}{N}, \cdots, 1\}, \forall z \in [0, t_u]$.

Theorem 1. *Given any initial state of $M(z,t)$, i.e., $M(z,0)$, for any $T > 0$,*

$$\Pr\{\lim_{N\to\infty}\sup_{0\le t\le T}\|M(z,t) - m(z)\| = 0\} = 1 \ \ a.s. \tag{5.4}$$

Proof. We prove this theorem by applying Kurtz's theorem [101] to $M(z,t)$, which yields the following properties:

- $M(z,t), \forall z \in [0, t_u]$ *is a density dependent Markov process with state space* $\mathbf{s}(m) = \{0, \frac{1}{N}, \frac{2}{N}, \cdots, 1\}, \forall m(z)$. Each $m(z,t), \forall z \in [0, t_u]$ admits a density of $\mathbf{Z}^{(N)}(t)$, which is a recurrent and irreducible continuous-time Markov process. Thus, $M(z,t)$ is a density dependent Markov process.
- *Convergence of initial condition of* $M(z,t)$: given any $\mathbf{Z}^{(N)}(0)$, the initial condition of $M(z,t)$, i.e., $M(z,0)$, for any $z \in [0, t_u]$, is deterministic.
- *Bounded jump rate*: as discussed, $\boldsymbol{Z}^{(N)}(t)$ is a recurrent and irreducible continuous-time Markov process. Each pseudonym changes at the age of no shorter than τ. Therefore the jump rate of $M(z,t)$ for any $z \in [0, t_u]$ is upper bounded by $1/\tau$, i.e., $\frac{\partial M(z,t)}{\partial t} \le \frac{1}{\tau}, \forall z \in [0, t_u]$.
- *Lipschitz property:* The differential equation for the evolution of $M(z,t), \forall z \in [0, t_u]$, satisfies the Lipschitz condition, of which the definition is given as below [134].

  Definition 1. *A function $f : [a,b] \to R$ is said to satisfy the Lipschitz condition if*

there is a constant L such that

$$|f(x) - f(x^{'})| \leq L|x - x^{'}|, \forall x, x^{'} \in [a, b]. \tag{5.5}$$

Given these properties of $M(z,t)$, Theorem 1 can be readily proved with Kurtz's theorem.

■

By capturing the evolution of $M(z,t)$ by a set of partial differential equations (PDEs), $m(z)$ can be obtained, as presented in the following theorem.

Theorem 2. *Given any initial state of $M(z,t)$, i.e., $M(z,0)$, $M(z,t), \forall z \in [0, t_u]$ has a unique stationary distribution, $m(z)$, which is given by*

$$m(z) = \begin{cases} \dfrac{1}{\tau + \int_{\tau}^{t_u} q(z)dz}, & \text{if } z \in [0, \tau) \\ \dfrac{q(z)}{\tau + \int_{\tau}^{t_u} q(z)dz}, & \text{if } \ z \in [\tau, t_u]. \end{cases} \tag{5.6}$$

Proof. The evolution of $M(z,t)$ satisfies the following PDEs for any $t \geq 0$:

$$\frac{\partial M(z,t)}{\partial t} = \begin{cases} -\dfrac{\partial M(z,t)}{\partial z}, & \text{if } z \in [0, \tau) \\ -\dfrac{\partial M(z,t)}{\partial z} + M(z,t)p(z|q(z)), & \text{if } \ z \in [\tau, t_u]. \end{cases} \tag{5.7}$$

We prove Eq.(5.7) by characterizing the evolution of the cumulative distribution function (CDF) of $M(z,t)$, denoted by $F(z,t)$. At any time t, $F(z,t)$ is calculated as

$$F_M(z,t) = \int_0^z M(s,t)ds, \tag{5.8}$$

where $F_M(z,t)$ gives the proportion of nodes with pseudonym age no longer than z at time t, and $F_M(\infty, t) = 1 \ \forall t$. Based on the population dynamics, the evolution of $F_M(z,t)$ can be written as

$$\frac{\partial F_M(z,t)}{\partial t} = -\frac{\partial F_M(z,t)}{\partial z} + \int_z^{t_u} f(s|q(s))M(s,t)ds, \tag{5.9}$$

where $q(z) = 1 - \int_{\tau}^{z} p(s)ds$ is the probability that a pseudonym does not change when its age reaches z; and $f(z|q(z))$ is the conditional PDF that a pseudonym changes at age z under the condition that it has not changed before.

In (5.9), $-\frac{\partial F_M(z,t)}{\partial z}$ is the decrease rate of $F_M(z,t)$: during $[t-\partial t, t)$, the pseudonyms with age in the range of $[z-\partial t, z)$ become older than z. As a result, these pseudonyms are removed from $F_M(z,t)$. $\int_{z}^{t_u} f(s|q(s))M(s,t)ds$ is the increase rate of $F_M(z,t)$: if node i with pseudonym age of no shorter than z changes the pseudonym during $[t-\partial t, t)$, the age is reset to 0, thus it contributes to the growth of $F_M(z,t)$. In contrast, if the user's pseudonym age is shorter than z at t, its pseudonym changes during $[t-\partial t, t)$ and its pseudonym age is reset to 0 which is still shorter than z. This causes no change to $F_M(z,t)$. In this sense, (5.9) captures the changing rate of the population of nodes with pseudonym ages no longer than z.

Since $p(z) = 0$ for $z < \tau$, the corresponding transition rate of $M(z,t)$ can be derived by taking derivative of $F(z,t)$, which is given by (5.7). Based on Theorem 1, $M(z,t)$ has a unique stationary distribution, which is obtained when the transition rate is 0, i.e.,

$$\frac{\partial M(z,t)}{\partial t} = 0, \forall z \in [0, t_u]. \tag{5.10}$$

Moreover, we have $\int_{0}^{t_u} M(z,t) = 1$. By substituting (5.7) into (5.10) and then solving (5.10), we can confirm that (5.10) has a unique solution, as given in (5.6). This conclude the proof of Theorem 2. ■

5.3.2 Anonymity Set Size and Throughput Loss

The anonymity set of each node u_i, denoted by n_i, $n_i \in \mathbb{N}^+$, includes itself and the peers who are indistinguishable during the pseudonym change procedure. Since the network of interest is homogeneous, when u_i changes its pseudonym at time t, the anonymity set size only depends on the distribution of the pseudonym ages in the network, i.e.,

$M(z,t), \forall z \in [0, t_u]$. Based on Theorems 1 and 2, we can establish the following corollaries.

Corollary 3. *Suppose that $N \gg 1$. When node u_i changes its pseudonym at age z with $p(z) > 0$, $\forall z \in [\tau, t_u]$, the anonymity set size of the node becomes stationary and is given by*

$$n_i(z) = \begin{cases} \dfrac{N(\int_\tau^{t_u} q(z)dz + (z-\tau))}{\tau + \int_\tau^{t_u} q(z)dz}, \text{ if } z \in [\tau, 2\tau) \\ N, \text{ if } z \in [2\tau, t_u]. \end{cases} \tag{5.11}$$

Proof. When node u_i changes its pseudonym at age z (with the corresponding PDF, $p(z) > 0$) at time t, the anonymity set size of the node is equal to the number of nodes who suspend transmissions but have yet to change pseudonyms until the expiry of the pseudonym of node u_i, denoted by $n_i^1(z,t)$, and the number of nodes who suspend transmissions while u_i suspends its transmission, denoted by $n_i^2(z,t)$. Therefore, $n_i(z,t) = n_i^1(z,t) + n_i^2(z,t)$, and

$$n_i^1 = N \int_\tau^{t_u} M(z,t)dz, \tag{5.12}$$

$$n_i^2 = \begin{cases} N \int_0^{z-\tau} M(\tau,t)ds, \text{ if } z \in [\tau, 2\tau) \\ N \int_0^{\tau} M(\tau,t)ds, \text{ if } z \in [2\tau, t_u]. \end{cases} \tag{5.13}$$

Based on Theorem 2, when $M(z,t)$ converges to $m(z)$, (5.11) is obtained by substituting (5.6) into (5.12) and (5.13). This concludes the proof. ■

Corollary 4. *Given that n_i changes its pseudonym at age z with $p(z) > 0$, $\forall z \in [\tau, t_u]$, for $N \gg 1$, the distribution of n_i is stationary, and n_i satisfies*

$$n_i \geq \frac{N \int_\tau^{t_u} q(z)dz}{\tau + \int_\tau^{t_u} q(z)dz}, \tag{5.14}$$

$$n_i \leq \min\left\{\frac{N\left(\int_\tau^{t_u} q(z)dz + (t_u - \tau)\right)}{\tau + \int_\tau^{t_u} q(z)dz}, N\right\}, \tag{5.15}$$

$$\mathbb{E}(n_i) = \sum_{n=0}^{N} n \cdot \hat{\pi}(n) = \frac{2N\int_\tau^{t_u} q(z)dz}{\tau + \int_\tau^{t_u} q(z)dz}, \tag{5.16}$$

where $\hat{\pi}(n)$ is the probability of $n_i = n$. From (5.11) *and* (5.16)*, the distribution and expectation of n_i only depend on the pseudonym change strategy $p(z)$.*

Proof. This corollary can be readily proved based on (4) and (9), and the proof is suppressed for brevity. ■

We are particularly interested in the condition under which the required anonymity set size can be achieved without cooperation between the nodes. Based on the above two corollaries, the following corollary is established.

Corollary 5. *Given the k-anonymity requirement, $n_i \geq k$ is required for each node u_i to preserve its identity privacy. The pseudonym change strategy $p(z)$ needs to satisfy the following conditions to achieve the privacy without coordination or negotiation among the nodes:*

$$\begin{cases} \int_\tau^{t_u} p(z) = 1 \\ \int_\tau^{t_u} q(z)dz \geq \dfrac{\tau k}{(N-k)}. \end{cases} \tag{5.17}$$

Proof. This corollary can be achieved by deriving the correlation among the pseudonym lifetime τ, pseudonym change strategy $p(z)$, and the k-anonymity requirement. Specifically, the minimum anonymity set size of each node should be no smaller than k, i.e., $\frac{N\int_\tau^{t_u}(z)dz}{\tau+\int_\tau^{t_u} q(z)dz} \geq k$ in (5.14), which leads to (5.17). ■

We are also interested in the impact of the pseudonym change strategy on the transmission performance in the system. Assume the transmission of each node u_i follows the

Poisson distribution with parameter λ_i. Since each node only transmits data during the lifetime of each pseudonym, the throughput loss during the use of a pseudonym which changes at age z, $z \in [\tau, t_u]$, is calculated as

$$\gamma(z) = 1 - \frac{\tau}{z}. \tag{5.18}$$

Given the pseudonym change strategy $p(z)$, the distribution of the ratio of the throughput loss per node, denoted by $\pi(\gamma)$, is given by

$$\pi(\gamma) = \frac{\tau}{(1-\gamma)^2} p(\frac{\tau}{1-\gamma}), \gamma \in [0, 1 - \frac{\tau}{t_u}]. \tag{5.19}$$

Based on (5.19), the expected throughput loss per node, denote by $\mathbb{E}(\gamma)$, is given by

$$\mathbb{E}(\gamma) = 1 - \tau \int_{\tau}^{t_u} \frac{p(z)}{z} dz. \tag{5.20}$$

Based on (5.17) and (5.18), the following corollary is presented.

Corollary 6. *Given any pseudonym change strategy, the minimum expected pseudonym age is $\frac{N\tau}{N-k}$, and the throughput loss is negligible if $N \gg k$ in the network.*

Proof. Based on (5.17), when the lower bound of $\int_{\tau}^{t_u} q(z)dz$, i.e., $\int_{\tau}^{t_u} q(z)dz = \frac{\tau k}{N-k}$, is satisfied in the pseudonym change strategy, we obtain:

$$\int_{\tau}^{t_u} P(z)dz = t_u - \frac{N\tau}{N-k}, \tag{5.21}$$

where $P(z) = \int_{\tau}^{z} p(s)ds$. Therefore, the expected pseudonym age is given by

$$\bar{z} = \int_{\tau}^{t_u} zp(z)dz = \frac{N\tau}{N-n_c}. \tag{5.22}$$

Based on (5.18), (5.20), the expected throughput loss ratio caused by the pseudonym change in the system can be approximated to $(\bar{z} - \tau)/\bar{z} = k/(N-k)$. When N is large, the throughput loss becomes negligible. ■

As revealed in Corollary 6, to preserve the identity privacy, i.e., to achieve the k-anonymity, is at the cost of the network capability (or in other words, the number of transmissions under each pseudonym). Every node has to either change its pseudonym more frequently, or activate multiple pseudonyms at an instant, so as to increase the anonymity set size and protect its identity privacy [128].

5.3.3 Evaluation of Pseudonym Change Strategies

In this section, several simple pseudonym change strategies with different PDFs, $p(z)$, are proposed and optimized.

a. Uniform pseudonym change strategy

This is the strategy where, when any node u_i's pseudonym expires, it selects a random silent period within $[0, t_u - \tau]$ to suspend its transmission, and changes the pseudonym after the silent period in a distributed fashion, i.e.,

$$p(z) = \begin{cases} 0, & \text{if } z \in [0, \tau); \\ \frac{1}{t_u - \tau}, & \text{if } z \in [\tau, t_u]. \end{cases} \tag{5.23}$$

Therefore, $q(z)$ and $f(z|q(z))$ are given by

$$q(z) = \begin{cases} 1, & \text{if } z \in [0, \tau); \\ \frac{t_u - z}{t_u - \tau}, & \text{if } z \in [\tau, t_u], \end{cases} \tag{5.24}$$

and

$$f(z|q(z)) = \begin{cases} 0, & \text{if } z \in [0, \tau); \\ \frac{1}{t_u - z}, & \text{if } z \in [\tau, t_u]. \end{cases} \tag{5.25}$$

By substituting (5.23), (5.24) and (5.25) into (5.6), $m(z)$ can be obtained as:

$$m(z) = \begin{cases} \frac{1}{\frac{1}{2}(\tau + t_u)}, & \text{if } z \in [0, \tau); \\ \frac{t_u - z}{\frac{1}{2}(\tau + t_u)(t_u - \tau)}, & \text{if } z \in [\tau, t_u]. \end{cases} \tag{5.26}$$

From (5.11), the anonymity set size is $n_i \in \left[\frac{N(t_u-\tau)}{\tau+t_u}, \frac{3N(t_u-\tau)}{\tau+t_u}\right]$, and the expected anonymity set size is $\mathbb{E}(n_i) = 2N\frac{t_u-\tau}{\tau+t_u}$. The minimum anonymity set size is expected to be no smaller than k, to preserve privacy. Therefore, $t_u \geq \frac{(N+k)\tau}{N-k}$ needs to be guaranteed. Denote the lower bound of t_u as t_u^l, $t_u^l = \frac{(N+k)\tau}{N-k}$, based on (5.19), the throughput loss in the system is lower bounded by $\gamma = \frac{t_u-\tau(1+\ln\frac{N+k}{N-k})}{t_u-\tau} > \frac{k}{2N}$, which can be simplified to $\gamma = \frac{(N+k)-(N-k)\left(1+\ln(\frac{N+k}{N-k})\right)}{2k}$.

b. Exponential pseudonym change strategy

When a distributed exponential pseudonym change strategy is applied, the nodes change their pseudonyms at any age z that is exponentially distributed within $[\tau, +\infty)$:

$$p(z) = \begin{cases} 0, & \text{if } z \in [0, \tau); \\ we^{-w(z-\tau)}, & \text{if } z \in [\tau, \infty]. \end{cases} \tag{5.27}$$

Therefore, $q(z)$ and $f(z|q(z))$ are given by

$$q(z) = \begin{cases} 1, & \text{if } z \in [0, \tau); \\ e^{-w(z-\tau)}, & \text{if } z \in [\tau, t_u], \end{cases} \tag{5.28}$$

and

$$f(z|q(z)) = \begin{cases} 0, & \text{if } z \in [0, \tau); \\ w, & \text{if } z \in [\tau, \infty). \end{cases} \tag{5.29}$$

By substituting (5.27), (5.28) and (5.29) into (5.6), $m(z)$ can be obtained as

$$m(z) = \begin{cases} \frac{1}{\tau+\frac{1}{w}}, & \text{if } z \in [0, \tau); \\ \frac{1}{\tau+\frac{1}{w}}e^{-w(z-\tau)}, & \text{if } z \in [\tau, \infty). \end{cases} \tag{5.30}$$

The anonymity set size is $n_i \in \left[N\frac{\frac{1}{w}}{\tau+\frac{1}{w}}, N\right]$, and the expected anonymity set size of each node is $\mathbb{E}(n_i) = 2N\frac{\frac{1}{w}}{\tau+\frac{1}{w}}$. Accordingly, $1/w$ should be no less than $\frac{\tau k}{N-k}$ to guarantee that the anonymity set size of each node is no smaller than k. By setting $w = \frac{N-k}{\tau k}$, the

throughput loss of the system is calculated as $1-\Gamma(0,\frac{N-k}{k})\frac{N-k}{k}e^{\frac{N-k}{k}}$, where $\Gamma(0,\frac{N-k}{k})=\int_{\frac{N-2k}{k}}^{\infty}e^{-1}e^{-t}dt$, is the upper incomplete Gamma function.

c. Linear pseudonym change strategy

When a distributed linear pseudonym change strategy is applied, the nodes change the pseudonyms at any age with a linearly decreasing PDF within $[\tau, t_u]$, as given by

$$p(z)=\begin{cases}0, & \text{if } z\in[0,\tau);\\ \frac{2(t_u-z)}{(t_u-\tau)^2}, & \text{if } z\in[\tau,t_u].\end{cases} \tag{5.31}$$

Therefore, $q(z)$ and $f(z|q(z))$ are given by

$$q(z)=\begin{cases}1, & \text{if } z\in[0,\tau);\\ \frac{(t_u-z)^2}{(t_u-\tau)^2}, & \text{if } z\in[\tau,t_u],\end{cases} \tag{5.32}$$

and

$$f(z|q(z))=\begin{cases}0, & \text{if } z\in[0,\tau);\\ \frac{2}{t_u-\tau}, & \text{if } z\in[\tau,t_u].\end{cases} \tag{5.33}$$

By substituting (5.31), (5.32) and (5.33) into (5.6), $m(z)$ can be obtained as:

$$m(z)=\begin{cases}\frac{1}{\tau+\frac{t_u-\tau}{3}}, & \text{if } z\in[0,\tau)\\ \frac{(t_u-z)^2}{(\tau+\frac{t_u-\tau}{3})(t_u-\tau)^2}, & \text{if } z\in[\tau,t_u].\end{cases} \tag{5.34}$$

The anonymity set size is $n_i\in\left[\frac{N(t_u-\tau)}{2\tau+t_u},\frac{4N(t_u-\tau)}{2\tau+t_u}\right]$, and the expected anonymity set size is $\frac{2N(t_u-\tau)}{2\tau+t_u}$. Since $\frac{N(t_u-\tau)}{2\tau+t_u}\geq k$, i.e., $t_u\geq\frac{\tau(N+2k)}{N-k}$ is required to guarantee the anonymity set size of each node is no smaller than k. Denote the lower bound of t_u as t_u^l, $t_u^l=\frac{\tau(N+2k)}{N-k}$, the throughput loss of the system is lower bounded by $1-\frac{2\tau(t_u(\ln\frac{t_u}{\tau}-1)+\tau)}{(t_u-\tau)^2}$, which is simplified to $1-\frac{2(N-k)\left((N+2k)(\ln(\frac{N+2k}{N-k})-1)+(N-k)\right)}{(3k)^2}$.

d. Triangle pseudonym change strategy

When an isosceles-triangle distributed pseudonym change strategy is applied, the PDF of the pseudonym changes at age z can be written as

$$p(z) = \begin{cases} 0, & \text{if } z \in [0, \tau); \\ \frac{4(z-\tau)}{(t_u-\tau)^2}, & \text{if } z \in [\tau, \frac{t_u+\tau}{2}); \\ \frac{4(t_u-z)}{(t_u-\tau)^2}, & \text{if } z \in [\frac{t_u+\tau}{2}, t_u]. \end{cases} \tag{5.35}$$

Therefore, $q(z)$ and $f(z|q(z))$ are given by

$$q(z) = \begin{cases} 1, & \text{if } z \in [0, \tau); \\ \frac{(t_u-\tau)^2-2(z-\tau)^2}{(t_u-\tau)^2}, & \text{if } z \in [\tau, \frac{t_u+\tau}{2}); \\ \frac{2(t_u-z)^2}{(t_u-\tau)^2}, & \text{if } z \in [\frac{t_u+\tau}{2}, t_u], \end{cases} \tag{5.36}$$

and

$$f(z|q(z)) = \begin{cases} 0, & \text{if } z \in [0, \tau); \\ \frac{4(z-\tau)}{(t_u-\tau)^2-2(z-\tau)^2}, & \text{if } z \in [\tau, \frac{t_u+\tau}{2}); \\ \frac{2}{t_u-z}, & \text{if } z \in [\frac{t_u+\tau}{2}, t_u]. \end{cases} \tag{5.37}$$

Similarly, $m(z)$ can be derived as:

$$m(z) = \begin{cases} \frac{1}{\frac{1}{2}(t_{sp}+\tau)}, & \text{if } z \in [0, \tau) \\ \frac{2(t_u-\tau)^2-4(z-\tau)^2}{(t_u+\tau)\cdot(t_u-\tau)^2}, & \text{if } z \in [\tau, \frac{t_u+\tau}{2}) \\ \frac{4(t_u-z)^2}{(t_{sp}+\tau)\cdot(t_u-\tau)^2}, & \text{if } z \in [\frac{t_u+\tau}{2}, t_u] \end{cases} \tag{5.38}$$

The anonymity set size is $n_i \in \left[\frac{N(t_u-\tau)}{\tau+t_u}, \frac{3N(t_u-\tau)}{\tau+t_u}\right]$, and the expected anonymity set size is $\mathbb{E}(n_i) = \frac{2N(t_u-\tau)}{\tau+t_u}$. To guarantee $\frac{N(t_u-\tau)}{\tau+t_u} \geq k$, $t_u \geq \frac{\tau(N+k)}{N-k}$ is required. Denote the lower bound of t_u as t_u^l, $t_u^l = \frac{\tau(N+k)}{N-k}$, the throughput loss of the system is lower bounded by $1 + \frac{4\tau^2 \ln\left(\frac{\tau+t_u}{2\tau}\right)+4\tau t_u \ln\left(\frac{\tau+t_u}{2t_u}\right)}{(t_u-\tau)^2}$, which can be simplified to $1 + \frac{(N-k)^2 \ln(\frac{k}{N-k})+(N^2-k^2)\ln(\frac{N}{N+k})}{k^2}$.

The above analysis indicates that the lifetime of the pseudonyms has little impact on the throughput loss ratio in the four example pseudonym change strategies. Moreover, the expected pseudonym age achieved for the four pseudonym change strategies is equal to the result derived in (5.22). In other words, given the age threshold and k-anonymity requirement, the expected pseudonym age is the same for all the continuous pseudonym change strategies.

5.4 Validation with Simulations

5.4.1 Simulation Setup

A lightweight discrete-time simulator is developed based on the Matlab platform to assess different pseudonym change strategies, where a every node is preassigned the pseudonyms. The length of every time-slot is μ seconds, $\mu \ll \tau$. For initialization, every node joins the network by randomly selecting a time-slot and the first pseudonym, and starting to send data packets under the pseudonym. We are particularly interested in the impact of the initial distribution of $\mathbf{Z}^{(N)}(0)$ on the steady-state performance of identity privacy. Three initial distributions of $\mathbf{Z}^{(N)}(0)$ are considered, with the Poisson, uniform or normal distributions, as described in section 5.3. The lifetime of each pseudonym is τ. The nodes also set up their silent periods (randomly) within $[0, t_u - \tau]$ based on the different strategies. Data packets are transmitted during the lifetime of an active pseudonym.

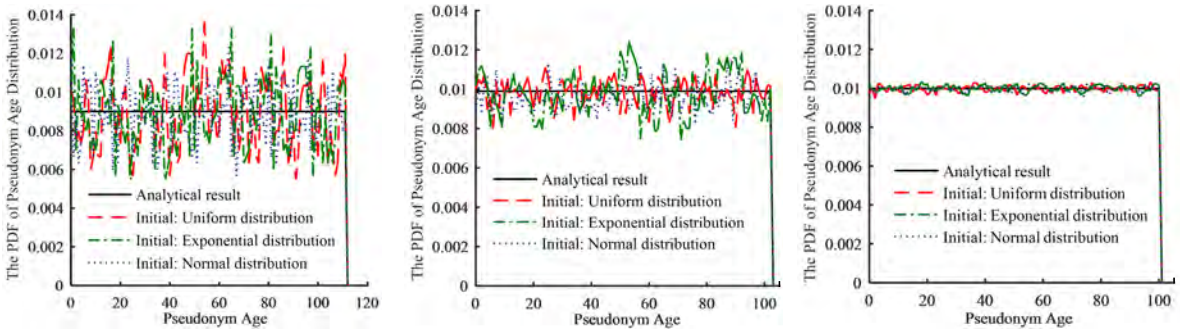
We assume that the packet transmissions of a node follow a Poisson distribution with the average transmit rate of λ (in packets per second), as described in section 5.3. A GPA attacker is set up to observe all the nodes' transmissions and their active pseudonyms in use. The attacker can initiate the linking attacks based on its observations on the transmissions and active pseudonyms in an attempt to establish the connections between the different pseudonyms of a node over time. The performance of the attacker is measured

by its uncertain level U_i [82]. In the simulations, we set $\lambda = 1$ packet per second, and $\tau = 100$ seconds. The other parameters, such as t_u, k and N are provided in the captions of individual simulation figures.

5.4.2 Convergence of the Pseudonym Age Distribution

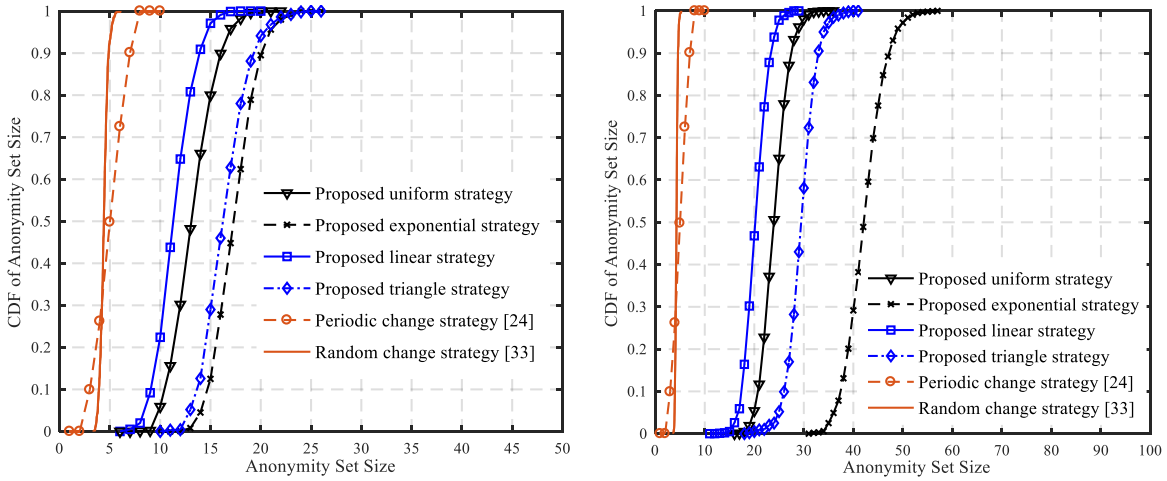
Fig. 5.3 demonstrates the age distribution in the uniform pseudonym change strategy with different initial pseudonym age distributions, and t_u is set to be the optimized value derived in the proposed model, i.e., $t_u = \frac{\tau(N+k)}{N-k}$. At least 50 pseudonyms are changed per node to plot each of the distributions. The simulation results validate the analytical model, and show that, as the number of nodes increases, the pseudonym age distribution converges increasingly fast and becomes close to the analytical result.

We also observe that the initial distribution of the pseudonym ages only affects the time for the pseudonym age distribution to stabilize and converge to the stationary distribution. The initial distribution has little impact on the final stationary distributions.



(a) $N = 50$ (b) $N = 500$ (c) $N = 5000$

Figure 5.3: The PDF of the pseudonym age distribution when applying the uniform pseudonym change strategy with $k = 10$ and different number of nodes in the network.



(a) $t_u = \frac{\tau N}{N-k}$ (b) $t_u = \frac{\tau(N+2k)}{N-k}$

Figure 5.4: The CDF of the anonymity set size with $N = 500$, $k = 10$, and $t_u = \frac{\tau N}{N-k}$ and $t_u = \frac{\tau(N+2k)}{N-k}$, optimized for the exponential strategy and linear strategy, respectively.

5.4.3 k-Anonymity versus Throughput Loss

Figs. 5.4a and 5.4b plot the CDF of the anonymity set size obtained under different pseudonym change strategies, where $k = 10$ and $N = 500$. We take $t_u = \frac{\tau N}{N-k}$, optimized for the exponential strategy, in Fig. 5.4a; and $t_u = \frac{\tau(N+2k)}{N-k}$, optimized for the linear strategy, in Fig. 5.4b. For comparison purpose, we also simulate a periodic pseudonym change approach developed in [61], where each node changes its pseudonym immediately after the expiration of the previous one; and a random pseudonym change approach developed in [81], where each node changes its pseudonym after a uniformly distributed random period within $[0, t_u]$.

Fig. 5.4a shows that only the proposed exponential strategy achieves the k-anonymity. Not all the nodes can achieve the k-anonymity in the other five strategies. Fig. 5.4b demonstrates that all the four proposed pseudonym change strategies can achieve the k-anonymity, with the exponential strategy achieving the largest anonymity set size. The reason is that these strategies require different silent periods for all the nodes to achieve

the k-anonymity. We also observe that the expected achievable anonymity set sizes of the four proposed strategies are almost twice of the k-anonymity requirement, and the largest achievable anonymity set size is nearly three times of the requirement. In contrast, the anonymity set size is much smaller than k in the periodic pseudonym change strategy [61] and the random change strategy [81]. The reason is that, when each node changes pseudonyms immediately with no silent period [61, 81], the node may not be able to gather enough peers to achieve the k-anonymity.

In Fig. 5.5a, we evaluate the average pseudonym age, the number of pseudonyms needed for a node to transmit 10^4 packets, and the throughput loss resulting from the pseudonym changes, as the number of nodes N increases. Different pseudonym change strategies and network scales are taken into account. The pseudonym change threshold, t_u, in the periodic and random pseudonym change strategy [61, 81] are set up as τ initially, and updated iteratively until the anonymity set size of each node always guarantees the k-anonymity. In the proposed four strategies, both the pseudonym age and throughput loss decreases, as the number of nodes increases. It is noted that the expected pseudonym ages of the four strategies are the same, and thus their throughput losses are nearly the same. When the number of nodes in the network is large, the throughput loss is negligible and the performances of the different pseudonym change strategies converge. In other words, a sufficiently large network can inherently preserve the identity privacy to the nodes without need for coordination or negotiation, and achieve the k-anonymity at a negligible loss of throughput. We see that, each node needs to change the pseudonyms more frequently to achieve the required k-anonymity under the two existing strategies [61, 81]. In the periodic [61] and random [81] pseudonym change strategies, without silent period, the pseudonym changes do not affect the throughput, but need a much larger number of pseudonyms, especially when the number of nodes is small or the k-anonymity requirement is stringent. This would increase the interactions between the TTP and the nodes with

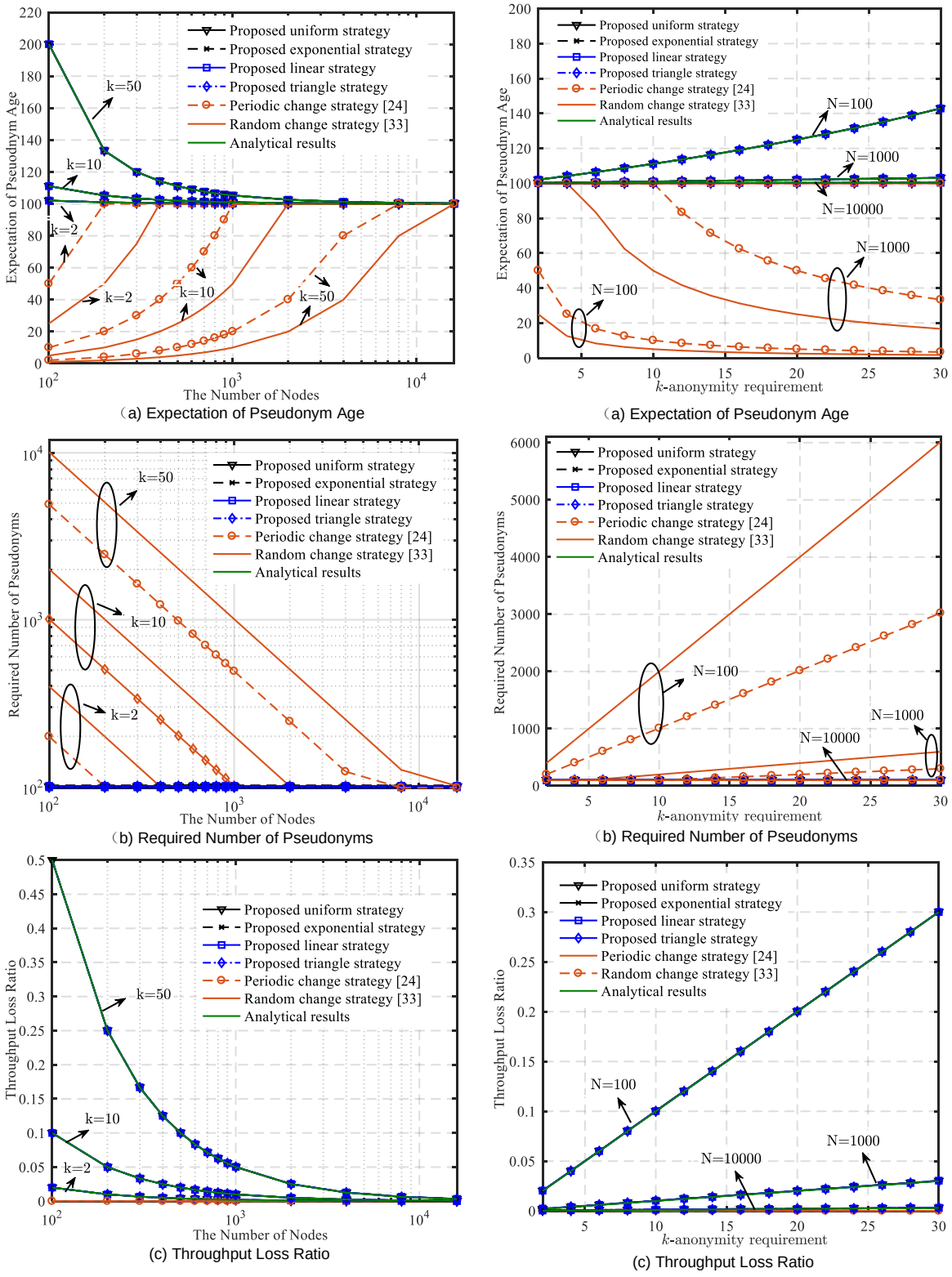


Figure 5.5: The pseudonym age and throughput loss of different pseudonym change strategies under different number of nodes and k-anonymity requirements in the network.

growing possibility of side information leakage in the existing approaches [61, 81].

In Fig. 5.5b, we evaluate the average pseudonym age, the number of pseudonyms needed, and the throughput loss, as the k-anonymity requirement becomes increasingly stringent. We see that the time required to gather enough nodes for the k-anonymity increases under each of the four proposed pseudonym change strategies, and the throughput loss also becomes increasingly severe, with the increasingly stringent k-anonymity requirement. For the two existing strategies [61, 81], the pseudonym age decreases with the increasingly stringent k-anonymity requirement. When the number of nodes is small, the difference of the pseudonym age is large among the six strategies. The difference diminishes as the number of nodes rises. This observation indicates that the design of the pseudonym change strategy can have strong impact on the performance of the network, especially when the network is small.

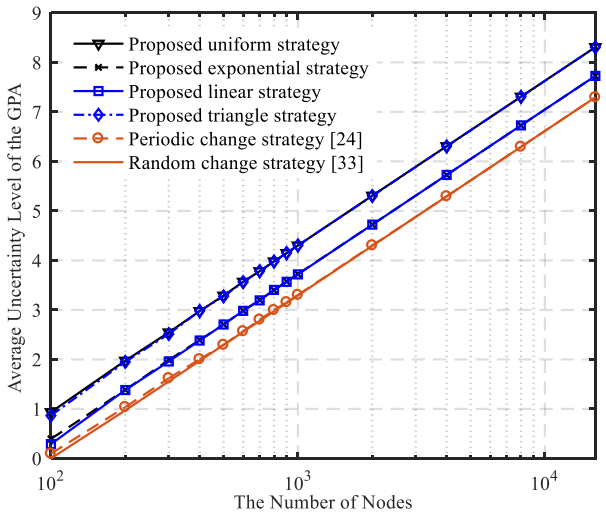


Figure 5.6: The simulated average uncertainty level of the GPA with the pseudonym lifetime $\tau = 100$ seconds and the silent period $\tau - t_u = 2$ seconds in the proposed pseudonym change strategies.

The GPA's uncertainty level is also simulated and plotted under different pseudonym change strategies. Fig. 5.6 plots the uncertainty level with the increasing number of

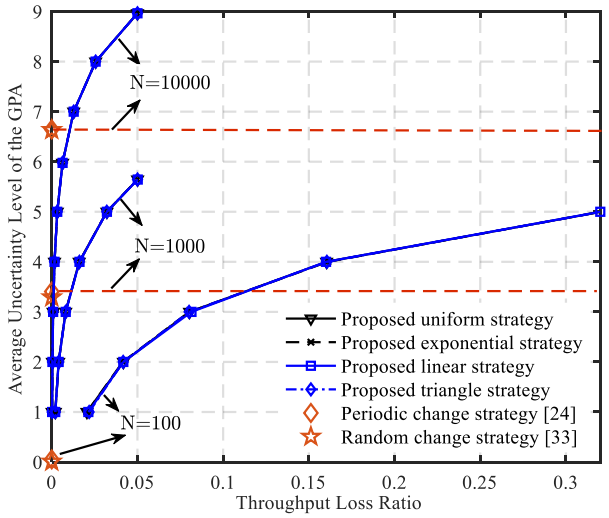


Figure 5.7: The simulated average uncertainty level of the GPA with the pseudonym lifetime $\tau = 100$ seconds and different throughput losses. The uncertain levels of the two existing approaches are provided as the reference lines.

nodes. In the periodic pseudonym change strategy [61], every node changes its pseudonym periodically every t_u seconds. In the random strategy, every node updates its pseudonym randomly within $[0, t_u]$ seconds [81]. The result shows that the GPA's uncertainty level of the existing two strategies is nearly the same, and much smaller than it is in the proposed four pseudonym change strategies. Fig. 5.7 plots the GPA's uncertainty level of the proposed pseudonym change strategies, as the throughput loss of the strategies increases. We also plot the uncertainty level under the two existing strategies [61, 81] for reference purpose, since the two existing strategies do not incur any throughput loss at the cost of privacy. It is shown that when the number of nodes is large, even a very short silent period is able to preserve the identity privacy of the nodes with nearly negligible cost of throughput in the proposed strategies. When the number of nodes is small, the ability of the proposed pseudonym change strategies to trade throughput for privacy becomes important. It allows the proposed strategies to substantially outperform the existing methods [61, 81] in terms of privacy.

5.5 Conclusion

In this chapter,fully uncoordinated pseudonym changing approach is proposed to preserve the privacy in distributed networks. The nodes in the network can change their pseudonym without any interaction or coordination from the infrastructures, by suspending their transmissions and then randomly change pseudonyms to confuse potential adversaries. The distributed network provides an inherent privacy for the nodes, when the network is large. A new analytical mode has been developed to analyze the privacy of fully distributed changes of pseudonyms, where the pseudonym aging and changing is modeled and formulated mathematically. Given the network size, pseudonym lifetime, and pseudonym change strategies, the silent period is analytically configured, under which the k-anonymity is achieved without any cooperation between the nodes, while minimizing the throughput loss of the network. Validated by simulations, the proposed analytical model confirms that, when N is large, the k-anonymity inherently provided by the network is able to preserve the privacy for nodes and incur negligible throughput loss. The throughput loss resulting from the pseudonym changes can diminish with the growth of N.

In the future, the proposed analytical model and algorithms applied for heterogeneous works where the nodes can have different requirements and transmission patterns will be developed. Another promising research topic is to integrate the proposed distributed pseudonym change strategies for practical applications.

Chapter 6

Thesis Conclusion and Future Work

Based on the above chapters, we comprehensively summarize the research contribution of the thesis in this chapter, and point out the future research directions.

6.1 Thesis Conclusion

This paper combines the hot field and the challenges in the development of the VANETs, optimizes the performance of the MAC protocols to improve the transmission reliability of the safety related messages, and studies the pseudonym schemes, so as to ensure the reliable, secure and privacy preserving transmission in the VANETs. The main research innovations and contributions of this paper can be summarized as follows.

1. The MAC Layer Context Linking Attack Resistant Pseudonym Scheme

   The thesis proposes an integrated pseudonym scheme, namely the MARP scheme, to guarantee not only the privacy preserving transmission of the safety messages, but also the efficiency and reliability of the transmission in the VANETs. The specific contributions of the work, corresponding to chapter 3, are summarized. Firstly, we analyze the dilemma that has occurred between the MAC layer operation and

the pseudonym schemes in the existing work, which designed the protocols to ensure the transmission QoS requirements, and the location privacy requirement separately. Based on the analysis of the probabilities to utilize the MAC layer context information under both the CSMA and TDMA protocols to link the new and old pseudonyms of the nodes by the adversary, we identify the new MAC layer context linking attack, and point that that the MAC layer protocol and the pseudonym scheme should be designed in a consistent approach. Then, we propose the MAC layer attack resistant pseudonym scheme, MARP scheme. In the MARP scheme, the nodes reserve the dedicated time slots to transmit the safety messages, and based on the CSMA to exchange the security related messages, so that the transmissions of different types of messages have little interference with each other. Based on the current pseudonym age, the nodes make the cooperation decision to change the pseudonyms and transmission slots simultaneously with their neighbors. By this way, the MARP scheme ensures the reliable and privacy preserving transmission for the VANETs. Finally, the thesis proposes an analytical model to evaluate the privacy performance of the nodes, in terms of the pseudonym age and the cooperation probability of the vehicles, and the time-to-confusion of the adversary, with taking the PDR and transmission latency of the messages into consideration. The simulation results corroborate the accuracy of the analytical model, and the superiority of the MARP scheme.

2. A MAC Layer Aware Pseudonym Scheme in the Software Defined Internet of Vehicles

   In this paper, we have presented a SDN based three-layer system architecture for the IoV. In SDIV, the RSUs are connected together according to the flow tables to manage the vehicles by the controller in the central cloud layer. Facilitated by the architecture, a MAC layer aware pseudonym scheme is proposed to allocate both

the channel resources and pseudonyms for the vehicles. The RSU clouds coordinates vehicles to adaptively change pseudonyms by determining whether the vehicles can achieve location privacy gains through changing pseudonyms, and access the wireless channels in a contention-free manner. Security analysis and extensive simulations are conducted to show that the scheme can significantly outperform the compared schemes in terms of the location privacy preservation, and the transmission efficiency and reliability for the IoV.

3. Uncoordinated Pseudonym Changes for Privacy Preserving in Distributed Networks

   Pseudonyms have been adopted to preserve identity privacy of nodes in distributed networks. Frequent and unlinkable changes of pseudonyms need to be enabled by having at least k nodes change together to confuse potential eavesdroppers. Existing approaches either depend on the coordination from central controllers, or involve interactive signaling between the nodes. This can potentially compromise the identity privacy of the nodes. This thesis proposes a fully uncoordinated approach to change pseudonyms in distributed networks, where each node uses a pseudonym until its expiration and then changes after a random delay. The approach can not only guarantees the privacy preserving of the nodes but also brings little degradation on the throughput of the network. A typical application scenario is to prevent adversaries from tracking vehicles and personnel in the IoV. Continuously growing data exchange between vehicles/personnel and the Internet has been increasingly facilitating eavesdropping and linking attacks on the vehicles/personnel. We develop a new model to analyze the time-varying population of changing pseudonyms. Critical conditions are analytically established, under which individual nodes can independently change their pseudonyms while their identity privacy is preserved. The conditions are validated by illustrative examples. Corroborated by simulations, the accuracy of the analytical model improves, as the number of nodes increases. The

analysis confirms that, the k-anonymity can be achieved at a negligible throughput loss in the case of large networks.

6.2 Future Research Directions

Limited by the research time, the author's research ability and the thesis' word limits, we only present the most important parts of our research work. There are still some related problems need to be solved. Considering to improve the research work presented in the thesis, the author's future research efforts will be put on the following aspects.

1. Performance Evaluation of the Pseudonym Schemes under Different Attack Models in the VANETs

   The quality of privacy of the nodes is highly dependent on the capability of the adversaries in the VANETs. Adversaries can launch different types of attacks, such as statistical attacks [135], intersection attacks [136], driving path prediction attacks of the vehicles, etc., to achieve the continuous tracking of the nodes in the VANETs. In the thesis, Chapters 3 and 4 have focused on the design of the pseudonym schemes by mainly considering the MAC layer context linking attack proposed in the thesis. We assume that as long as the privacy level of each vehicle node is above 0, it can temporarily confuse the adversary. The proposed pseudonym schemes ignore the impact of other types of attacks when evaluating the privacy level that can be achieved by the vehicles. In the future, the author will consider the impact of different types of attacks, and the capabilities of the adversaries to compromise the location privacy of the nodes, to design the more robust and scalable pseudonym schemes in the VANETs.

2. The Design and Optimization of the Pseudonym Management System in the SDIV.

In Chapter 4, we propose the SDN based IoV architecture (SDIV), under which the MAP scheme is proposed to exploit the local RSU cloud to coordinate the channel accessing and pseudonym changing of the vehicles, however, the allocation of the pseudonym resources between different RSUs is not considered. In the future,we consider to improve the pseudonym management efficiency. The central cloud will promptly schedule the pseudonyms in the distributed pseudonym pools under different RSUs in the local cloud with diverse traffic densities and privacy preferences. Thus, the pseudonym utilization can be further improved, and the location privacy of the vehicles can also be enhanced by ensuring them the plenty of pseudonyms.

3. The Integration of the Distributed Pseudonym Change Approach for Practical Applications

   In Chapter 5, we propose the uncoordinated pseudonym change approach for the homogeneous distributed networks, where all the nodes are supposed to have the same privacy requirements and transmission patterns. In the future, we will extend the proposed uncoordinated pseudonym change approach and the analytical model to heterogeneous networks where the nodes have different privacy requirements, transmission patterns and connectivity typologies. We will also investigate the feasibility of integrating the proposed distributed pseudonym change strategies for practical applications and carry out more simulation by using the network simulators, such as the NS2/3, OPNET, etc.

Appendix A

Pseudocode for the Algorithm

A.1 Algorithms Proposed in the MARP Scheme

Algorithm 1 Time Slot Reservation

for each vehicle V_i**:**
during each slot ts_k **in the DT**:
sets s_k as "00"
if received exactly one message **then**
    sets s_k as "01"
else if received from a one-hop neighbour saying s_k is "01" **then**
    sets s_k as "10"
else if received more than one message from different vehicles or a collision **then**
    sets s_k as "11"
end if
updates the FI and randomly selects a new transmission slot ts_i with status "00" during the next frame
broadcasts the BSM in ts_i
waits for $N_{ts} - 1$ slots following ts_i
if all the neighbours update s_i as "01" **then**
    considers the reservation is successful
else
    repeats the above process until it reserves a transmission slot successfully
end if

Algorithm 2 Time Slot Shuffle

for each vehicle V_i**:**
updates the slot status list after each time slot,
if (the s_i is updated by all the neighbours as "01") & (does not need to construct a mix-zone) **then**
    broadcasts the BSMs periodically in the current time slot
else
    releases the current slot and does Algorithm 1 until it reserves a new transmission slot in the DT
end if

Algorithm 3 Mix-zone Construction

for each vehicle V_i**:**
updates the age of the current pseudonym
if the pseudonym age $\geq \tau$ **then**
  transmits the BSM with null safety information and the $ChangeReq$ indicator
  if $|C_{v_i}| \geq 1$ **then**
    randomly chooses a silent period within $[ts_{min}, ts_{max}]$
    releases the current transmission slot
    remains silent until the silent period expires
    uses a new pseudonym to implement Algorithm 1 until it reserves a new time slot
  end if
end if

A.2 Algorithms Proposed in the MAP Scheme

Algorithm 4 The Channel Coordination Procedure

for each RSU:
update the slot status and the No. of the SCH
when switching to the CCH:
in ts_j**:**
if $(N_\gamma(U_i) \geq AS_{th}$ && $G_i \geq G_{th})$ || $(U_i(t) == 0)$,$\forall i \in \nu(j)$
then $MC_i = 1$
$RSU_j \rightarrow v_i, \forall i \in \nu(j)$:
$coord\_msg = (Cert_j||SI||MC||GID_j||T_m||payload||Sig(CM))$
end if
in other time slots, $\forall ts_i \in \mathbb{N}_{ts}, ts_i \neq ts_j$**:**
$SI_i =$ "00"
if receives from one vehicle correctly
then $SI_i =$ "01"
if pre-warning a collision
then $SI_i =$ "10"
end if
end if
else if detects a transmission collision
then $SI_i =$ "11"
end if
when switching to the specific SCH:
broadcasts CRL, provides services for the surrouding vehicles

Algorithm 5 The Channel Accessing Procedure

for each v_i**:**
when switching to the CCH:
if $MC_i == 1$
then goes to Algorithm 8
else then
receives and verifies the $coord\_msg$ from RSU_j
end if
if it has occupied a time slot ts_i
if $SI_i ==$ "01" **then**
broadcasts the safety message in ts_i
else if $SI_i ==$ "10" **then**
broadcasts the safety message in ts_i
if there exists available time slot **then**
randomly selects a new available slot to reserve,
when reserves the new slot successfully, releases ts_i
end if
else if $SI_i ==$ "11" **then**
releases ts_i immediately
end if
else if it has not reserved a time slot**then**
selects an available time slot, and transmits the sasfety message in the slot
end if
when switching to the specific service channel:
exchanges the security related messages with the RSU_j

Algorithm 6 The Pseudonym Request Procedure

for each v_i**:**
if needs to request new pseudonyms
then $v_i \rightarrow RSU_m$:
$req = RSU_j||E_{SK_{ik}}(PseudoReq||PK_i||Cert_i||Cert_{ik}$
$||timestamp)$
if RSU_m is not the registration RSU of V_i
then receives the data record of v_i from RSU_j
$record = \{PK_i||PID_i, SK_{PID_i}, Cert_{PID_i}||Cert_i\}$
end if
if v_i's identity is verified
then $RSU_m \rightarrow V_i$:
$rep = E_{PK_{PIDik}}(\{PID_{ik}, SK_{PID_{ik}}, Cert_{PID_{ik}}\}_{k=1}^{w})$
end if
if v_i receives, and verifies rep
then $V_i \rightarrow RSU_m$:
$ack = RSU_m||E_{PK_m}(Cert_{PID_{ik}}||timestamp)$
end if
RSU_m updates the data record of v_i, and transmits to RSU_j

Algorithm 7 The Pseudonym Change Coordination Algorithm

for each RSU_j**:**
calculates the current $U_i(t)$ for each $v_i, i \in nu(j)$
does the descending sort for $U_i(t), \forall v_i \in nu(j)$
while not reaching the least $U_i(t)$
if $N_\gamma(U_i(t)) \geq AS_{th}^i$ && $G_i \geq G_{th}, \exists V_k \in \gamma(U_i(t))$
determines that all the vehicles in $\gamma(U_i(t))$ to change pseudonyms simultaneously
return
else i++
goes to the Algorithm 8

Algorithm 8 The Pseudonym Change Procedure

for each v_i**:**

$v_i \rightarrow RSU_j$

$join\_ack = RSU_j||E_{GID}(Cert_{PID_{ik}}||PID_{ik}$
$||payload||timestamp)$

$RSU_j \rightarrow v_i$:

$res = E_{PK_{ik}}(GID||SK_{G_{ji}}||timestamp||Cert_{G_{ji}}||Cert_j)$

$RSU_j \rightarrow v_i$:

$slot\_alloc = E_{PK_{GID_i}}(ts_i^{'}||Cert_j||timestamp)$

receives new slot allocation information, $v_i \rightarrow RSU_j$:

$success = RSU_j||E_{PK_j}(ack||SafetyData$
$||timestamp||Cert_{G_{ji}})$

else if $U_i(t) == 0$

then v_i releases ts_i, selects a silent period from $[ts_{min}, ts_{max}]$ to keep silent after the silent period expires

end if

V_i: $PID_{i,k} \rightarrow PID_{i,k+1}, k = 1, 2, \ldots, w$, goes back to Algorithm 5

Appendix B

List of Acronyms

BSM	Basic Safety Message
CA	Certificate Authority
CCH	Control Channel
CCHI	Control Channel Interval
CDF	Cumulative Distribution Function
CGA	Cryptographically Generated Address
CSMA	Carrier Sense Multiple Access
CRL	Certificate Revocation List
CITS	Collaborative Intelligent Transportation System
DIFS	Distributed Inter-frame Spacing
DSRC	Dedicated Short Range Communication
ECDSA	Elliptic Curve Digital Signature Algorithm
ELP	Electronic License Plate
ETSI	European Telecommunications Standards Institute
FCC	Federal Communication Commission
FDMA	Frequency Division Multiple Access

GPA	Global Passive Adversary
GPS	Global Positioning System
HSM	Hardware Security Module
IEEE	Institute of Electrical and Electronic Engineers
IoT	Internet of Things
IoV	Internet of Vehicles
MAC	Medium Access Control
MANET	Mobile AdHoc Networks
OBU	On-Board Unit
OFS	OpenFlow Switch
PCA	Pseudonym Certificate Authority
PDE	Partial Differential Equations
PDF	Probability Density Function
PKI	Public Key Infrastructure
SCH	Service Channel
SCHI	Service Channel Interval
SDMA	Space Division Multiple Access
SDN	Software Defined Network
SI	Synchronized Interval
RSU	Roadside Unit
TA	Trusted Authorities
TPD	Tamper Proof Device
TTP	Third Trusted Parties
TDMA	Time Division Multiple Access
VANET	Vehicular Ad Hoc Network
WAVE	Wireless Access in the Vehicular Environment

WSN	Wireless Sensor Network

References

[1] F. Yang, S. Wang, J. Li, Z. Liu, and Q. Sun, "An overview of internet of vehicles," *China Communications*, vol. 11, no. 10, pp. 1–15, 2014.

[2] O. Kaiwartya, A. H. Abdullah, Y. Cao, A. Altameem, M. Prasad, C. Lin, and X. Li, "Internet of vehicles: Motivation, layered architecture, network model, challenges, and future aspects," *IEEE Access*, vol. 4, pp. 5356–5373, 2016.

[3] J. Contreras-Castillo, S. Zeadally, and J. A. Guerrero-Ibanez, "Internet of vehicles: Architecture, protocols, and security," *IEEE Internet of Things Journal*, vol. 5, no. 5, pp. 3701–3709, 2018-10.

[4] Y. L. Morgan, "Notes on dsrc & wave standards suite: Its architecture, design, and characteristics," *IEEE Communications Surveys Tutorials*, vol. 12, no. 4, pp. 504–518, 2010.

[5] H. Hartenstein and L. P. Laberteaux, "A tutorial survey on vehicular ad hoc networks," *IEEE Communications Magazine*, vol. 46, no. 6, pp. 164–171, 2008.

[6] E. Schoch, F. Kargl, M. Weber, and T. Leinmuller, "Communication patterns in vanets," *IEEE Communications Magazine*, vol. 46, no. 11, pp. 119–125, 2008.

[7] R. Stanica, E. Chaput, and A. Beylot, "Properties of the mac layer in safety vehicular ad hoc networks," *IEEE Communications Magazine*, vol. 50, no. 5, pp. 192–200, 2012.

[8] X. Lin, R. Lu, C. Zhang, H. Zhu, P. Ho, and X. Shen, "Security in vehicular ad hoc networks," *IEEE communications magazine*, vol. 46, no. 4, pp. 88 – 95, 2008.

[9] R. S. Raw, M. Kumar, and N. Singh, "Security challenges, issues and their solutions for vanet," *International Journal of Network Security & Its Applications*, vol. 5, no. 5, 2013.

[10] R. G. Engoulou, M. Bellaïche, S. Pierre, and A. Quintero, "Vanet security surveys," *Computer Communications*, vol. 44, pp. 1–13, 2014.

[11] *IEEE Guide for Wireless Access in Vehicular Environments (WAVE) - Architecture*, Std. IEEE 1609.0, 2014.

[12] *Trial-Use Srandard for Wireless Access in Vehicular Environments (WAVE)-Resource Maneger*, Std. IEEE 1609.1, 2010.

[13] *IEEE Approved Draft Standard for Wireless Access in Vehicular Environments–Security Services for Applications and Management Messages*, IEEE Std. P1609.2a/D8, 2017.

[14] *Trial-Use Srandard for Wireless Access in Vehicular Environments (WAVE)-Networking Services*, Std. IEEE 1609.3, 2010.

[15] *IEEE Standard for Wireless Access in Vehicular Environments (WAVE) – Multi-Channel Operation*, IEEE Std. IEEE 1609.4, 2016.

[16] E. T. C. I. T. Syst., "Intelligent transport system (its); framework for public mobile networks in cooperative its (c-its)," Tech. Rep., 2012.

[17] K. Sjoberg, P. Andres, T. Buburuzan, and A. Brakemeier, "Cooperative intelligent transport systems in europe: current deployment status and outlook," *IEEE Vehicular Technology Magazine*, vol. 12, no. 2, pp. 89–97, 2017.

[18] *Intelligent Transport Systems (ITS); Communications Architecture*, Std. ETSI EN 302-665 V1.1.1, 2010.

[19] *Intelligent Transport Systems; Access layer specification for intelligent transport systems operating in the 5 GHz frequency band*, Std. ETSI EN 302 663 V1.2.1, 2013.

[20] *GeoNetworking, Part 4: Geographical Addressing and Forwarding for Point-to-Point and Point-to-Multipoint Communications; Subpart 1: Media-Independent Functionality*, Std. ETSI EN Standard 302 636-4-1 V1.2.1, 2014.

[21] *GeoNetworking, Part 5: Transport Protocols; Sub-part 1: Basic Transport Protocol*, Std. ETSI EN Standard 302 636-5-1 V1.2.1, 2014.

[22] *ETSI TS 102 867 v1. 1.1-Intelligent Transport Systems (ITS); Security; ITS communications security architecture and security management*, Std. ETSI TS 102 940, 2012.

[23] B. Lonc and P. Cincilla, "Cooperative its security framework: Standards and implementations progress in europe," in *World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2016 IEEE 17th International Symposium on A*, 2016, pp. 1–6.

[24] *Feasibility Study on LTE-based V2X Services (Release 14).v14.0.0*, Std. 3GPP TR 36.885, 2016-06.

[25] R. Mishra, A. Singh, and R. Kumar, "Vanet security: Issues, challenges and solutions," in *2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)*, 2016, pp. 1050–1055.

[26] H. Hasrouny, A. E. Samhat, C. Bassil, and A. Laouiti, "Vanet security challenges and solutions: A survey," *Vehicular Communications*, vol. 7, pp. 7–20, 2017.

[27] A. Pathan, *Security of Self-Organizing Networks: MANET, WSN, WMN, VANET*. Boca Raton, USA: CRC Press, 2011.

[28] H. Jayasree and A. Damodaram, "Anonymity and accountability in web based transactions," *Advanced Computing*, vol. 3, no. 2, p. 193, 2012.

[29] L. Zhang, Z. Liu, R. Zou, J. Guo, and Y. Liu, "A scalable csma and self-organizing tdma mac for ieee 802.11 p/1609. x in vanets," *Wireless Personal Communications*, vol. 74, no. 4, pp. 1197–1212, 2014.

[30] D. N. M. Dang, H. N. Dang, V. Nguyen, Z. Htike, and C. S. Hong, "Her-mac: A hybrid efficient and reliable mac for vehicular ad hoc networks," in *2014 IEEE 28th International Conference on Advanced Information Networking and Applications*, 2014, pp. 186–193.

[31] V. Nguyen, T. Z. Oo, P. Chuan, and C. S. Hong, "An efficient time slot acquisition on the hybrid tdma/csma multichannel mac in vanets," *IEEE Communications Letters*, vol. 20, no. 5, pp. 970–973, 2016.

[32] V. Nguyen, T. Z. Oo, N. H. Tran, and C. S. Hong, "An efficient and fast broadcast frame adjustment algorithm in vanet," *IEEE Communications Letters*, vol. 21, no. 7, pp. 1589–1592, 2017.

[33] C. Song, G. Tan, C. Yu, N. Ding, and F. Zhang., "Apdm: An adaptive multi-priority distributed multichannel mac protocol for vehicular ad hoc networks in unsaturated conditions," *Computer Communications*, vol. 104, pp. 119–133, 2017.

[34] D. Lee, S. H. Ahmed, D. Kim, J. Copeland, and Y. Chang, "An efficient sch utilization scheme for ieee 1609.4 multi-channel environments in vanets," in *2016 IEEE International Conference on Communications (ICC)*, 2016, pp. 1–6.

[35] Z. Liu, T. Liang, J. Guo, and L. Zhang, "Priority-based access for dsrc and 802.11p vehicular safety communication," in *2012 International Conference on Connected Vehicles and Expo (ICCVE)*, 2012, pp. 103–107.

[36] Z. Liu, R. Zou, H. Zhang, and L. Zhang, "An adaptive and reliable integrated mac mechanism for vanets," in *2013 16th International Symposium on Wireless Personal Multimedia Communications (WPMC)*, 2013, pp. 1–5.

[37] P. Zhou, Y. Liu, J. Wang, W. Deng, and H. Oh, "Performance analysis of prioritized broadcast service in wave/ieee 802.11p," *Computer Networks*, vol. 107, pp. 233–245, 2016.

[38] M. S. Almalag, S. Olariu, and M. C. Weigle, "Tdma cluster-based mac for vanets (tc-mac)," in *2012 IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, 2012, pp. 1–6.

[39] R. Zhang, X. Cheng, L. Yang, X. Shen, and B. Jiao, "A novel centralized tdma-based scheduling protocol for vehicular networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 1, pp. 411–416, 2015.

[40] Y. Lai, P. Lin, W. Liao, and C. Chen, "A region-based clustering mechanism for channel access in vehicular ad hoc networks," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 1, pp. 83–93, 2011.

[41] R. S. Tomar, S. Verma, and G. S. Tomar, "Cluster based rsu centric channel access for vanets," *Transactions on Computational Science XVII*, pp. 150–171, 2013.

[42] R. Stanica, E. Chaput, and A. Beylot, "Local density estimation for contention window adaptation in vehicular networks," in *2011 IEEE 22nd International Symposium on Personal, Indoor and Mobile Radio Communications*, 2011, pp. 730–734.

[43] D. B. Rawat, D. C. Popescu, G. Yan, and S. Olariu, "Enhancing vanet performance by joint adaptation of transmission power and contention window size," *IEEE Transactions on Parallel and Distributed Systems*, vol. 22, no. 9, pp. 1528–1535, 2011.

[44] F. Borgonovo, A. Capone, M. Cesana, and L. Fratta, "Adhoc mac: New mac architecture for ad hoc networks providing efficient and reliable point-to-point and broadcast services," *Wireless Networks*, vol. 10, no. 4, pp. 359–366, 2004.

[45] L. Miao, F. Ren, C. Lin, and A. Luo, "A-adhoc: An adaptive real-time distributed mac protocol for vehicular ad hoc networks," in *2009 Fourth International Conference on Communications and Networking in China*, 2009, pp. 1–6.

[46] H. A. Omar, W. Zhuang, and L. Li, "Vemac: A tdma-based mac protocol for reliable broadcast in vanets," *IEEE Transactions on Mobile Computing*, vol. 12, no. 9, pp. 1724–1736, 2013.

[47] K. Bilstrup, E. Uhlemann, E. G. Strom, and U. Bilstrup, "Evaluation of the ieee 802.11p mac method for vehicle-to-vehicle communication," in *2008 IEEE 68th Vehicular Technology Conference*, 2008, pp. 1–5.

[48] X. Jiang and D. H. Du, "Ptmac: A prediction-based tdma mac protocol for reducing packet collisions in vanet," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 11, pp. 9209–9223, 2016.

[49] G. M. Abdalla, M. A. Abu-Rgheff, and S. Senouci, "Space-orthogonal frequency-time medium access control (soft mac) for vanet," in *2009 Global Information Infrastructure Symposium*, 2009, pp. 1–8.

[50] R. S. Tomar and S. Verma, "Rsu centric channel allocation in vehicular ad-hoc networks," in *Wireless Communication and Sensor Networks (WCSN), 2010 Sixth International Conference on*, 2010, pp. 1–6.

[51] W. Guo, L. Huang, L. Chen, H. Xu, and J. Xie, "An adaptive collision-free mac protocol based on tdma for inter-vehicular communication," in *Wireless Communications & Signal Processing (WCSP), 2012 International Conference on*, 2012, pp. 1–6.

[52] S. V. Bana and P. Varaiya, "Space division multiple access (sdma) for robust ad hoc vehicle communication networks," in *ITSC 2001. 2001 IEEE Intelligent Transportation Systems. Proceedings (Cat. No.01TH8585)*, 2001, pp. 962–967.

[53] H. Nakata, T. Inoue, M. Itami, and K. Itoh, "A study of inter vehicle communication scheme allocating pn codes to the location on the road," in *Proceedings of the 2003 IEEE International Conference on Intelligent Transportation Systems*, vol. 2, 2003, pp. 1527–1532.

[54] V. Jayaraj, H. C., and S. R.G., *A survey on hybrid MAC protocols for vehicular ad-hoc networks*, vol. 6, pp. 29 – 36, 2016.

[55] J. Petit, F. Schaub, M. Feiri, and F. Kargl, "Pseudonym schemes in vehicular networks: A survey," *IEEE Communications Surveys and Tutorials*, vol. 17, no. 1, pp. 228–255, 2015.

[56] S. Gisdakis, T. Giannetsos, and P. Papadimitratos, "Security, privacy, and incentive provision for mobile crowd sensing systems," *IEEE Internet of Things Journal*, vol. 3, no. 5, pp. 839–853, 2016.

[57] G. Calandriello, P. Papadimitratos, J.-P. Hubaux, and A. Lioy, "Efficient and robust pseudonymous authentication in vanet," in *Proceedings of the fourth ACM international workshop on Vehicular ad hoc networks*, 2007, pp. 19–28.

[58] R. Lu, X. Lin, T. H. Luan, X. Liang, and X. Shen, "Pseudonym changing at social spots: An effective strategy for location privacy in vanets," *IEEE Transactions on Vehicular Technology*, vol. 61, no. 1, pp. 86–96, 2012.

[59] *Public Key Cryptography for the Financial Services Industry, The Elliptic Curve Digital Signature Algorithm (ECDSA)*, Std. American National Standard X9.62-2005, 2005.

[60] *UMAC: Message Authentication Code using Universal Hashing*, Std. rfc4418, 2006.

[61] *Intelligent transport systems(ITS); security; stage 3 mapping for ieee 1609.2*, ETSI TS Std. 102 867 v1. 1.1, 2012.

[62] *Dedicated Short Range Communications (DSRC) Message Set Dictionary*, SAE Std. J2735 V, 2016.

[63] J. Freudiger, M. Raya, and M. Félegyházi, "Mix-zones for location privacy in vehicular networks," in *ACM Workshop on Wireless Networking for Intelligent Transportation Systems (WiN-ITS)*, 2007.

[64] P. Balaji and L. Ling, "Mobimix: Protecting location privacy with mix-zones over road networks," in *2011 IEEE 27th International Conference on Data Engineering*, 2011, pp. 494–505.

[65] X. Liu and X. Li, "Privacy preservation using multiple mix zones," *Location Privacy Protection in Mobile Networks*, pp. 5–30, 2013.

[66] R. Yu, J. Kang, X. Huang, S. Xie, Y. Zhang, and S. Gjessing, "Mixgroup: Accumulative pseudonym exchanging for location privacy enhancement in vehicular social networks," *IEEE Transactions on Dependable and Secure Computing*, vol. 13, no. 1, pp. 93–105, 2016.

[67] B. Abdelwahab and M. Samira, "Urban pseudonym changing strategy for location privacy in vanets," *International Journal of Ad Hoc and Ubiquitous Computing*, pp. 49–64, 2016.

[68] ——, "Tapcs: Traffic-aware pseudonym changing strategy for vanets," *Peer-to-Peer Networking and Applications*, vol. 10, no. 4, pp. 1008–1020, 2017.

[69] K. Emara, W. Woerndl, and J. Schlichter, "Caps: Context-aware privacy scheme for vanet safety applications," in *Proceedings of the 8th ACM Conference on Security & Privacy in Wireless and Mobile Networks*, vol. 21, 2015.

[70] Y. Pan and J. Li, "Cooperative pseudonym change scheme based on the number of neighbors in vanets," *Journal of Network and Computer Applications*, vol. 36, no. 6, pp. 1599–1609, 2013.

[71] B. Ying, D. Makrakis, and Z. Hou, "Motivation for protecting selfish vehicles' location privacy in vehicular networks," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 12, pp. 5631–5641, 2015.

[72] L. Fischer, A. Aijaz, C. Eckert, and D. Vogt, "Secure revocable anonymous authenticated inter-vehicle communication (sraac)," in *4th Conference on Embedded Security in Cars (ESCAR 2006), Berlin, Germany*, 2006.

[73] F. Schaub, F. Kargl, Z. Ma, and M. Weber, "V-tokens for conditional pseudonymity in vanets," in *2010 IEEE Wireless Communication and Networking Conference*, 2010, pp. 1–6.

[74] W. Whyte, A. Weimerskirch, V. Kumar, and T. Hehn, "A security credential management system for v2v communications," in *2013 IEEE Vehicular Networking Conference*, 2013, pp. 1–8.

[75] N. Bißmeyer, J. Petit, and K. M. Bayarou, "Copra: Conditional pseudonym resolution algorithm in vanets," in *Wireless On-demand Network Systems and Services (WONS), 2013 10th Annual Conference on*, 2013, pp. 9–16.

[76] G. P. Corser, H. Fu, and A. Banihani, "Evaluating location privacy in vehicular communications and applications," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 9, pp. 2658–2667, 2016.

[77] S. Al-Shareeda and F. Özgüner, "Preserving location privacy using an anonymous authentication dynamic mixing crowd," in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, 2016, pp. 545–550.

[78] Y. Wei and Y. Chen, "Safe distance based location privacy in vehicular networks," in *2010 IEEE 71st Vehicular Technology Conference*, 2010, pp. 1–5.

[79] J. H. Song, V. W. S. Wong, and V. C. M. Leung, "Wireless location privacy protection in vehicular ad-hoc networks," in *2009 IEEE International Conference on Communications*, 2009, pp. 1–6.

[80] B. Hoh, M. Gruteser, H. Xiong, and A. Alrabady, "Achieving guaranteed anonymity in gps traces via uncertainty-aware path cloaking," *IEEE Transactions on Mobile Computing*, vol. 9, no. 8, pp. 1089–1107, 2010.

[81] Y. Pan, F. L. J. Li, and B. Xu, "An analytical model for random pseudonym change scheme in vanets," *Cluster Computing*, vol. 17, no. 2, pp. 413–421, 2014.

[82] A. Beresford and F. Stajano, "Location privacy in pervasive computing," *IEEE Pervasive Computing*, vol. 2, no. 1, pp. 46–55, 2003.

[83] J. Freudiger, M. H. Manshaei, J. Le Boudec, and J. Hubaux, "On the age of pseudonyms in mobile ad hoc networks," in *2010 Proceedings IEEE INFOCOM*, 2010, pp. 1–9.

[84] J. Freudiger, M. H. Manshaei, J.-P. Hubaux, and D. C. Parkes, "Non-cooperative location privacy," *IEEE Transactions on Dependable and Secure Computing*, vol. 10, no. 2, pp. 84–98, 2013.

[85] W. Marius, S. Pavel, D. Frank, and R. Kurt, "A classification of location privacy attacks and approaches," *Personal and ubiquitous computing*, vol. 18, no. 1, pp. 163–175, 2014.

[86] K. Emara, W. Woerndl, and J. Schlichter, "Vehicle tracking using vehicular network beacons," in *2013 IEEE 14th International Symposium on "A World of Wireless, Mobile and Multimedia Networks" (WoWMoM)*, 2013, pp. 1–6.

[87] B. Bloessl, C. Sommer, F. Dressler, and D. Eckhoff, "The scrambler attack: A robust physical layer attack on location privacy in vehicular networks," in *2015 International Conference on Computing, Networking and Communications (ICNC)*, 2015, pp. 395–400.

[88] D. Förster, H. Löhr, A. Grätz, J. Petit, and F. Kargl, "An evaluation of pseudonym changes for vehicular networks in large-scale, realistic traffic scenarios," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 10, pp. 3400–3405, 2018.

[89] E. Schoch, F. Kargl, T. Leinmuller, S. Schlott, and P. Papadimitratos, "Impact of pseudonym changes on geographic routing in vanets," in *Proc. Third European Workshop Security and Privacy in Ad Hoc and Sensor Networks (ESAS)*, 2006.

[90] K. Emara, "Safety-aware location privacy in vanet: Evaluation and comparison," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 12, pp. 10 718–10 731, 2017.

[91] J. Kang, R. Yu, X. Huang, and Y. Zhang, "Privacy-preserved pseudonym scheme for fog computing supported internet of vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 8, pp. 2627–2637, 2018.

[92] S. Lefèvre, J. Petit, R. Bajcsy, C. Laugier, and F. Kargl, "Impact of v2x privacy strategies on intersection collision avoidance systems," in *2013 IEEE Vehicular Networking Conference*, 2013, pp. 71–78.

[93] M. Hadded, P. Muhlethaler, A. Laouiti, R. Zagrouba, and L. A. Saidane, "Tdma-based mac protocols for vehicular ad hoc networks: a survey, qualitative analysis, and open research issues," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 4, pp. 2461–2492, 2015.

[94] M. J. Booysen, S. Zeadally, and G.-J. Van Rooyen, "Survey of media access control protocols for vehicular ad hoc networks," *IET communications*, vol. 5, no. 11, pp. 1619–1631, 2011.

[95] *IEEE Standard for Information technology– Local and metropolitan area networks– Specific requirements– Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 6: Wireless Access in Vehicular Environments*, IEEE Std. Std 802.11p, 2010.

[96] Z. Tong, H. Lu, M. Haenggi, and C. Poellabauer, "A stochastic geometry approach to the modeling of dsrc for vehicular safety communication," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 5, pp. 1448–1458, 2016.

[97] X. Yin, X. Ma, K. S. Trivedi, and A. Vinel, "Performance and reliability evaluation of bsm broadcasting in dsrc with multi-channel schemes," *IEEE Transactions on Computers*, vol. 63, no. 12, pp. 3101–3113, 2014.

[98] R. Zou, Z. Liu, L. Zhang, and M. Kamil, "A near collision free reservation based mac protocol for vanets," in *2014 IEEE Wireless Communications and Networking Conference (WCNC)*, 2014, pp. 1538–1543.

[99] E. Fonseca, A. Festag, R. Baldessari, and R. L. Aguiar, "Support of anonymity in vanets-putting pseudonymity into practice," in *2007 IEEE Wireless Communications and Networking Conference*, 2007, pp. 3400–3405.

[100] T. Aura, "Cryptographically generated addresses (cga)," Tech. Rep., 2005.

[101] S. N. Ethier and T. G. Kurtz, *Markov processes: characterization and convergence*. Wiley, 1986.

[102] M. Khabbaz, M. Hasna, C. M. Assi, and A. Ghrayeb, "Modeling and analysis of an infrastructure service request queue in multichannel v2i communications," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 3, pp. 1155–1167, 2014.

[103] G. Xiaohu, L. Zipeng, and L. Shikuan, "5g software defined vehicular networks," *IEEE Communications Magazine*, vol. 55, no. 7, pp. 87–93, 2017.

[104] Y. Ibrar, A. Iftikhar, A. Ejaz, G. Abdullah, I. Muhammad, and G. Nadra, "Overcoming the key challenges to establishing vehicular communication: Is sdn the answer?" *IEEE Communications Magazine*, vol. 55, no. 7, pp. 128–134, 2017.

[105] L. Guiyang, L. Jinglin, Z. Lin, Y. Quan, L. Zhihan, and Y. Fangchun, "sdnmac: A software-defined network inspired mac protocol for cooperative safety in vanets," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 6, pp. 2011–2024, 2018.

[106] J. Liu, J. Wan, B. Zeng, Q. Wang, H. Song, and M. Qiu, "A scalable and quick-response software defined vehicular network assisted by mobile edge computing," *IEEE Communications Magazine*, vol. 55, no. 7, pp. 94–100, 2017.

[107] J. Kang, R. Yu, X. Huang, and Y. Zhang, "Privacy-preserved pseudonym scheme for fog computing supported internet of vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 8, pp. 2627–2637, 2018.

[108] A. Boualouache, S. Senouci, and S. Moussaoui, "A survey on pseudonym changing strategies for vehicular ad-hoc networks," *IEEE Communications Surveys and Tutorials*, vol. 20, no. 1, pp. 770–790, 2018.

[109] Z. Tong, H. Lu, M. Haenggi, and C. Poellabauer, "A stochastic geometry approach to the modeling of dsrc for vehicular safety communication," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 5, pp. 1448–1458, 2016.

[110] X. Yin, X. Ma, K. S. Trivedi, and A. Vinel, "Performance and reliability evaluation of bsm broadcasting in dsrc with multi-channel schemes," *IEEE Transactions on Computers*, vol. 63, no. 12, pp. 3101–3113, 2014.

[111] Y. Kim, M. Lee, and T. J. Lee, "Coordinated multichannel mac protocol for vehicular ad hoc networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 8, pp. 6508–6517, 2016.

[112] X. Huang, R. Yu, J. Kang, N. Wang, S. Maharjan, and Y. Zhang, "Software defined networking with pseudonym systems for secure vehicular clouds," *IEEE Access*, vol. 4, pp. 3522–3534, 2016.

[113] G. Luo, S. Jia, Z. Liu, K. Zhu, and L. Zhang, "sdnmac: A software defined networking based mac protocol in vanets," in *2016 IEEE/ACM 24th International Symposium on Quality of Service (IWQoS)*, 2016, pp. 1–2.

[114] D. Boneh and X. Boyen, "Short signatures without random oracles and the sdh assumption in bilinear groups," *J. Cryptology*, vol. 21, no. 2, pp. 149–177, 2008.

[115] D. Lucarelli, I.-J. Wang *et al.*, "Decentralized synchronization protocols with nearest neighbor communication," in *Proceedings of the 2nd international conference on Embedded networked sensor systems*, 2004, pp. 62–68.

[116] M. E. Andrés, N. E. Bordenabe, K. Chatzikokolakis, and C. Palamidessi, "Geo-indistinguishability: Differential privacy for location-based systems," *arXiv preprint arXiv:1212.1984*, 2012.

[117] F. Chapeau-Blondeau and A. Monir, "Numerical evaluation of the lambert w function and application to generation of generalized gaussian noise with exponent 1/2," *IEEE Transactions on Signal Processing*, vol. 50, no. 9, pp. 2160–2165, 2002.

[118] Bitcoin. [Online]. Available: https://bitcoin.org/

[119] M. A. Ferrag, L. Maglaras, and A. Ahmim, "Privacy-preserving schemes for ad hoc social networks: A survey," *IEEE Communications Surveys and Tutorials*, vol. 19, no. 4, pp. 3015–3045, 2017.

[120] L. Sweeney, "K-anonymity: A model for protecting privacy," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 10, no. 5, pp. 557–570, 2002.

[121] A. Beresford and F. Stajano, "Mix zones: User privacy in location-aware services," in *Proc. Pervasive Computing and Comm. Workshops*, 2004, pp. 127–131.

[122] S. Gao, J. Ma, W. Shi, G. Zhan, and C. Sun, "Trpf: A trajectory privacy-preserving framework for participatory sensing," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 6, pp. 874–887, 2013.

[123] E. B. Sasson, A. Chiesa, C. Garman, and et. al, "Zerocash: Decentralized anonymous payments from bitcoin," in *2014 IEEE Symp. on Security and Privacy*, 2014, pp. 459–474.

[124] T. Ruffing, P. Moreno-Sanchez, and A. Kate, "Coinshuffle: Practical decentralized coin mixing for bitcoin," in *Proc. of the 19th European Symp. on Research in Computer Security (ESORICS)*, 2014, pp. 345–364.

[125] Cornelius, C. Kapadia, A. Kotz, and D. Peebles, "Coinshuffle: Practical decentralized coin mixing for bitcoin," in *Proc. of the 19th European Symp. on Research in Computer Security (ESORICS)*, 2014, pp. 345–364.

[126] D. Eckhoff, C. Sommer, T. Gansen, R. German, and F. Dressler, "Strong and affordable location privacy in vanets: Identity diffusion using time-slots and swapping," in *2010 IEEE Vehicular Networking Conference*, 2010, pp. 174–181.

[127] Z. Liu, Z. Liu, L. Zhang, and X. Lin, "Marp: A distributed mac layer attack resistant pseudonym scheme for vanet," *IEEE Transactions on Dependable and Secure Computing*, 2018.

[128] S. Du, H. Zhu, X. Li, K. Ota, and M. Dong, "Mixzone in motion: Achieving dynamically cooperative location privacy protection in delay-tolerant networks," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 9, pp. 4565–4575, 2013.

[129] F. Martin, J. Walter, and I. Baumgart, "Sybil-resistant pseudonymization and pseudonym change without trusted third parties," in *Proc. of the 14th ACM Workshop on Privacy in the Electronic Society*, 2015, pp. 65–74.

[130] Q. Li and G. Cao, "Providing privacy-aware incentives in mobile sensing systems," *IEEE Transactions on Mobile Computing*, vol. 15, no. 6, pp. 1485–1498, 2016.

[131] B. Ying, D. Makrakis, and Z. Hou, "Motivation for protecting selfish vehicles' location privacy in vehicular networks," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 12, pp. 5631–5641, 2015.

[132] B. Ying, D. Makrakis, and Z. Hou, "Motivation for protecting selfish vehicles' location privacy in vehicular networks," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 12, pp. 5631–5641, 2015.

[133] A. Boualouache, S. Senouci, and S. Moussaoui, "Towards an efficient pseudonym management and changing scheme for vehicular ad-hoc networks," in *2016 IEEE Global Communications Conference (GLOBECOM)*, 2016, pp. 1–7.

[134] A. Turner, "Convergence of markov processes," Master's thesis, University of Cambridge, Cambridge, 2002.

[135] G. Danezis, "Statistical disclosure attacks," in *International Information Security Conference (IFIP)*, 2003, pp. 421–426.

[136] Raymond and Jean-Franc¸ois., "Traffic analysis: Protocols, attacks, design issues, and open problems," *Designing Privacy Enhancing Technologies*, 2001.