Secure and privacy-preserving schemes in vehicular networks

Kong, Qinglei

2018

Kong, Q. (2018). Secure and privacy-preserving schemes in vehicular networks. Doctoral thesis, Nanyang Technological University, Singapore.

<http://hdl.handle.net/10356/75653>

<https://doi.org/10.32657/10356/75653>

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

**SECURE AND PRIVACY-PRESERVING SCHEMES IN
VEHICULAR NETWORKS**

**KONG QINGLEI
SCHOOL OF ELECTRICAL & ELECTRONIC
ENGINEERING
2018**

SECURE AND PRIVACY-PRESERVING SCHEMES IN VEHICULAR NETWORKS

KONG QINGLEI

School of Electrical & Electronic Engineering

A thesis submitted to the Nanyang Technological University
in partial fulfillment of the requirement for the degree of
Doctor of Philosophy

2018

To my family and teachers

Acknowledgements

I would like to express my gratitude to my supervisor Professor Ma Maode, I want to thank you for your encouragement in my research and your continuous support in the project which makes my PhD study possible. I would also show my sincere gratitude to Professor Lu Rongxing, who is also my supervisor, for his guidance, patience and encouragement during my whole PhD study progress. Prof Lu has helped me a lot in build a positive research attitude, learn how to solve complicated problems and develop dense interests in academic researches. Under their busy schedules, they have been always offering me for advice, encouragement and insightful comments.

I would especially like to thank my amazing parents for the love, support and constant encouragement they brought me. My parents are the people who made everything possible, they are always trusting in me and encouraging me to pursue my life and career goals.

I would show my appreciation for the colleagues in the Media Technology Laboratory. Among them, Dr. Huang Ling, Ms. Wang Xiaohong, and Mr. Liu Ankang, your support, laughter and suggestions have made me inspired and smiling. You are and always will be my good friends.

Thankfulness also are shown for the research associate position in Nanyang Technological University, especially for the Development Of NTU/NXP Smart Mobility Test-bed project, which makes my position possible.

Last but definitely not least, I would like to dedicate this thesis to my grandparents. It was you to accompany me and make a happy childhood memory possible. Although it has been years since you have passed, I still take your love with me, every second.

Table of Contents

Acknowledgements	iii
Table of Contents	v
Abstract	ix
List of Figures	xi
List of Tables	xiii
List of Abbreviations	xv
1 Introduction	1
1.1 Existing Standards for Vehicular Networks	1
1.1.1 Vehicular Ad Hoc Networks	1
1.1.2 LTE-A Vehicular Networks	2
1.2 Vehicular Cloud	4
1.2.1 Characteristics of Vehicular Cloud	5
1.2.2 Applications of Vehicular Cloud	6
1.3 Motivations and Objectives	7
1.3.1 Motivations	7
1.3.2 Objectives	8
1.4 Research Contributions	9
1.5 Thesis Outline	11
2 Literature Review	13
2.1 Related Works in Vehicular Cloud	13
2.1.1 Security Threats and Security Requirements in Vehicular Cloud	14
2.1.2 Related works on securing vehicular cloud	16
2.1.2.1 Related works on securing VANETs	17
2.1.2.2 Related works on securing Mobile Cloud	18
2.1.3 Vehicular Data Dissemination	19
2.2 Secure Data Processing	22
2.2.1 Privacy-preserving data aggregation	22
2.2.2 Secure scalar product computation	25
2.2.3 Secure location-based data queries	27

2.3	Secure Key Management in LTE-A networks	31
2.3.1	EPS-AKA process	31
2.3.1.1	Extended key architecture	32
2.3.2	Review of EPS handover key management	33
2.3.3	Mobile relay networks and Coordinated Multi-Point	35
2.3.3.1	Mobile relay networks	35
2.3.3.2	CoMP	36
3	Achieve location privacy-preserving range query in vehicular sensing	39
3.1	Introduction	39
3.2	System Model, Security Requirement and Design Goal	42
3.2.1	System Model	42
3.2.2	Security Requirements	44
3.2.3	Design Goals	45
3.3	Proposed Privacy-preserving Range Query Scheme	46
3.3.1	Preliminaries	46
3.3.2	System Initialization	47
3.3.3	On-Board Data Query Generation	48
3.3.4	Data Report Generation	49
3.3.5	Privacy Preserving Data Filtering	51
3.3.6	Data Report Aggregation	52
3.4	Security Analysis	53
3.5	Performance Evaluation	55
3.5.1	Parameter Setup	55
3.5.2	Computation Complexity	56
3.5.3	Communication Overhead	58
3.6	Summary	60
4	A privacy-preserving sensory data sharing scheme in Internet of Vehicles	63
4.1	Introduction	64
4.2	System Model, Security Requirements, and Design Goals	66
4.2.1	System Model	66
4.2.2	Security Requirements	68
4.2.3	Design Goals	70
4.3	Preliminaries	71
4.3.1	Modified Paillier Cryptosystem	71
4.3.2	Chinese Remainder Theorem	72
4.4	Proposed Sensory Data Sharing Scheme	72
4.4.1	System Initialization	73
4.4.2	Sensory Data Collection	73
4.4.3	Sensory Data Acquisition	76
4.4.4	Discussion	78
4.5	Security Analysis	80
4.6	Performance Evaluation	83

4.6.1	Parameter Setup	84
4.6.2	Computation Complexity	85
4.6.3	Communication Overhead	87
4.7	Summary	88
5	A Privacy-preserving and Verifiable Querying Scheme in Vehicular Fog Data Dissemination	89
5.1	Introduction	90
5.2	Model, Requirements and Design Goal	92
5.2.1	System Model	92
5.2.2	Security Requirements	93
5.2.3	Design Goals	94
5.3	Preliminaries	95
5.3.1	Paillier Cryptosystem	95
5.3.2	Bilinear Maps	96
5.4	Proposed Verifiable Data Querying and Dissemination Scheme	96
5.4.1	System Initialization	97
5.4.2	Group Formulation	98
5.4.3	Data Query Generation	99
5.4.3.1	Data Query Generation of V_i	99
5.4.3.2	Data Query Generation of V_a	101
5.4.4	Data Report Aggregation and Reading	102
5.4.5	Data Query Verification	103
5.5	Security Analysis	105
5.6	Performance Evaluation	106
5.6.1	Analysis on the Number of Data Querying Vehicles	107
5.6.2	Computation Complexity	108
5.6.3	Communication Overhead	110
5.7	Summary	110
6	Secure handover key management protocols in LTE-A Vehicular Networks	113
6.1	Achieve secure handover session key management via mobile relay	113
6.1.1	Introduction	113
6.1.2	System Model, Security Requirement and Design Goal	117
6.1.2.1	System Model	117
6.1.2.2	Security Requirements	118
6.1.2.3	Design Goals	119
6.1.3	Preliminaries	120
6.1.3.1	Bilinear Pairing	120
6.1.3.2	Proxy Re-encryption	121
6.1.4	Proposed Handover Key Management Scheme via Mobile Relay	121
6.1.4.1	System Initialization	121
6.1.4.2	Handover Key Management Scheme in MRN Scenario	123
6.1.5	Security Analysis	126

6.1.6	Performance Evaluation	128
6.1.6.1	Computational Cost	128
6.1.6.2	Communication and Storage Overhead	133
6.2	Achieving secure CoMP joint transmission handover	134
6.2.1	Introduction	134
6.2.2	Model, Requirements and Design Goal	135
6.2.2.1	System Model	136
6.2.2.2	Security Requirements	136
6.2.2.3	Design Goal	137
6.2.3	Proposed Handover Key Management Scheme for CoMP	138
6.2.3.1	Preliminaries	138
6.2.3.2	System Initialization	139
6.2.3.3	Secure Handover Key Management Process	140
6.2.4	Security Analysis	142
6.2.5	Performance Evaluation	143
6.3	Summary	145
7	Conclusions and Future Work	147
7.1	Contributions	147
7.2	Future Work	149
	Appendix	151
A.	Author's Publications	151
	Bibliography	153

Abstract

In the thesis, secure and privacy-preserving schemes are designed in vehicular networks, and the specific details of the completed works are shown as following.

Due to the increasing popularity of intelligent vehicles and the recent flourishing of cloud computing, it becomes a recent research trend of extending cloud computing to vehicles by leveraging under-utilized on-board capabilities of vehicles. In vehicular cloud, three security schemes have been designed.

In the first scheme, a novel location privacy-preserving data query scheme is proposed, with the proposed scheme, a data requester can retrieve the data generated by vehicular sensors from the distributive vehicular on-board storage with high accuracy. Since the majority of intelligent vehicles are generally considered as not fully utilized, these vehicles are candidates to provide distributive data collection and storage services. By exploiting the homomorphic Paillier cryptosystem technique and the structured binary scalars to represent the positions of data requesters and the data generating vehicles, the proposed scheme can achieve the location matching of the involved entities with privacy-preservation.

In the second scheme, an efficient data sharing scheme and location privacy-preservation in Internet of vehicles (IoV) is proposed, which enables the collection and distribution of the data captured by vehicular sensors. The data captured by vehicular sensors records a myriad of physical phenomena about the surrounding environments, which enables the data sharing among vehicles and deployed roadside infrastructures to further improve traffic safety and on-board experience in the intelligent transportation system. With the proposed scheme, the multi-dimensional sensory data captured at different locations are first structured by the Chinese Remainder Theorem, then the modified Paillier cryptosystem is exploited to achieve the location privacy-preserving sensory data aggregation. Meanwhile, the proposed scheme exploits the proxy re-encryption technique to achieve the sensory data acquisition at the network edge, without the involvement of the trusted central entity.

In the third scheme, a secure request-response based vehicular data dissemination scheme in the parking lot scenario is proposed. The roadside units (RSUs) deployed in vehicular ad hoc networks (VANETs) can act as information servers to provide information to vehicles under their coverage area. The proposed scheme exploits an invertible matrix to structure multiple data requests, and encrypts the data requests with the homomorphic Paillier Cryptosystem technique. Based on the data requests aggregation, the RSU can recover the individual data request while protecting the unlinkability between the data query and its origin. In addition, the RSU can verify the correctness of the recovered data requests without privacy disclosure by exploiting an identity-based batch verification technique.

To improve the quality of the vehicle-to-infrastructure (V2I) communications, the LTE-A network is brought to vehicular network for its dense deployment, high bandwidth and high-speed movement support. Specifically, two secure handover key establishment schemes are designed.

In the first scheme, a secure handover session key management scheme is designed in mobile relay LTE-A network. To be more specific, a session key is successfully established between the on-board user equipment (UE) and the Donor evolved Node B (DeNB), which is firstly produced by the UE and then securely delivered to the DeNB, which achieves the security goal of the forward and backward key separations. Meanwhile, to decrease the communication and computational burden of the proposed scheme, the proxy re-encryption technique is also employed, i.e., the session keys are initially encrypted by the public key of the mobility management entity (MME) and then re-encrypted by the mobile relay node (MRN), so that the target DeNB can recover the value of the session key with its private key without involving the core entity MME.

In the second scheme, a novel secure coordinated multi-point (CoMP) joint transmission handover key establishment scheme in LTE-A vehicular networks is proposed. Specifically, to achieve the diversity gain brought by the CoMP joint transmission and accommodate to protect the backward/forward key separation, the session key is initially produced by the vehicle and then securely delivered towards the cooperating eNBs, and then decrypted by each cooperating eNB respectively.

List of Figures

1.1	VANET architecture	2
1.2	LTE vehicular network architecture	3
1.3	Vehicular cloud architecture	5
2.1	EPS-AKA procedure [1]	32
2.2	Extended key hierarchy in the EPS [1]	32
2.3	Message Flow during Intra-MME Handover	33
3.1	Data query architecture.	43
3.2	Registration and secret sharing among the data requester and vehicles.	47
3.3	Defined grid cell architecture/locations of vehicles and the data query area in grid cells.	48
3.4	Computation complexity of data server. (a) Computation complexity of the data server with the traditional scheme; (b) Computation complexity of the data server with the proposed scheme.	58
3.5	Comparison of the computation complexity of the data requester and vehicle. (a) Computation complexity of the data requester; (b) Computation complexity of one data uploading vehicle.	58
3.6	Vehicles to data server communication overhead. (a) Communication overhead of vehicles during data uploading with the traditional scheme; (b) Communication overhead of vehicles during data uploading with the proposed scheme.	59
3.7	Communication overhead from the data requester to the data server.	60
4.1	Vehicular Sensory Data Sharing System	67
4.2	Message flow during data sharing	74
4.3	Probability of data querying failure	79
4.4	Computation complexity of data server. (a) Computation complexity of the data server with the traditional scheme; (b) Computation complexity of the data server with the proposed scheme.	83
4.5	Communication cost of Vehicle-to-RSU during data collection. (a) Cost of the proposed scheme; (b) Cost of the traditional scheme.	86
4.6	Communication cost of Vehicle-to-RSU during data acquisition. (a) Cost of the proposed scheme; (b) Cost of the traditional scheme.	86
5.1	Vehicular data querying architecture under consideration	93
5.2	Key distribution	97
5.3	Secret distribution among vehicles	98

5.4	Vehicular data querying message flow	99
5.5	Computation complexity of the proposed vehicular data querying scheme. (a) Computation complexity of the RSU; (b) Computation complexity of each vehicle.	107
5.6	Computation complexity of the proposed vehicular data querying scheme. (a) Computation complexity of the RSU; (b) Computation complexity of each vehicle.	108
6.1	System model under consideration	118
6.2	Message flow in MRN scenario during UE handover	123
6.3	Computational delay of DeNB, when $T_{MRN} \leq T_{DeNB}$	129
6.4	Computational delay of DeNB, when $T_{MRN} > T_{DeNB}$	129
6.5	Comparison of computational delay of DeNB w.r.t. the number of UE	130
6.6	Comparison of computational delay of DeNB w.r.t. the number of MRN	131
6.7	Computational delay of UE w.r.t. the number of UE	132
6.8	Computational delay of UE w.r.t. the number of MRN	132
6.9	Maximum computational delay of UE w.r.t. the number of MRN and UE	133
6.10	CoMP joint transmission handover in LTE-A vehicular networks	137
6.11	Message flows during CoMP joint transmission handover	140
6.12	eNB valid serving ratio with different vehicle density and different average vehicle speed	144
7.1	Structure of the proposed schemes in this thesis	148

List of Tables

1.1	Comparisons between Vehicular Mobile cloud	6
2.1	Comparative Study of Secure Aggregation Schemes	23
3.1	Notations of the Proposed Range Querying Scheme	46
4.1	Notations of Proposed Sensory Data Sharing Scheme	72
4.2	Parameters Settings	84
5.1	Notations of Proposed Querying and Dissemination Scheme	97
6.1	Notations of Secure Session Key Establishment in MRN handover	122
6.2	Notations of Proposed Key Establishment Scheme in CoMP Handover	138
6.3	Parameters Settings	144

List of Abbreviations

VANET	Vehicular Ad Hoc Network
DSRC	Dedicated Short Range Communications
V2V	Vehicle-to-Vehicle
V2I	Vehicle-to-infrastructure
RSU	Roadside Unit
QoS	Quality of Service
ICT	Information Communications Technology
OBU	On-board Unit
EPS	Evolved Packet System
EPC	Evolved Packet Core
E-UTRAN	Evolved Universal Terrestrial Radio Access Network
eNB	evolved Node Base
MME	Mobility Management Entity
HSS	Home Subscriber Server
ITS	Intelligent Transportation System
PIR	Private Information Retrieval
OT	Oblivious Transfer
AuC	Authentication Center
HLR	Home Location Register
AKA	Authentication and Key Agreement
CoMP	Coordinated Multi-Point

WAVE	Wireless Access in Vehicular Environment
PKI	Public Key Infrastructure
IoT	Internet of Things
UE	User Equipment
USIM	Universal Subscriber Identity Module
ASME	Access Security Management Entity
RRC	Radio Resource Control
AS	Access Stratum
NAS	Non-Access Stratum
KDF	Key Derivation Function
NCC	Next-hop Chaining Counter
NH	Next Hop
ICI	Inter-Cell Interference
CS/CB	Coordinated Scheduling or Beamforming
JP	Joint Processing
CSI	Channel State Information
TA	Trusted Authority
MTC	Machine Type Communication
MRN	Mobile Relay Node
CAPEX	Capital Expenditure
OPEX	Operating Expenditure
SeNB	Serving eNB
TeNB	Target eNB
S-GW	Serving Gateway
P-GW	Packet Gateway

Chapter 1

Introduction

1.1 Existing Standards for Vehicular Networks

The formulation of vehicular network by connecting the vehicles and roadside infrastructures can support applications in Intelligent Transport System (ITS), which can significantly enhance traffic driving safety to reduce fatalities on the road and improves the traffic efficiency to reduce the traffic congestion.

1.1.1 Vehicular Ad Hoc Networks

IEEE 802.11p [2] is considered as the de facto standard to support ITS applications in vehicular ad hoc networks (VANET), which is in accordance with the dedicated short range communications (DSRC) band (75MHz bandwidth). One vehicle can communicate with other vehicles through vehicle to vehicle communication (V2V) or communicate with fixed equipment next to the road (which is referred to as road side unit (RSU) in VANET) through vehicle to infrastructure communication (V2I), as illustrated in Fig. 1.1. VANETs can be exploited for providing a broad range of safety and non-safety applications, such as vehicle collision avoidance, traffic congestion alert, vehicular navigation service, location-based value added services [3]. However, due to the lack of pervasive roadside infrastructure deployment and the defined limited

transmission range, VANET can only provide the short-range and intermittent V2I connectivity and it also suffers from the constraints in scalability and unbounded delay.

FIGURE 1.1. VANET architecture

1.1.2 LTE-A Vehicular Networks

Given the wide area capillary deployment of eNBs in LTE-A network, the integration of LTE-A network into the vehicular networks as a potential access technology can help the vehicles to improve the quality of V2I communications even at high moving speed [4]. In comparison with the IEEE 802.11p standard, the LTE-A standard proposed by 3GPP offers superior performances in terms of higher throughput and lower latencies. LTE can satisfy the requirements of quality of service (QoS) in vehicular applications such as infotainment and Internet connections [5], and it also can be applicable to scenarios with low traffic density and challenging propagation conditions. However, the centralized LTE architecture may not be able to support the direct V2V communications.

FIGURE 1.2. LTE vehicular network architecture

Fig. 1.2 illustrates the architecture of the evolved packet system (EPS) in LTE vehicular network, which consists of the evolved packet core (EPC) and the evolved universal terrestrial radio access network (E-UTRAN). The E-UTRAN only contains the evolved NodeB (eNB) and is responsible for the radio communications between the mobile equipment and the EPC, and the EPC consists of the mobility management entity (MME), home subscriber server (HSS), serving gateway (S-GW), and packet gateway (P-GW). The details are shown as follows:

- The MME is responsible for the high-level operation of the mobile equipment, and it sends control-plane signalling related to security and mobility management issues. For security functions, it also performs the ciphering and integrity protection of Non-Access Stratum (NAS) signaling.
- The HSS includes the functions of both authentication center (AuC) and home location register (HLR), such as storing user subscription information and generating security information.
- The S-GW acts as the router of the user-plane signaling between eNB and P-GW, which handles the user-plane packet data transmitted by the E-UTRAN and delivers towards P-GW.

- The P-GW interfaces with the external packet data network, and performs the IP-related functions like packet routing and address allocation.

1.2 Vehicular Cloud

With the rapid development of the automotive industry and information communications technology (ICT), intelligent vehicles are increasingly commonplace. Such vehicles are assumed to be equipped with various on-board devices, such as on-board computers, radio transceivers, sensors, and storage equipment. Meanwhile, the majority of intelligent vehicles are generally regarded as not fully utilized [6], which can be treated as potential computation, communication, and sensing equipment, and are perfect candidates to perform the role of providing on-demand services, without requiring additional investment on fixed roadside infrastructures [7]. To make full exploitation of the under-utilized resources, the concept of vehicular cloud computing is defined as [8]: *a group of largely autonomous vehicles whose corporate computing, sensing, communication and physical resources can be coordinated and dynamically allocated to authorized users*. By exploiting the vehicular cloud, one can leverage the under-utilized capabilities of vehicles, by integrating on-board computing resources with traditional cloud computing architecture, and provide on-demand solutions for unpredicted events in a proactive fashion within the vehicular networks [9, 10]. Similar to VANET, there mainly exist two types of vehicular clouds: infrastructure-based vehicular cloud and autonomous vehicular cloud [8]. For the infrastructure-base vehicular cloud, the vehicles can access the services through infrastructures deployed along the roadside. For the autonomous vehicular cloud, vehicles are organized to support emergencies in an ad hoc fashion.

The vehicular cloud architecture mainly consists of three components: vehicles, communication interface and cloud, as shown in Fig. 1.3. The involved vehicles are equipped with on-board units (OBUs), including sensors, actuators, communication devices, GPS devices, and embedded computers [11], which enables the vehicles to perform various sensing, networking and data processing tasks. The communication component consists both vehicle-to-vehicle (V2V)

FIGURE 1.3. Vehicular cloud architecture

and vehicle-to-infrastructure (V2I) communications, which is responsible for the data interaction among vehicles, roadside units (RSUs) and the cloud. The cloud component possess massive and complex computation and storage capability, which also supports remote access.

1.2.1 Characteristics of Vehicular Cloud

Mobile devices can both act as cloud users and service providers in mobile cloud, which generates a significant portion of mobile resources and services [12]. As a special component of mobile cloud, vehicular cloud holds a few unique characteristics:

- *High mobility.* In vehicular cloud, the vehicles are constantly moving, such that the V2V and V2I interaction time is stringent, which leads to high requirement in terms of computation and communication delay.
- *Fluctuation in spatial-temporal domain.* Vehicles follow the real-world road condition, and the number of vehicles fluctuates with respect to space and time. Thus, the number of vehicles at specific time and location is highly predictable.

- *Strong processing capability.* Modern vehicles are normally equipped with strong on-board computer, various sensors, and self-charging batteries, which enables the vehicular cloud to perform various tasks by coordinating the on-board resources and roadside infrastructures.

We use Table 1.1 to compare the differences between vehicular cloud and mobile cloud, in terms of connectivity, data transmission, processing capability and applications.

TABLE 1.1. Comparisons between Vehicular Mobile cloud

	Vehicular Cloud	Mobile Cloud
Connectivity	Intermittent with Stringent Delay requirement	Continuous with no Delay Requirement
Data Transmission	Trans-Vehicle-to-vehicle, infrastructure	vehicle-to-Point-to-point
Capability	Strong Processing Power, large Battery	Low Processing Power, Limited Energy Supply
Applications	Urban Sensing, Content Distribution, Social Computing, Mobile Healthcare Intelligent Transportation	

1.2.2 Applications of Vehicular Cloud

To fully exploit the excess and under-utilized computing, sensing and communicating resources, a few applications are developed under the vehicular cloud.

- *Parking lot datacenter.* To fully exploit the long-term parking computational resources of vehicles at the airport, a datacenter at the airport is developed in [13], such that the parking vehicles can formulate a vehicular cloud with their on-board processing units, and assign computational tasks through resource scheduling. Since the computing and storage resources of vehicles are unused, a computer cluster can be created with the vehicles in the car park, and formed vehicular cloud can provide ample data storage.
- *Intelligent traffic management.* With the introduction of the vehicular network, the traffic and environmental information can also be collected by vehicles through the on-board

sensors, instead of the traditional traffic data generation through the deployed inductive loops, roadside cameras and roadside sensors. A Navigator Assisted Vehicular route Optimizer (NAVOPT) is proposed to sense the segment traffic congestion [14], which enables the construction of the traffic load map for the traffic delay estimation. Dynamic traffic light management [8] is also proposed to solve the traffic congestion, by coordinating the on-board resources of the participating vehicles.

- *Disaster evacuation.* An intelligent disaster management system is proposed in [15], which integrates the vehicular network and the mobile cloud computing technology. In order to perform the emergency evacuation management, the self-organized vehicles can exploit the collected real-time data and calculate the available resources with the possible evacuation routes.
- *Vehicular sensory data sharing.* Vehicles are equipped with a plethora of on-board sensors, which record a myriad of physical phenomena. By interacting the data generated by on-board sensors with the neighboring vehicles, various value-added services can be produced [10]. A proactive monitoring application MobEyes is designed in [16], which exploits the vehicular sensors on road to record the phenomenon of the surrounding events for further investigation. The autonomous driving application called CarSpeak [17] is developed to enable a vehicle to access the sensory data captured by the neighboring vehicles.

1.3 Motivations and Objectives

1.3.1 Motivations

Nowadays, a typical intelligent vehicle is usually equipped with a strong processing equipment, a large on-board storage device and various on-board sensors, however, these on-board devices are mostly under-utilized. To fully exploit these on-board resources, various schemes

needs to be devised to provide various safety-related and value-added applications in vehicular cloud. Meanwhile, the vehicles are moving in dynamic topologies, and the vehicular-generated/queried data are mostly location-dependent. Thus, to protect the privacy-related information of the vehicles, various security schemes are needed to preserve the privacy of the involved vehicles in vehicular cloud.

In LTE-A networks, the session keys established between UEs and eNBs are designed to protect the confidentiality of the control-plane signalling and the user-plane data transmission, and guarantee the data integrity of the control-plane signalling. The current key establishment process is realized through the synchronization of key chaining counter, which is under the potential threat of desynchronization attack and not applicable towards various network architectures proposed in LTE-A networks, such as mobile relay and coordinated multipoint (CoMP) joint transmission.

1.3.2 Objectives

The research objectives of this thesis are existed in two aspects. The first objective is to investigate the secure and privacy-preserving problems existed in the vehicular cloud data processing. To be more specific, this thesis aims to protect the sensory data generated by vehicular sensors, and preserve the privacy of the involved vehicles (during both the data collecting and querying processes). Meanwhile, since VANETs are experiencing the problems of scarce transmission bandwidth and stringent delay requirements, we aim to improve the communication and computation cost introduced by the secure and privacy-preserving data processing.

The second objective is to study the secure handover schemes existed in LTE-A vehicular networks. As the LTE-A networks are subject to various network topologies, two secure handover key management schemes are designed under the mobile relay and Coordinated Multi-Point (CoMP) scenarios, such that the control-plane signalling and data-plane transmission can be protected.

1.4 Research Contributions

The research in this thesis focus on the development of a series of secure and privacy-preserving schemes in vehicular networks. Specifically, the main contributions of this thesis are shown as follows:

- A novel range query scheme is proposed, which helps the data requester to accurately retrieve the sensed data from the distributed on-board storage in VANETs with location privacy preservation. The proposed scheme exploits structured scalars to denote the locations of data requesters and vehicles, and achieves the privacy-preserving location matching with the homomorphic Paillier cryptosystem technique. Detailed security analysis shows that the proposed range query scheme can successfully preserve the location privacy of the involved data requesters and vehicles, and protect the confidentiality of the sensed data. In addition, performance evaluations are conducted to show the efficiency of the proposed scheme, in terms of computation delay and communication overhead. Specifically, the computation delay and communication overhead do not dependent on the length of the scalar, and it is only proportional to the number of vehicles.
- To improve the communication and computational burden introduced by the querying processes, a novel efficient and location privacy-preserving data sharing scheme in IoV is proposed, which enables the collection and distribution of the data captured by vehicular sensors. During the vehicular sensory data collection phase, the proposed scheme structures the multi-dimensional sensory data captured at different locations and exploits the modified Paillier cryptosystem to achieve the location privacy-preserving sensory data aggregation, in which the aggregation result can be fully recovered by the trusted central entity. During the vehicular sensory data acquisition phase, the proposed scheme exploits the proxy re-encryption technique to achieve the location privacy-preserving aggregated sensory data querying at the network edge, i.e., without the involvement of the trusted central entity. Numerical analysis are performed to validate the effectiveness of the proposed scheme, i.e., low data querying failure probability. Security analysis and

performance evaluation are also conducted to validate the security properties and show the computation and communication efficiency of the proposed scheme, respectively.

- To disrupt the linkage between a data request and its origin vehicle, a secure querying scheme is proposed in vehicular data dissemination. The proposed scheme introduces an invertible matrix to structure multiple data requests, and encrypts the data requests with the homomorphic Paillier Cryptosystem technique. For the data communication between the vehicles and the RSU, the data aggregation is performed on the ciphertext at the RSU without decryption. Based on the aggregation result, the RSU can recover the data requests while protecting the unlinkability between the data query and its origin. In addition, the RSU can verify the correctness of the recovered data requests by exploiting batch verification, without violating the privacy of the data requesting vehicles. Through security analysis, we demonstrate that the proposed scheme can achieve the security goals of unlinkability, verifiability, and confidentiality. Through performance evaluation, we demonstrate that the proposed scheme can adapt to the fluctuating number of the data querying vehicles, and significantly reduce the computation complexity and communication overhead.
- For the vehicular network scenario with high mobility and various network topologies, two novel secure handover key management scheme in LTE-A vehicular networks are proposed. In the first scheme, a secure handover session key management scheme via mobile relay in LTE-A networks, in which the session keys initially encrypted with the public key of the MME will be re-encrypted by a mobile relay, so that other DeNBs can later decrypt the session keys with their own private keys while without the direct involvement of the MME. In the second scheme, a novel secure CoMP joint transmission handover key management scheme in LTE-A vehicular networks, in which the session key is generated by the vehicle and securely delivered towards the cooperating eNBs, and then decrypted by each cooperating eNB respectively.

1.5 Thesis Outline

The organization of the rest of the thesis is shown as follows. Chapter 2 reviews the related works, in terms of the previous works on securing vehicular cloud, secure data processing and secure key management in LTE-A networks. A novel location privacy-preserving range query scheme is proposed in Chapter 3, which helps to accurately retrieve the sensed data from the distributed on-board storage in VANETs. In Chapter 4, a novel efficient and location privacy-preserving data sharing scheme in IoV is proposed, which enables the collection and distribution of the data captured by vehicular sensors. A secure and verifiable request-response based vehicular data dissemination scheme in parking lot scenario is proposed in Chapter 5. In Chapter 6, two novel secure handover key management scheme in LTE-A vehicular networks are proposed, in terms of mobile relay and CoMP joint transmission respectively. Finally, conclusions and future research work are identified in Chapter 7.

Chapter 2

Literature Review

This chapter first provides a general review of vehicular cloud, which includes the security threats and requirements of vehicular cloud, vehicular data dissemination, and the related works on securing vehicular cloud. Then a brief review about the secure data computation is performed, which consists of the privacy-preserving data aggregation, secure scalar product computation, and location-based data queries. In addition, a brief review related to the secure key management in LTE-A networks is presented, such as the EPS-authentication and key agreement (AKA) process, EPS key management, and mobile relay and coordinated multipoint (CoMP) techniques.

2.1 Related Works in Vehicular Cloud

In this section, we briefly review the security threats and security requirements in vehicular cloud, vehicular data dissemination, and the related works in securing vehicular cloud.

2.1.1 Security Threats and Security Requirements in Vehicular Cloud

As vehicular cloud is a hybrid technology consists of the combination of VANETs, cloud computing and wireless sensing network, it inherits the security and privacy threats of all the involved techniques [7]. In vehicular cloud environment, an attacker can equally share the same physical machine and infrastructure as their target vehicles, which offers them the great opportunity to gain more advantages than the attacker in the traditional cloud. Meanwhile, similar to cloud computing, the confidentiality and integrity of the sensitive information stored in vehicular cloud is of concern to the users. In case of data loss or data leakage, the data confidentiality, integrity and availability could be violated [18]. Specifically, the security threats in vehicular cloud can be classified in the following categories [7, 19].

- *Data tampering.* In vehicular cloud, the attacker can modify, alter and forge the information maintained in the vehicular data storage or during the vehicular data transmission.
- *Information disclosure.* In vehicular cloud, the attacker can access the content of the vehicular data kept in the distributed vehicular data storage, or violate the vehicular data during the data transmission process.
- *Identity spoofing.* In vehicular cloud, an attacker pretends to be a legitimate user to access vehicular data or gain illegitimate advantages, by abusing the identities or the security credentials of the authorized users.
- *Repudiation.* The attacker can manipulate or forge the identification of the vehicular data, actions and the related operations, due to the lack of the tracking and user log mechanism in vehicular cloud. That is, repudiation indicates the capability of denying a fact or an action.
- *Denial of service.* This is a common threat that the attacker attempts to consume the system resources in vehicular cloud, and disable the vehicular cloud resources provided to the legitimate users.

- *Privacy violation.* In vehicular cloud, the attackers can infer or uncover the personally identifiable information such as identities and geographic locations, based on the vehicular data storage, or during the data transmission.

In addition, since vehicular cloud involves the sharing of the computation and storage equipment of vehicles, an adversary may attack the vehicular cloud through malware injection [20]; to be more specific, an attacker can access the same physical and virtual machines in vehicular cloud as their targets [19]. To resist the security threats mentioned above, the following security requirements should be satisfied in the vehicular cloud [7, 19].

- *Confidentiality.* The content of the data stored on the vehicular storage, and the content of the vehicular cloud related data transmitted among the vehicles and infrastructures, should not be disclosed.
- *Integrity.* The stored or transmitted data should not be tampered or modified in vehicular cloud, and the received data should be correctly verified. For instance, the data owner can periodically check the correctness of the data stored in vehicular cloud, or data processing performed by vehicular cloud [21, 22].
- *Authentication.* Generally speaking, authentication means the verification of whether someone or something is, and determines who or what it claims to be. In vehicular cloud, the owner of the vehicular data should be authenticated, while the sender and receiver of the vehicular data should also be properly authenticated. Three metrics can be adopted such as ownership of identity (such as identity card or security token), knowledge of things (such as password or personal identification number), and biometrics (such as signature, voice or fingerprint).
- *Non-repudiation.* In general, non-repudiation refers to the association actions or changes to a unique individual. For example, in vehicular cloud, a user cannot deny the authenticity of their signature on vehicular data storage or the origin of a received message.

- *Availability.* The hardware, software and communication infrastructures and resources in vehicular cloud, should be available for the authorized vehicles or users when they are being requested.
- *Privacy preservation.* Since a vehicles identity is highly related to its owners identity, the tracking of a vehicle can lead to the violation of the vehicle owner's identity. Vehicles are also highly location dependent, however, the disclosure of location information or moving trajectory can often result in the leakage of office and home address, which violate the location privacy of vehicles. Thus, the location privacy of vehicles should be preserved in vehicular cloud.
- *Sybil Attack.* By creating a large number of pseudo-identities, an adversary can represent multiple entities in vehicular network to acquire a disproportionate level of control over the network, e.g., disseminate compromised information to gain unauthorized privileges.
- *Collusion Attack.* In vehicular systems, several users (in terms of vehicles and RSUs) may deviate from the defined sessions by sharing their own contents, or even collude (two or more users) to reveal the contents owned by any other entity.

Due to the fluctuating characteristic of vehicles in the spatial-temporal domain, the security schemes designed in vehicular cloud should be adaptive to the dynamically changing number of participating vehicles, and the proposed security schemes should satisfy the stringent requirements in the terms of processing delay and transmission overhead even when there exist a high amount of traffic demand.

2.1.2 Related works on securing vehicular cloud

A large amount of papers have been published related to the security and privacy issues in vehicular cloud. In this subsection, we briefly review some of the highly related works.

2.1.2.1 Related works on securing VANETs

IEEE 1609.2 Standard [23] is proposed for wireless access in vehicular environments (WAVE), i.e., protect security services for applications and management messages. The standard describes the methods to secure WAVE management and application messages, and defines the necessary administrative functions to support the core security functions. Moreover, the standard protects the messages from eavesdropping, spoofing, alternation and replay attacks, and it also preserves the privacy of vehicles from being identified by unauthorized parties.

Similar to other communication networks, security issues need to be addressed in vehicular networks. On one hand, an involved vehicle must be verified to prove its registration, and the exchanged message flows must be signed and integrity protected [24]. Due to the dynamically moving characteristics of the vehicles, it is challenging for the vehicles to conduct the mutual authentication and privacy preservation through the DSRC. The proposed scheme develops a group-based signature scheme [25], which achieves the function with which each group member can verify the other members and does not require the involvement of the identity. On the other hand, the privacy of a participating vehicle also needs to be preserved. More specifically, private information such as the vehicle's real identity and parameter setup could possibly be revealed to a third-party (e.g., other participating vehicles) [19]. Without privacy preservation, vehicles may be reluctant to get involved in the collaboration process. There exist various privacy preservation schemes for vehicular networks (e.g. preventing vehicles from revealing their identities or sensitive information to irrelevant entities) (see [26, 27]).

In vehicular clouds, an attacker could provide or access the same resources as their targets; thus, it is challenging to distinguish an attacker from a normal service provider or users [19, 28]. Thus, secure access control involves distinguishing malicious attackers from legitimate nodes by verifying the identities of the users, before allowing them to access the network [29]. There are existing research on ensuring integrity and providing identity verification in vehicular networks (see [25, 30, 31]), but these solutions are unlikely to be suitable for vehicular clouds, as the vehicular cloud is vulnerable to risks associated with both vehicular networks and cloud computing.

2.1.2.2 Related works on securing Mobile Cloud

In mobile cloud computing, besides resorting to the cloud computing and data storage for computationally intensive operations, mobile devices can also be treated as service nodes to collect information from the surrounding environment [32]. The sensed information, such as the location coordinates, information related to health situations, should be processed and maintained with security and privacy preservation. And a brief review about the secure data collection in mobile cloud computing is performed as follows.

A mobile cloud data processing framework is proposed in [32], which achieves the security and privacy protections through trust management and private data isolation. A simple and straightforward way to achieve the trustworthiness of the sensed data is to exploit a piece of trusted hardware or software for raw data signing [33]. However, in order to save the scarce transmission bandwidth, the mobile device sensed data reporting usually happens on the compressed sensor data, which leads to the unavailability of this approach [34]. Since mobile devices are typically contain a plethora of personally identifiable information, as a consequence, the privacy-sensitive data applications based on mobile sensors could violate the anonymity of the mobile devices. A privacy-aware architecture called AnonySense is proposed in [35], which is designed for the anonymous opportunistic crowdsourcing tasking and reporting. Meanwhile, the proposed scheme achieves the sensory data querying and receiving with privacy preservation. By exploiting the blind signature technique, an anonymity-preserving reputation scheme is proposed in [36]. In the proposed scheme, each user periodically update pseudonyms and links the reputation scores generated by different users, while prevents attackers from associating different pseudonyms. In addition, mobile devices are equipped with localization modules, which are able to determine the geographical positions with high accuracy, and could violate the location privacy of the device owner.

To achieve both anonymity and the trustworthiness of the sensory data, a novel mobile sensing provenance model is proposed in [37], which achieves the trust assessment of the sensory data and user anonymity. Meanwhile, a trust assessment algorithm is also designed, in

order to calculate the anonymous user reputation level. Two novel Min and k -th Min time-series data computation protocols among a group of mobile phone users are proposed in [38], which enables the data aggregator to compute the minimum and k -th minimum value with the XOR-homomorphic encryption consists of a probabilistic coding scheme and a cipher system. To promote the user participation in mobile sensing applications, two privacy-aware incentive mechanisms are proposed in [39], which pay credits towards the contributing users for the data they collected and prevent the dishonest users from abusing the system. A privacy-preserving system is proposed in [40], which the Internet-connected mobile users upload the collected sensor data towards the cloud server. Based on the collected information, the cloud server performs privacy-preserving collaborative learning with user-determined privacy levels.

In comparison with mobile cloud, there are some differences exist in vehicular cloud: 1) the vehicles suffer from high mobility, and the introduced communication and computation requirements in vehicular cloud are more stringent; 2) vehicles usually have enough energy supply and are equipped with enough processing capability. Thus, vehicular cloud is capable of performing more complex computation tasks.

2.1.3 Vehicular Data Dissemination

As an important application of vehicular cloud, vehicular data dissemination can be mainly classified into two categories based on the involved entities: V2I data dissemination and V2V data dissemination. For the V2I data dissemination, it mainly focuses on the traffic safety applications and traffic comfort applications [41]. For the traffic safety applications, the ultimate goal is to decrease the amount of traffic accidents or disseminate the traffic warning information when necessary through message broadcasting. The safety-related messages are usually disseminated in the push-based model, such that hazardous road conditions [42], emergency message broadcasting [43], and collision warning system [44]. While for the traffic comfort applications, they are mainly designed to improve the on-board experiences, such as available parking lots [45] and the faraway traffic congestion [46]. The pull-based model is often

exploited for on-demand traffic comfort applications with delay-tolerance and service delivery [41], in which a data requester sends a query towards a fixed site with data storage, and receives a data reply from the fixed data site. In [47], a pull-based data-downloading scheme is proposed, where a vehicle requests data from the RSU when it is under the coverage of the RSU, or from the neighboring vehicles when it is out of the coverage of the RSU. A pull-based data dissemination scheme is proposed in [48] based on a vehicular cloud service platform, which supports the vehicular user information query.

In V2V data dissemination, the vehicular data dissemination does not involve any fixed infrastructure and communicate among vehicles in the ad hoc mode. Since the V2V data dissemination is realized through the intermittently connected wireless links, the current works mainly focuses on the reliable data delivery with stringent delay requirement [49]. Due to the dynamically changing topology of vehicles and the transmission range of a single vehicle is limited, a V2V data dissemination scheme with vehicular caching scheme is proposed to release network congestion [50], which is particularly suitable for the sparsely distributed network with low transmission delay. An enhanced vehicular data dissemination at intersections is proposed in [51], in which a vehicle keeps the emergency message received at an intersection, and then disseminate the received message at the neighboring intersections, which achieves the benefit that the messages can be properly disseminated towards all the related road intersections. By exploiting the predictable vehicle mobility brought about by the constraint traffic patterns and fixed road layout, a carry-and-forward vehicular data delivery strategy with delay-tolerant is proposed in [52], where the vehicle maintains the received packet until a new vehicle enters the vicinity and forwards the packet. An efficient vehicular data dissemination scheme is proposed in [53], which enables the data aggregation and diffusion between vehicles within its transmission coverage. A simple and efficient adaptive data dissemination protocol is proposed in [54], which balances between the high packet delivery ratio and an acceptable transmission delay with bandwidth decrease.

The position-based data routing strategies are designed for VANET communications, which are particularly suitable for the highly dynamic network topology [55] and it is a combination of the

vehicle-to-infrastructure and vehicle-to-vehicle communications. For the position-based data routing, it can be divided into two categories: (improved) greedy forwarding and (predictive) directional greedy forwarding. The position-based data routing protocols are based on the geographical position of the destination derived from the GPS devices. Various positioning-based solutions are also proposed, such as [56, 57], however, none of the above works consider the problem of security issues.

Similar to VANETs, secure data processing is also a critical problem. For the pull-based VANETs data dissemination targeted at the traffic safety applications, it also faces the similar security challenges in VANETs. In pull-based VANETs data dissemination, the external attacks also include the eavesdropping, modifying, identity spoofing, repudiation and privacy disclosure. Even though the encryption techniques can protect the content and data integrity of the transmitted messages, the data dissemination process may still not be secure due to the multi-hop characteristic of VANETs. IEEE 1609.2 protocol [23] proposes a public key infrastructure (PKI) solution to protect the confidentiality and data integrity for the key establishment and data dissemination, and we briefly some related works in the following. A time-efficient and secure vehicular communications mechanism is proposed for the sequential message authentication in [58], which exploits the message authentication code based hash chains to achieve the vehicular messages authentication. A cooperative message authentication protocol is proposed in [59], which allows the recipient vehicles to identify the messages broadcasted by the malicious vehicles. Through cooperative message authentication, the proposed scheme achieves the goal of low complexity data verification. A secure clustering algorithm for vehicle-to-cloud authentication is designed in [60], in which a trust evaluation metric is also devised to calculate the trust level of vehicles and formulate the secure clusterings with the trust parameters. Meanwhile, due to the multi-hop characteristic of the vehicular data dissemination, insider attacks such as Sybil attack and false data injection, can be possibly launched to compromise one or a few vehicles, and further threat the whole network. However, for the pull-based data dissemination schemes, it even faces more security challenges. Since the data is disseminated in a request-response paradigm, the vehicle needs to deliver data query request towards the RSU,

which may threaten the location and identity privacy of the vehicles. On the other hand, during the data dissemination phase, the data content response, which may contain the sensitive information related to the vehicle's interest, and further violate the vehicle privacy.

However, the existing schemes related to vehicular data dissemination schemes are mainly designed for the dissemination strategies, and are less concerned about the security issues. Although there are security issues in the push-based vehicular dissemination schemes, these issues can be mostly solved with the existing solutions in VANETs. However, the existing security schemes in VANETs cannot solve the security issues met in the pull-based schemes, and new security strategies need to be designed to deal with the specific security threats for the pull-based vehicular data dissemination.

2.2 Secure Data Processing

In this section, we briefly review some of the works on secure and privacy-preserving data processing, in terms of the privacy-preserving data aggregation, the secure scalar product computation, and the secure location-based data queries, which are highly related to the rest of this thesis.

2.2.1 Privacy-preserving data aggregation

Due to the large volume of data generation in vehicular cloud, it is necessary for the RSU (or a coordinating vehicle) to aggregate the sensory data reports generated by the data uploading vehicles, such that the introduced computation and communication overhead can be largely reduced. There exists many privacy-preserving data aggregation schemes [61–65] in the era of Internet of Things (IoT), which can greatly save the scarce transmission bandwidth and the limited computation resource. Similar to other IoT applications, vehicular sensors can generate a plethora of information, and becomes an important component of the IoT architecture. In the rest of the subsection, a brief review of some privacy-preserving data aggregation schemes will

TABLE 2.1. Comparative Study of Secure Aggregation Schemes

Reference	Scenario	Security Feature
[66, 67]	Wireless Sensor Networks	Aggregation with Decryption
[61, 62, 68]	Wireless Sensor Networks	Aggregation without Decryption
[69–71]	Smart Grid	Privacy Preservation
[72, 73]	Smart Grid	Fault Tolerance
[63, 71, 74]	Smart Grid	Multi-dimensional
[64, 65, 75]	IoT	Differential Privacy

be performed. Meanwhile, a table is offered for the comparative study of the existing secure aggregation schemes, as shown in Table 2.1

Single-dimension data aggregation. Secure data aggregation has been extensively studied in wireless sensor networks (WSNs), which aims to combine and summarize data at intermediate nodes to reduce the data transmission overhead. During the data uploading phase, the data aggregator sensor node collects information from the neighboring nodes, aggregates the data from the neighboring nodes, and delivers the aggregated result towards the base station [66]. Earlier data aggregation schemes exploits the symmetric key cryptography for the plain data aggregation [67], in order to achieve security goal of data integrity. Since the sensory data may contain some sensitive information and the intermediate nodes can be easily compromised, the end-to-end data confidentiality is needed. With the symmetric key cryptography algorithms, it is almost impossible to achieve both end-to-end confidentiality and in-network data aggregation.

The homomorphic property enables the data computation on the encrypted data. With the homomorphic addition and multiplication cryptographic techniques, the intermediate sensors can perform both additive and multiplicative operations on encrypted data. A concealed data aggregation with privacy homomorphism is proposed in [68], which achieves the end-to-end data concealment and enables the data aggregation on ciphertexts. To reduce the communication overhead introduced by homomorphic cryptosystem, the intermediate nodes aggregate the encrypted data and upload the aggregation results towards the base station. A provably secure encryption scheme is proposed to enable the efficient additive aggregation of the encrypted data with the pseudorandom function [61]; meanwhile, the proposed scheme can also support the efficient statistical computation in terms of mean, variance, and standard deviation.

A privacy-preserving data aggregation scheme is proposed in [62], which enables an untrusted data aggregator to periodically compute the sum of all the participants' data without disclosing the content of each individual data report.

Besides WSN, data aggregation schemes have also been widely designed in IoT. A privacy-preserving dual-functional aggregation scheme is proposed in smart grid [69], which exploits the lattice cryptographic technique to support both the mean and variance aggregation. A novel secure data aggregation scheme is also proposed in [70], which achieves both privacy preservation and data integrity, and the proposed scheme also guarantees differential privacy and fault tolerance. In [71], a novel privacy-preserving set aggregation scheme is devised for smart grid, which allows more accurate data analytics in smart grid. In [76], a fault-tolerant smart-grid data aggregation scheme is proposed, which exploits the homomorphic Paillier encryption technique to obtain the aggregated data without the individual data disclosure. A privacy-preserving data aggregation with fault tolerance for smart grid communications is proposed in [72], and with the proposed scheme, the malfunctioning smart meters can be flexibly adjusted and the proposed scheme can also resist the differential attacks. A privacy-preserving verifiable data aggregation scheme is proposed in [73], which enables the data requester to verify the correctness of the retrieved results.

Multi-dimensional data aggregation. Since the data generated in IoT are multi-dimensional in nature, secure multi-dimensional data aggregation mechanisms have also been devised. In [63], Lu et.al proposed a privacy-preserving multi-dimensional data aggregation scheme in smart grids, which utilized the homomorphic Paillier cryptosystem to achieve data aggregation. The proposed scheme exploits a super-increasing sequence to compress the multi-dimension electricity usage data into one single-dimension, and performs the privacy-preserving data aggregation on the reformatted report. A anonymous multi-dimensional smart-grid data aggregation scheme is proposed in [74], which enables the operation center can calculate both additive and non-additive aggregation results. By exploiting the Boneh-Goh-Nissim homomorphic encryption [77], a novel multi-functional privacy-preserving data aggregation scheme for smart grid

communications is proposed in [65]. With the proposed scheme, the multiple statistical function computation is performed on collected data, which can also resist differential attacks. A data collection and query scheme is proposed in [75], in which the users' personal information are collected by the body sensors and converted into multi-dimension data; meanwhile, with the collected the multi-dimension data, the cloud server also provides query services. A lightweight privacy-preserving data aggregation scheme for the fog computing enhanced IoT environment is devised in [64]. The proposed scheme exploits the homomorphic Paillier encryption and the Chinese Remainder Theorem to structure the hybrid multi-dimensional data report into one composite data report; meanwhile, the proposed scheme also filters the injected false data at the network edge at an early stage.

However, the above schemes are based on the assumption that the number of data reports in data dimension is pre-defined, and they are not adaptive to situation that the number of sensory data reports generated in each dimension is fluctuating. In vehicular cloud, the data reports in each data dimension is fluctuating with respect to the spatial-temporal domain, and the number of data reports in each data dimension also needs to be counted.

2.2.2 Secure scalar product computation

The ultimate goal of the secure scalar product computation is that one entity (either the server or one participant) can obtain the scalar product of two or more participants. With the secure scalar product computation schemes, only the scalar product result can be learnt and the contents of the private vectors are protected. In some existing earlier works [78–80], the private scalar product computation are mainly designed for the purpose of privacy-preserving data mining. A frequent itemset mining protocol for vertically partitioned transaction database is proposed in [78], in which the transaction database is denoted as binary matrix. The proposed scheme enables different sites to collaborate, such that the global association rules can be identified with individual transaction data disclosure. In [79], a provably secure scalar product computation protocol is devised for privacy-preserving data mining, which enables two entities

to conduct the dot product computation of their scalars. In the proposed scheme, the homomorphic encryption technique is exploited to achieve the private scalar computation. An efficient privacy-preserving scalar product protocol of horizontally partitioned data mining is proposed in [80], which balances the knowledge discovery and privacy preservation.

Besides data mining, secure scalar product are also widely devised in privacy preserving personal profile matching, since some personal profiles may contain the sensitive information which are highly related to identity privacy. To perform the privacy-preserving profile matching, secure product computation schemes over two scalars are proposed in [81–86]. In [81], a novel privacy-preserving profile matching protocol for the similar attributes is devised, which leverages the secure product computation with the polynomial secret sharing. To identify the patients with similar symptoms, a secure same symptom handshake scheme is proposed in [82]. By exploiting the bilinear pairing technique, one patient can verify the other patient is with the same symptom. A fine-grained privacy-preserving profile matching scheme is proposed with the Paillier homomorphic Cryptosystem in [83], which enables two users to measure the level of similarity in their fine-grained personal files. In the proposed scheme, profile matching can be performed between two participating users without scalar content disclosure. An efficient privacy-preserving binary scalar product computation protocol with the lightweight non-homomorphic encryption is proposed in [84], which enables the user to identify the similarity in symptom characters without revealing users' symptoms. A privacy-preserving distributed profile matching protocol for mobile social network is proposed in [85], which exploits the elliptic curves based cryptography for homomorphic encryption construction. A privacy-preserving proximity test computation scheme is designed in [86], in order to identify the number of responding users in location-based services. The proposed scheme exploits the bilinear pairing technique to achieve the secure computation, while preserving the location privacy of the users.

In addition, secure scalar product computation are widely proposed in the era of cloud computing. In cloud computing, the cloud server executes the scalar computation, and the sensitive information personal contained may violate privacy and degrade service reputation. The proposed scheme enables the cloud to perform the dot product computation of their two entities'

scalars without privacy disclosure, and achieves the goals of neural network learning [87], similarity computation [88–90]. A secure multi-party computation algorithm for back-propagation neural network learning is proposed in [87] by exploiting the Boneh-Goh-Nissim homomorphic encryption technique, which allows the cloud server to execute secure scalar product and addition operations with privacy preservation in terms of the data content disclosure. In [88], a user privacy-preserving scalar production computation scheme is proposed, which combines the cosine similarity with some obfuscation mechanism to achieve the similarity computation. The proposed scheme denotes the users' data as vectors and performs geometrical transformation, which still keeps the angle between different scalars. In big data analysis, an efficient and privacy-preserving scalar product computation scheme is proposed in [89], which can also achieve the cosine similarity computation. The proposed scheme enables the scalar product computation without the involvement of the homomorphic encryption. To achieve both privacy preservation and high efficiency in big data analysis, a novel vector dot product computation scheme is proposed in [90]. By generating of a scalar noise for each individual scalar element, the proposed scheme achieves the binary and numerical data types with high efficiency and be adaptive to the high data volume.

However, in the above schemes, each element in each scalar needs to be encrypted separately, which leads to the result that the size of the ciphertexts is proportional to the length to the scalar. In addition, two corresponding elements needed to be computed individually, which results in the computation overhead be proportional to the length of the scalar. In vehicular cloud, since the computation and communication resource is limited, the elements in a given scalar is first structured to one-dimension and then encrypted, which is highly efficient in terms of communication and computation overhead.

2.2.3 Secure location-based data queries

Location privacy preservation is considered as one of the most important security requirements in vehicular networks. Without location privacy preservation, it is possible for the adversaries

to intercept messages in specific regions, and trace vehicles' moving patterns and physical locations through information analysis, such that the passenger's or driver's habits, routines, and personal associations can be disclosed [91]. To preserve location privacy, there mainly exist three categories of solutions: pseudonyms based schemes, location cloaking based schemes, and cryptography based schemes. In the following of this subsection, the location privacy preservation schemes are reviewed, and their pros and cons in secure location-based data queries are also analyzed.

The first category of solutions are through exploiting pseudonyms [92, 93]. [94] identifies the requirements of pseudonyms in terms of *time-limited*, *uniqueness*, *availability*, *pseudonym change block* and *link to other identifiers*. A lightweight conditional privacy-preservation protocol is proposed in [92], which enables the generation of the unlinkable pseudo-identities, and allows efficient revocation verification. The proposed scheme achieves the privacy preservation of vehicles, anonymous authentication of safety messages, and the real identity traceability by the authority. To prevent the long-term pseudonym tracking, a periodically pseudonym changing strategy at social spots is proposed in [93], such that the location privacy of the message broadcasting vehicles can be preserved. A privacy preservation pseudonym generation scheme is proposed in [95], which enables the vehicles to generate the pseudonyms with high computation efficiency, guarantees the desired level of anonymity, and allows vehicle revocation when necessary. To preserve the location privacy and ensure message integrity in VANETs, an anonymous batch authentication scheme with hash message authentication code (HMAC) is proposed in [96]. A mix-zone scheme to protect the location of vehicles is proposed in [97], in which an intersection located RSU manages mix-zone and the vehicles change their pseudo IDs and the corresponding public keys at intersections. With the proposed mix-zone scheme, an adversary cannot distinguish and link two pseudo-IDs generated by the same vehicle.

However, the pseudonym-based solutions are based on the idea of masking users' real identities, which is efficient in access control and authentication schemes [98]. In location-based services, the location information can still be disclosed [99]. Especially during the location-based data querying process, the pseudonym-based schemes may disclose the querying location, and

by associating the querying location with a particular individual, the real identity of the data requester can still be revealed [98].

The second category of solutions are through location cloaking [100–102], in which the identification of the exact location of each participant can be prevented by introducing uncertainty or error into the position information. *K-Anonymity* is a commonly used technique in location cloaking, which preserves the location privacy of one user through hiding one user among a group of K users [103]. In *K-Anonymity*, a user is considered private, if an attack based on the user cannot identify the target user with the probability of higher than $1/K$ among the other $K - 1$ users [98]. To improve the problem when there is little diversity in sensitive attributes of *K-Anonymity*, *l-diversity* is proposed by decreasing the data representation granularity [104]. To further improve the performance of the group based anonymity, *t-closeness* method is devised to improve *l-diversity* by additionally maintaining the data distribution of sensitive fields [105]. An efficient participatory sensing mechanism is devised with the locality-sensitive hashing (LSH) technique, which achieves both location privacy preservation and *K-Anonymity* [102]. A location privacy protection scheme is devised in [101], which enables the users to show their own location privacy preferences, and achieves a tradeoff among the location privacy preservation, data query accuracy and the user scalability.

For location-based services, a few schemes are devised based on location cloaking [106, 107]. A dummy-location selection algorithm is proposed in [106], which achieves the *k-anonymity* for users in location-based services. The proposed scheme exploits the entropy metric to choose dummy locations, in order to guarantee the widespread of the selected dummy locations. In [108], a fine-grained spatial cloaking scheme is proposed in location-based services. The proposed scheme exploits the modified Hilbert Curve to expand the cloaking region and achieve the *K-Anonymity*. Even though location cloaking can achieve a satisfactory level of location privacy preservation, it may not be able to guarantee that there are enough users in the neighborhood and not be able to achieve high accuracy in their query results [107]. Meanwhile, in location-cloaking based privacy-preserving data querying schemes, it is hard to difficult to balance the trade-off between the privacy preservation and the accuracy of the location-based

data query results [109]. In addition, a trusted third party is needed to store the original locations of the data, which is under the threat of the single point of attack [110].

The third type of solutions rely on the cryptographic techniques, which can effectively preserve the location privacy through encryption [111, 112]. One straightway to preserve location privacy is to encrypt the location information, since there is no location information disclosed. A location privacy preservation scheme is proposed in [113], which is based on the idea of dividing the location information into pieces and distributing the location information into numerous entities.

Various cryptography-based location privacy-preserving data query schemes have also been proposed in VANETs [99, 107, 110, 114–116]. An efficient spatial range query solution is proposed in [99], which permits data queries over encrypted location based data. In the propose scheme, a predicate-only encryption scheme for inner product range, which can detect whether a specific location position is within a given circular with location privacy-preservation. In [110], an efficient privacy-preserving location-based service system in the outsourced cloud is proposed, which preserves the privacy of the user query and the location-based data, by exploiting an improved homomorphic encryption technique over the composite order group. A fine-grained privacy-preserving location-based service (LBS) framework is proposed in [114], which achieves both location privacy and accurate location-base service results without the involvement of any trusted third party. The private information retrieval (PIR) technique is also exploited in [98], in order to enable the user to query the location-based database. The oblivious transfer technique is exploited in [115] to achieve the privacy-preserving location-based data query, in which the data server cannot learn the content of a users query, while the user cannot obtain more than they are entitled [117]. The authors in [107] propose a hybrid approach based on location cloaking and the additive homomorphic encryption, which achieves both efficiency and privacy preservation. A spatial range query scheme over ciphertext is also proposed in [116], which achieves the location privacy of the user’s query and the location-based service confidentiality.

For the cryptography-based location privacy preservation schemes, heavy computation and

communication overhead can be introduced, and influence the performance of the whole system, especially for the vehicular cloud scenario with stringent delay requirement and limited transmission bandwidth. In addition, the above cryptography-based location-based data querying schemes are mainly designed for an outsourced cloud environment with fixed data storage and fixed data access structure, which cannot be directly applied to our vehicular cloud scenario with distributed storage. Thus, based on the above analysis, the challenge becomes one to design mechanisms to efficiently protect user privacy with high location based data querying accuracy.

2.3 Secure Key Management in LTE-A networks

In this section, we briefly review the extended key architecture and evolved packet system - authentication and key agreement (EPS-AKA) process, and the EPS handover key management process.

2.3.1 EPS-AKA process

The main purpose of the EPS-AKA session key generation mechanism is to mutually authenticate the user equipment (UE) and the MME on behalf of the HSS, and establish the session keys between a UE and the MME and that between a UE and the eNB, as shown in Fig. 2.1.

The LTE system exploits the EPS-AKA process to achieve the mutual authentication between the UE and the EPC, as shown in Fig. 2.1. The MME which on behalf of the HSS, delivers a random challenge $RAND$ and an authentication token $AUTN$, and KSI_{ASME} (utilized for session key identification) towards the requesting UE, in order to achieve the goal of network authentication. After receiving $(RAND, AUTN, KSI_{ASME})$ the UE verifies $AUTN$, and calculates a response RES if it is successfully verified. Then the UE calculates the intermediate session keys CK, IK, K_{ASME} with $RAND$, the permanent master key K . After receiving RES , the MME checks the value of RES to verify the authenticity of the requesting UE.

FIGURE 2.1. EPS-AKA procedure [1]

2.3.1.1 Extended key architecture

FIGURE 2.2. Extended key hierarchy in the EPS [1]

As shown in Fig. 2.2, during the registration of a UE, the EPS-AKA session key generation procedure occurs between a UE, the connecting eNB, the MME, and the HSS.

Upon mutual authentication, based on the permanent master key K (stored on the Universal Subscriber Identity Module (USIM) and the AuC), the UE and the MME generate the first intermediate key K_{ASME} after binding to the serving network identity (SN_{ID}), which is valid until the EPS-AKA procedure. The MME acts as the access security management entity (ASME), which protects the access security and functions as a key generator and distributor within the

EPS security architecture. Based on K_{ASME} , the MME derives three keys K_{NASenc} , K_{NASint} and K_{eNB} , where K_{NASenc} and K_{NASint} are two transient keys for encryption and integrity checks in the non-access Stratum (NAS), and a specific session key K_{eNB} shared between a UE and a unique eNB. K_{eNB} is designed for the AS protection, and it needs to be updated when the UE attaches to a different eNB. Given K_{eNB} , three transient keys K_{RRCenc} , K_{RRCint} and K_{UPenc} are also derived, where K_{RRCenc} and K_{RRCint} are for the encryption and integrity protection of radio resource control (RRC) signaling traffic in the access stratum (AS), and K_{UPenc} is exploited to encrypt the user-plane data traffic in the AS.

2.3.2 Review of EPS handover key management

We make a brief introduction of the EPS handover key management mechanism, which was originally presented in 3GPP Release 11 for LTE networks [118]. During an intra-MME handover process, it is the task of the current source eNB to provide the next hop session keys shared between the target eNB and the roaming UEs, and the session key transfer is realized by the secure X2 link established between eNBs [1]. To ensure backward key separation, the source eNB generates the next-hop session key K_{eNB}^* for the target eNB based on the current session key K_{eNB} by applying a one-way hash function, which is also called the key derivation function (KDF):

FIGURE 2.3. Message Flow during Intra-MME Handover

$$K_{eNB}^* = KDF(K_{eNB}, \alpha) \quad (2.1)$$

where α denotes the physical parameter related to the target eNB, and the above key update process is also denoted as the horizontal key derivation. However, after several horizontal key derivations, a set of session keys are linked to each other, which is known as the handover key chaining, and the current session key could be derived from the previous eNBs. To guarantee forward key separation, the vertical key derivation is introduced. After each intra-MME handover process, the MME provides fresh keying material NH_{NCC} to the target eNB for the next hop session key generation:

$$K_{eNB}^* = KDF(NH_{NCC}, \alpha) \quad (2.2)$$

where NH_{NCC} represents that the next hop (NH) key has been updated next-hop chaining counter (NCC) times, and the current security association between the target eNB and UE is based on the value of NCC value. And the value of NH_{NCC} and the initial value NH_0 are generated as follows

$$\begin{aligned} NH_{NCC} &= KDF(K_{ASME}, NH_{NCC-1}) \\ NH_0 &= KDF(K_{ASME}, K_{eNB}) \end{aligned} \quad (2.3)$$

where K_{ASME} is the first intermediate key used to protect the security of the NAS layer. Thus, a two-hop based forward key separation is maintained, that is, the source eNB cannot derive the following session keys after two handovers. Fig. 2.3 depicts the message flow during an intra-MME handover process, by implementing the vertical key derivation.

In the proposed handover key management scheme, a desynchronization attack may happen to disrupt the update of the NCC value in *message (2)* or after *message (6)*. To be more specific, if an eNB which is compromised by an adversary through physical, host, and network protocol vulnerabilities, when a UE is connected to the compromised eNB, the compromised eNB can disrupt the update of the NCC value, which results in the desynchronization of the target eNB. The desynchronization attack leads the target eNB to give up the forward key separation, and only able to perform the horizontal key derivation, in which a set of session keys can be linked

to each other. To reduce the loss brought about by the desynchronization attacks, root key update is an efficient way to prevent further session key exposure [1]. However, the frequent root key update involves the transmission of the UE's identity, and it introduces heavy traffic and computation load to the MME in the core network, especially in this MRN scenario with the involvement of tens of on-board UEs.

2.3.3 Mobile relay networks and Coordinated Multi-Point

In this subsection, we briefly review some techniques in LTE-A networks, which is highly related to the following chapters.

2.3.3.1 Mobile relay networks

For public transportations, due to the fast moving well-shielded carriage, data transmission will suffer from high penetration path loss, severe Doppler frequency shift, and low handover success rate caused by a large number of on-board mobile devices performing frequent handovers simultaneously. To circumvent the above problems, the concept of mobile relay is proposed in the 3GPP in LTE-A networks [119]. A mobile relay consists of an outer antenna mounted on the top of the moving transportation providing a wireless backhaul with a Donor evolved Node B (DeNB) located along the roadside, while the wireless connection to the on-board users is realized by the inner antenna installed inside the carriage [120, 121]. As such, the Doppler effect in this network scenario can be mitigated by referring to the speed and track information [122].

The study of the mobile relay node (MRN) has attracted great attention in LTE-Advanced networks, and we make a brief review on some current directions of the MRN related to our work [120] [123] [124]. One direction mainly concentrates on the comparison between architectures for the mobile relays. In [120], the authors investigate the network architecture of the mobile relays, that is, the packet data-gateway and the serving-gateway selection during mobile relay handover processes. The other direction is related to the mobility management scheme, that is, group handover [123] [124]. The group handover is investigated under the scenario that

when on-board UEs are travelling in a public transportation, they will be stay in close distance with each other and perform handovers simultaneously which may lead to the network congestion and blocked handover [125] [126]. The group handover strategy treats the on-board UEs under the coverage of one MRN as a group, so that the signalling overhead and handover delay can be reduced and it can also decrease the rate of radio link failure. The authors in [127] proposed a novel eNodeB-to-eNodeB handover scheme for high speed moving vehicular femtocell networks, by utilizing an outside transceiver to perform handover, the connections between outside transceivers and eNodeBs are maintained seamlessly, and the outage probability is reduced compared to conventional handover scheme. However, very limited efforts have been made on the issues of security and key management scheme in mobile relay networks [128].

2.3.3.2 CoMP

CoMP transmission is mainly designed to combat the inter-cell inference (ICI) in cellular wireless networks, whose basic idea is to aggregate and coordinate neighboring eNBs, such that the interference at the UEs can be effectively mitigated [129] and the overall system performance can be enhanced. CoMP is mainly designed to improve the service quality at cell edges, a UE at the cell edge is able to be served by multiple eNBs, in order to improve the signal quality under cell edge conditions. There are two major categories of approaches in CoMP: coordinated scheduling or beamforming (CS/CB), in which each UE is only connected to one eNB; and joint processing (JP), in which each UE can be connected to multiple eNBs and transmitted simultaneously. A number of researches have been performed in CoMP. In [130], a decentralized scheme is proposed in both CB and JP approaches, which minimizes the sum of the transmission power with given QoS requirements. In [131], various uplink CoMP schemes with constraint backhaul and imperfect channel state information (CSI) are analyzed.

For the high-speed mobility scenario, vehicles usually suffer from frequent handovers, which result in the consequence of constantly experiencing high inter-cell interference and low network performance at cell edges [132]. To promote the handover performance, CoMP-based

joint processing handover schemes are proposed in LTE-A networks [132, 133], which provides simultaneous multiple transmissions for each UE, so that the received signal strength and the received signal spatial interference can be increased and reduced respectively.

Chapter 3

Achieve location privacy-preserving range query in vehicular sensing

A novel range query scheme is proposed in this chapter, which helps the data requester to accurately retrieve the sensed data from the distributed on-board storage in VANETs with location privacy preservation. The proposed scheme exploits structured scalars to denote the locations of data requesters and vehicles, and achieves the privacy-preserving location matching with the homomorphic Paillier cryptosystem technique. Detailed security analysis shows that the proposed range query scheme can successfully preserve the location privacy of the involved data requesters and vehicles, and protect the confidentiality of the sensed data. In addition, performance evaluations are conducted to show the efficiency of the proposed scheme, in terms of computation delay and communication overhead. Specifically, the computation delay and communication overhead do not depend on the length of the scalar, and it is only proportional to the number of vehicles.

3.1 Introduction

Nowadays, the fast development of automotive industry and the wide deployment of on-board sensors have created a huge opportunity of vehicular sensing [134]. Other than protecting the

normal operations of vehicles, the data generated by on-board sensors (e.g., chemical spill detectors, vibration sensors, and acoustic detectors [135]) can also provide unprecedented spatial-temporal coverage and witness unpredictable incidents with virtually zero investment in the deployment and maintenance of fixed surveillance infrastructures [136]. Meanwhile, since modern vehicles are normally not constrained by energy supply and equipped with adequate on-board storage (up to terabytes), they can keep the continuously harvested data in their on-board storage, which avoids the network congestion caused by the sensed data uploading. Motivated by the practical profits of vehicular sensing, various vehicular sensing based applications have appeared. For example, MobEyes [135] proposes a proactive urban monitoring application with the data collected from cameras and chemical detection sensors, a road surface condition detection application [137] is devised by opportunistically gathering data from vibration and GPS sensors. Meanwhile, to attract sufficient participants to join in the sensing process, an incentive mechanism for mobile crowdsensing has been devised in [138].

However, the wide application of vehicular sensing has met several challenges [7, 139, 140]. The first challenge is related to the location privacy of data requesters and data uploading vehicles. To retrieve the wanted data from the distributed on-board storage, each data query should specify the corresponding time and location. However, if the location information of a data query is disclosed, it may bring social reputation or economic damage to the querying location. Meanwhile, since vehicles are dynamically moving, their data reports should also be spatial-temporal tagged. The location information of a vehicle can be correlated with certain personal affairs (such as churches and hospitals), or it may reveal the identities of the vehicle owners (such as residences and offices). Without location privacy preservation, data requesters and vehicles are reluctant to issue data queries and upload sensed data, which leads to the under utilization of vehicular sensing. Thus, the location privacy of data requesters and vehicles should be preserved.

The second challenge is related to the accuracy of data query results. Due to the sheer volume of on-board data generation and limited transmission bandwidth, it is impossible for vehicles to upload all the sensed data immediately, and some of the sensed data are maintained in their

on-board storage (except those real-time data for applications with stringent delay requirements) [141]. Since vehicles are dynamically and opportunistically moving, the sensory data maintained in their on-board storage captured during a past time period may or may not be generated within the target query area. Moreover, it is highly possible that the queried sensory data are partially generated within the target query area. Thus, it is difficult for the data requester to accurately identify and acquire the wanted data from the on-board storage of the massive and dynamically moving vehicles. However, most of the current secure range query schemes are devised for the outsourced central cloud storage [99, 142], which cannot be directly applied to the distributive vehicular storage. Thus, a secure and accurate range query scheme is needed for the distributed on-board storage scenario to fully exploit the potential of vehicular sensing.

In this chapter, to overcome the above challenges, we propose a novel privacy-preserving range query scheme from the distributed on-board storage in vehicular ad hoc networks (VANETs) in a practical scenario: acquiring the data harvested by the on-board air pollution sensors within the defined query area, i.e., an industrial area, to monitor the air quality. The proposed scheme exploits the Paillier cryptosystem [143] to preserve the location privacy of data requesters and vehicles, and protect the confidentiality of the sensed data. Specifically, the contributions of this chapter are threefold.

First, the proposed scheme structures each multi-dimension scalar in one dimension, where the multi-dimension scalars denote the positions of the data requesters and vehicles in the format of grid cells; meanwhile, the proposed scheme supports secure scalar product computation for location matching. Thus, the location-based data query and data reports can be transmitted to the data server with high efficiency; meanwhile, the data harvested within the target query area can still be identified with location privacy preservation.

Second, since the vehicles are dynamically moving, the average of the sensory data captured during a short-time period may contain both the sensory data generated within and outside the query area. The proposed scheme enables the identification of the number of data reports generated within the target query area, and guarantees the accuracy of the sensory data captured within the target area.

Third, we give detailed security analysis to show that the proposed scheme is secure under the defined security model and achieves the security requirements in terms of location privacy preservation and confidentiality. Meanwhile, we conduct comparative performance evaluation to show that the proposed scheme is more efficient than the existing homomorphic secure scalar product schemes in terms of computational complexity and communication overhead.

The remainder of this chapter is organized as follows. We describe the system model, the security requirements, and the design goals in Section 3.2. We recall the preliminaries and propose our privacy-preserving range query scheme in Section 3.3, followed by our security analysis and performance evaluations in Sections 3.4 and 3.5, respectively. Finally, we conclude our work in Section 3.6.

3.2 System Model, Security Requirement and Design Goal

3.2.1 System Model

In the system model, we present the proposed range query scheme in a practical vehicular sensing application: help the data requester to acquire the data generated by vehicles (data harvested by the on-board air pollution sensors [144]) located within the target query area (an industrial district) during a past short-time period, which can help to investigate the air quality of the given industrial district. The system model consists of five entities, as shown in Fig. 3.1.

- The **data requester** aims to gather the air quality data in a given industrial district from the on-board air pollution sensors. The data requester submits a location-based data query ((1) *Data Query*) towards the data server and waits to hear the reply, as shown in Fig. 3.1.
- When the **data server** receives a data query request from a data requester, it forwards the received data query request ((2) *Data Query*) to all the vehicles through road side units (RSUs). Meanwhile, the data server receives all the on-board sensed data reports from

FIGURE 3.1. Data query architecture.

RSUs, performs data filtering according to the data query, and delivers the data query response towards the data requester((6) *Data Query Result*), as shown in Fig. 3.1.

- Each **RSU** serves as a gateway between the data server and data uploading vehicles, it helps to broadcast received data queries towards all the vehicles under its coverage ((3) *Data Query*) and forward the received on-board sensed data reports towards the data server ((5) *Data Report*), as shown in Fig. 3.1.
- Each **data uploading vehicle** is equipped with the required on-board sensor and enough on-board storage. When a piece of data is harvested by the on-board sensor of a vehicle, it is maintained in the on-board storage with time and location tagged. When a vehicle receives a data query, it first checks whether the queried data is still maintained. If the vehicle owns the queried data, it uploads the location-based sensed data to the nearest RSU((4) *Data Report*), as shown in Fig. 3.1.
- The **trusted authority** is a trusted and powerful entity, which is mainly responsible for

the system bootstrap, key materials management, and the registration of new data requesters and vehicles, as shown in Fig. 3.1.

The wireless connection between a vehicle and an RSU is realized through the IEEE 802.11p standard, a short- to medium-range communication technology operating at 5.85–5.925 GHz band with 3–27 Mbps data rates [145], which is mainly designed for the intelligent transportation systems radio service. The connections between RSUs and the data server, and those between data requesters and the data server, are realized through either wired links or any other links with high bandwidth and low transmission delay.

3.2.2 Security Requirements

In the security model, we consider the trusted authority is fully trusted, while the data server and RSUs are assumed to be honest-but-curious, that is, they follow the protocols, but they may try to infer the generation location and the content of data queries and data reports, which may violate the location privacy and confidentiality. Therefore, to preserve the location privacy of data requesters and vehicles, and protect the confidentiality of each individual sensed data report, the following security requirements should be satisfied:

Location Privacy. Protect the location privacy indicates that the location information contained in each data query and data report should be protected [139], and the location information in this context indicates the grid cell scalars, which the data query and data generation locations are mapped to. Even if the data server obtains all the possible data queries and data reports, it cannot identify the grid cell scalar contained in any location-based data query or data report. Location privacy protection also includes that the data uploading vehicles cannot learn the grid cell scalar of the data query, or vice versa. In this way, the location privacy of data requesters and data uploading vehicles can be preserved.

Confidentiality. Protect the confidentiality of individual sensed data report means that, even if the data server obtains all the possible on-board sensed data reports, it cannot recover the

content of an individual sensed data report [63]. Thus, the individual on-board sensed data can achieve the security requirement of confidentiality .

Note that there may exist other types of attacks such as impersonation, eavesdropping, and violation of data integrity [146]. Since we mainly focus on the privacy-preserving range query from the distributed on-board storage in VANETs, these security threats are beyond our study scope. We also assume that there is no collusion attack (any format of collusion among the vehicles, RSUs, and server by disclosing their private keys towards any other entity) in the system, which is in accordance with previous research on secure data query [110].

3.2.3 Design Goals

Under the aforementioned system model and security requirements, our design goal is to develop a privacy-preserving range query scheme from the distributed on-board storage in VANETs. Specifically, the following three objectives should be achieved.

According to the above statement, if the proposed range query scheme does not take security into consideration, the location privacy of data requesters and data uploading vehicles could be threatened. Then, data requesters and vehicles may not be willing to participate in the range query process, and the system cannot properly operate. Therefore, the proposed scheme should achieve the security goal of location privacy.

To accurately reflect the air quality, it is important for data requesters to accurately obtain the data harvested by vehicles in the given query area. Since vehicles are dynamically moving and their positions are opportunistically changing, each received sensed data report may or may not be able to reflect air quality. Thus, to achieve the goal of the high accuracy in air quality monitoring, the data server should accurately filter all the received data reports.

Since vehicles are featured with the fast-moving characteristic, the connections between RSUs and vehicles are relatively short and intermittent, and the communication overhead introduced by the data query and sensed data reports should also be minimized. To extract the desired data

TABLE 3.1. Notations of the Proposed Range Querying Scheme

Notation	Definition
p_1, q_1	Parameters of Paillier Cryptosystem
$pk = (n, g)$	Public Key of Paillier Cryptosystem
$sk = (\lambda, \mu)$	Secret Key of Paillier Cryptosystem
$\vec{u}$	Grid Cell Scalar
(x_0, y_0)	Bottom-left Vertex
α	Secret Value Shared between Data Server and Vehicles

from the massive amount of sensory data reports efficiently, the computational complexities brought to data server should also be deliberately evaluated.

3.3 Proposed Privacy-preserving Range Query Scheme

In this section, we propose the privacy-preserving range query scheme from the distributed on-board storage in VANETs, which mainly consists of five parts: preliminaries, system initialization, data query generation, data report generation, and data filtering. Meanwhile, the definitions of some important notations are described in Table 3.1.

3.3.1 Preliminaries

In our proposed scheme, the Paillier cryptosystem is exploited due to its additive homomorphic property and the homomorphic multiplication property of one plaintext and one ciphertext [143], which has been widely employed in many privacy-preserving data processing applications [63]. Specifically, the Paillier cryptosystem consists of three components: key generation, encryption, and decryption.

- *Key Generation:* Given the security parameter κ , two large positive prime numbers p_1, q_1 are first chosen, where $|p_1| = |q_1| = \kappa$. Then, the RSA modulus $n = p_1 q_1$ and $\lambda = \text{lcm}(p_1 - 1, q_1 - 1)$ are computed. Define a function $L(u) = \frac{u-1}{n}$, after choosing a generator $g \in \mathbb{Z}_{n^2}^*$, $\mu = (L(g^\lambda \bmod n^2))^{-1} \bmod n$ is further calculated. Then, the public key is $pk = (n, g)$, and the corresponding private key is $sk = (\lambda, \mu)$.

FIGURE 3.2. Registration and secret sharing among the data requester and vehicles.

- *Encryption*: Given a message $m \in \mathbb{Z}_n$, choose a random number $r \in \mathbb{Z}_n^*$, and the ciphertext can be calculated as $c = E(m) = g^m \cdot r^n \bmod n^2$.
- *Decryption*: Given the ciphertext $c \in \mathbb{Z}_{n^2}^*$, the corresponding message can be recovered as $m = D(c) = L(c^\lambda \bmod n^2) \cdot \mu \bmod n$.

3.3.2 System Initialization

For the range query system under consideration, we assume a trusted authority, located at the management authority of vehicles and traffics, will bootstrap the whole system.

Given a security parameter κ , the trusted authority selects two large prime numbers p_1 and q_1 , where $|p_1| = |q_1| = \kappa$. The trusted authority also calculates the Paillier cryptosystem's public key ($n = p_1 q_1, g$), and the corresponding private key (λ, μ) . Moreover, the trusted authority chooses a secure cryptographic hash functions $H()$, where $H : \{0, 1\}^* \rightarrow \mathbb{Z}_n^*$. In addition, the trusted authority also chooses a random number α . The trusted authority publishes the system parameter as $params = \{n, g, H()\}$.

The trusted authority assigns the private key (λ, μ) towards the data server, but the trusted authority does not share α with the data server. During the registration of each data uploading vehicle (w vehicles in total) or each data requester, the trusted authority checks its eligibility (whether the vehicle has the on-board air pollution sensors installed) and securely returns α towards it, as shown in Fig. 3.2.

FIGURE 3.3. Defined grid cell architecture/locations of vehicles and the data query area in grid cells.

3.3.3 On-Board Data Query Generation

A data requester, the environmental protection agency in this context, aims to monitor the air quality of an industrial district through collecting the data generated by the air pollution sensor (*Data_Type*), during a past time period (from the starting time t_s to the ending time t_e). We exploit the grid cell system defined in [147] to construct a grid cell architecture, which reflects the locations of the data requester and vehicles: the data requester also specifies a large region contains the target industrial district, which is large enough for the data requester and vehicles to be comfortable with revealing the fact that they are somewhere within this query area, and the bottom-left vertex of the large region is with the bottom-left vertex (x_0, y_0) . The large region is constructed into a grid cell structure, which is divided into $k = a \times b$ equal-sized grid cells with the side length of L , i.e., $L = 0.8$ km, when we take an industrial area with the coverage of 2.4 km^2 as an example [148]. Given location L_i with coordinates (x_i, y_i) , the corresponding grid cell identifier can be computed as $\lceil \frac{x_i - x_0}{L} \rceil + \lfloor \frac{y_i - y_0}{L} \rfloor \cdot a$. As shown in Fig. 3.3, the data query area and the moving trajectory of vehicles are denoted with grid cells.

The data requester generates a query area that includes a group of grid cells $\mathcal{I}_d$, which is overlapped with the target industrial district. Meanwhile, the data requester generates a grid cell scalar $\vec{u} = \{u_1, \dots, u_k\}$ to denote the defined grid cell architecture shown in Figure 3.3, which is initially set to zero. If $i \in \mathcal{I}_d$, i.e., grid cell i belongs to the target industrial district, we set the

corresponding position in the scalar $\vec{u}$ to $u_i = 1$. Moreover, the data requester selects a number β , which satisfies the condition that $2 \cdot k < 2^\beta$ and $2^{(2k+1)\beta} < n$. In addition, we structure the multi-dimension grid cell scalar to one composite value by calculating $m_u = \sum_{i=1}^k u_i \cdot 2^{i \cdot \beta}$, and generate the ciphertext of m_u , which is

$$\begin{cases} C_{u,1} = g^{m_u + H(\alpha||1||TS)} \cdot r_{u,1}^n \bmod n^2, \\ C_{u,2} = g^{m_u \cdot H(\alpha||1||TS)^{-1} + H(\alpha||2||TS)} \cdot r_{u,2}^n \bmod n^2, \end{cases} \quad (3.1)$$

where TS is the current timestamp, and $r_{u,1}, r_{u,2} \in \mathbb{Z}_n^*$ are two random numbers.

Finally, the data requester formulates a data request and delivers it to the data server:

$$Request \leftarrow \langle Data_Type || x_0 || y_0 || a || b || L || \beta || t_s || t_e || \Delta t || C_{u,1} || C_{u,2} || TS \rangle, \quad (3.2)$$

where Δt is the on-board sensed data sampling interval.

After receiving *Request*, the data server further delivers the data request

$$Request_1 \leftarrow \langle Data_Type || x_0 || y_0 || a || b || L || \beta || t_s || t_e || \Delta t || TS \rangle \quad (3.3)$$

towards the data uploading vehicles through RSUs.

3.3.4 Data Report Generation

For each registered data uploading vehicle, the previously sensed data (on the scale of a few past days) and the related information (such as number, time, data type, sensed data, location coordinates, etc.) are maintained in its on-board storage in the format of $\langle n, t, Data_Type, data, (x, y) \rangle$.

Upon receiving *Request₁*, it checks its on-board storage to identify the data in consistent with the data query. If the queried data exists, the vehicle performs the following steps.

The vehicle identifies the on-board air pollution sensor data ($data_1, data_2, \dots, data_d$) harvested at the past time points $(t_s, t_s + \Delta t, \dots, t_e)$, where $d = \frac{t_e - t_s}{\Delta t} + 1$, $d < k$, Δt is the sampling

interval. Then the vehicle calculates the sum of the sensed data reports, which is denoted as $sum_v = \sum_{i=1}^d data_i$ and it satisfies the condition that $\sum_{v=1}^w sum_v < n$, where w denotes the total number of data uploading vehicles. In addition, the vehicle generates the ciphertext of the on-board sensed data sum_v :

$$C_{v,0} = g^{sum_v + H(\alpha||0||TS)} \cdot r_{v,0}^n \bmod n^2, \quad (3.4)$$

where $r_{v,0} \in \mathbb{Z}_n^*$ is a random number.

Then, the vehicle maps the location coordinates of the on-board sensed data into grid cell architecture shown in Figure 3.3, and generates a vector $\vec{v} = \{v_1, \dots, v_k\}$ with Algorithm 1.

Algorithm 1: Vehicle Scalar Generation

Data: Given location coordinates $\{(x_1, y_1), (x_2, y_2), \dots, (x_d, y_d)\}$ and sets $\vec{v} = \{v_1, \dots, v_k\} = \vec{0}$.

Compute:

- 1: **for** $i = 1 : d$ **do**
- 2: Given location coordinates $L_i(x_i, y_i)$, the vehicle calculates $j = \lceil \frac{x_i - x_0}{L} \rceil + \lfloor \frac{y_i - y_0}{L} \rfloor \cdot a$;
- 3: **if** $j \in \{1, 2, \dots, k\}$ **then**
- 4: Calculates $v_j = v_j + 1$;
- 5: **end if**
- 6: **end for**

Output: $\vec{v}$

Then, the vehicle structures the multi-dimensional scalar $\vec{v}$ into one dimension : $m_v = \sum_{j=1}^k v_j \cdot 2^{(k+1-j) \cdot \beta}$, and generates the ciphertext

$$\begin{cases} C_{v,1} = g^{m_v + H(\alpha||1||TS)^{-1}} \cdot r_{v,1}^n \bmod n^2, \\ C_{v,2} = g^{m_v \cdot H(\alpha||1||TS) - H(\alpha||2||TS) - \eta_v} \cdot r_{v,2}^n \bmod n^2, \end{cases} \quad (3.5)$$

where $r_{v,1}, r_{v,2} \in \mathbb{Z}_n^*$ are two random numbers, and $\eta_v = \eta_{v,1} \cdot 2^\beta + \dots + \eta_{v,k} \cdot 2^{k \cdot \beta} + \eta_{v,k+1} \cdot 2^{(k+2) \cdot \beta}$, such that $\eta_{v,i}$ is a random number that satisfies the condition: $\eta_{v,i} + k < 2^\beta$ when $i \in \{1, \dots, k\}$ and $\eta_{v,k+1} \in [1, k \cdot 2^{(k-2) \cdot \beta})$.

Finally, the vehicle formulates the *Data_Report* and delivers it to its nearest RSU:

$$Data_Report \leftarrow \langle C_{v,0} \| C_{v,1} \| C_{v,2} \rangle. \quad (3.6)$$

3.3.5 Privacy Preserving Data Filtering

Upon receiving all the data reports, each RSU forwards all the collected on-board sensed data reports to the data server, and the data server filters within all the received data reports to identify the on-board sensed data reports generated within the industrial district.

The data server first computes the scalar product of $\vec{u}$ and $\vec{v}$ to identify the data reports generated within the industrial district, based on the idea that the data uploading vehicle must share at least one grid cell with the industrial district. Then, the data server decrypts $C_{u,1}$ with its private key sk and recovers the value of $m_u + H(\alpha || 1 || TS)$. Meanwhile, the data server also computes

$$\begin{aligned} C_{u,v} &= \frac{(C_{v,1})^{m_u + H(\alpha || 1 || TS)}}{C_{u,2} \cdot C_{v,2}} \bmod n^2 \\ &= \frac{(g^{m_v + H(\alpha || 1 || TS)^{-1}} \cdot r_{v,1}^n)^{m_u + H(\alpha || 1 || TS)}}{g^{m_u H(\alpha || 1 || TS)^{-1} + m_v H(\alpha || 1 || TS) - \eta_v} \cdot r_{u,2}^n \cdot r_{v,2}^n} \bmod n^2 \\ &= g^{m_u m_v + \eta_v + 1} \cdot \left(\frac{r_{v,1}^{m_u + H(\alpha || 1 || TS)}}{r_{u,2} \cdot r_{v,2}} \right)^n \bmod n^2. \end{aligned} \quad (3.7)$$

Based on $C_{u,v}$, the data server recovers the value of $s_{u,v} = m_u m_v + \eta_v + 1$ with the private key sk . With $s_{u,v}$, the data server can calculate the scalar product of $\vec{u} \cdot \vec{v}$,

$$k_v = \vec{u} \cdot \vec{v} = \frac{(s_{u,v} \bmod 2^{(k+2) \cdot \beta}) - (s_{u,v} \bmod 2^{(k+1) \cdot \beta})}{2^{(k+1) \cdot \beta}}, \quad (3.8)$$

where k_v indicates the pieces of sensed data generated within the industrial district. If $k_v = \vec{u} \cdot \vec{v} \geq 1$, at least one piece of sensed data is generated in industrial district, and the vehicle becomes a member of the group $\mathcal{N}_{in}$; otherwise, all the data reports are generated out of the industrial district, and the vehicle becomes a member of the group $\mathcal{N}_{out}$, such that $|\mathcal{N}_{in}| + |\mathcal{N}_{out}| \leq w$. Finally, the data server calculates $K_{in} = \sum_{v \in \mathcal{N}_{in}} k_v$.

Correctness. The correctness of Eq. (3.8) can be clearly illustrated with the following equation:

$$\begin{aligned}
s_{u,v} &= m_u m_v + \eta_v + 1 \\
&= u_1 v_1 \cdot 2^{(k+1)\beta} + u_2 v_1 \cdot 2^{(k+2)\beta} + \dots + u_k v_1 \cdot 2^{2k\beta} \\
&\quad + u_1 v_2 \cdot 2^{k\beta} + u_2 v_2 \cdot 2^{(k+1)\beta} + \dots + u_k v_2 \cdot 2^{(2k-1)\beta} \\
&\quad + \dots \\
&\quad + u_1 v_k \cdot 2^{2\beta} + u_2 v_k \cdot 2^{3\beta} + \dots + u_k v_k \cdot 2^{(k+1)\beta} \\
&\quad + \eta_{v,k+1} 2^{(k+2)\beta} + \eta_{v,k} 2^{k\beta} + \dots + \eta_{v,1} 2^\beta + 1 \\
&= \left(\sum_{i=1, i < j}^k u_i v_j \cdot 2^{(k+1+i-j)\beta} + \eta_{v,k} 2^{k\beta} + \dots + \eta_{v,1} 2^\beta + 1 \right) \\
&\quad + \left(\sum_{i=1, i > j}^k u_i v_j \cdot 2^{(k+1+i-j)\beta} + \eta_{v,k+1} 2^{(k+2)\beta} \right) \\
&\quad + \left(\sum_{i=1, i=j}^k u_i v_j \cdot 2^{(k+1)\beta} \right).
\end{aligned} \tag{3.9}$$

As shown in Equation (3.9), $s_{u,v}$ consists of three parts: (i) when $i < j$, $(\sum_{i=1, i < j}^k u_i v_j \cdot 2^{(k+1+i-j)\beta} + \eta_{v,k} 2^{k\beta} + \dots + \eta_{v,1} 2^\beta + 1) < 2^{(k+1)\beta}$; (ii) when $i > j$, $(\sum_{i=1, i > j}^k u_i v_j \cdot 2^{(i-j-1)\beta} + \eta_{v,k+1}) \cdot 2^{(k+2)\beta}$; and (iii) when $i = j$, $(\sum_{i=1, i=j}^k u_i v_j \cdot 2^{(k+1)\beta})$. Thus, with Equation (??), $\vec{u} \cdot \vec{v}$ can be computed.

3.3.6 Data Report Aggregation

The data server first aggregates the group of sensed data generated outside the industrial district $\mathcal{N}_{out}$, which is

$$\begin{aligned}
C_{out} &= \prod_{v \in \mathcal{N}_{out}} C_{v,0} \bmod n^2 \\
&= g^{\sum_{v \in \mathcal{N}_{out}} (sum_v + H(\alpha || 0 || TS))} \cdot \left(\prod_{v \in \mathcal{N}_{out}} r_{v,0} \right)^n \bmod n^2.
\end{aligned} \tag{3.10}$$

Then, the data server decrypts C_{out} with the private key sk , and calculates the average of the sensed data generated outside the industrial district, which is $S_{out} = \sum_{v \in \mathcal{N}_{out}} (sum_v + H(\alpha||0||TS))$. Meanwhile, the data server aggregates the sensed data within the group $\mathcal{N}_{in}$, which is

$$\begin{aligned} C_{in} &= \prod_{v \in \mathcal{N}_{in}} C_{v,0} \bmod n^2 \\ &= g^{\sum_{v \in \mathcal{N}_{in}} (sum_v + H(\alpha||0||TS))} \cdot \left(\prod_{v \in \mathcal{N}_{in}} r_{v,0} \right)^n \bmod n^2. \end{aligned} \quad (3.11)$$

In addition, the data server decrypts $S_{in} = \sum_{v \in \mathcal{N}_{in}} (sum_v + H(\alpha||0||TS))$ with the private key sk ; meanwhile, the data server delivers S_{in} , S_{out} , K_{in} , $|\mathcal{N}_{in}|$ and $|\mathcal{N}_{out}|$ towards the data requester.

After receiving S_{in} , S_{out} , K_{in} , $|\mathcal{N}_{in}|$ and $|\mathcal{N}_{out}|$, the data requester first calculates the average of the sensed data outside the industrial district, which is

$$Ave_{out} = \frac{(S_{out} - |\mathcal{N}_{out}| \cdot H(\alpha||0||TS)) \bmod n}{d \cdot |\mathcal{N}_{out}|}. \quad (3.12)$$

Then, the data requester calculates the average of the sensed data inside the industrial district, which is

$$Ave_{in} = \frac{(S_{in} - (|\mathcal{N}_{in}| \cdot d - K_{in}) \cdot Ave_{out} - |\mathcal{N}_{in}| \cdot H(\alpha||0||TS)) \bmod n}{K_{in}}. \quad (3.13)$$

3.4 Security Analysis

Following the earlier discussed security requirements, we analyze the security properties of the proposed privacy-preserving range query scheme in this subsection. Specifically, our analysis will focus on how the proposed scheme can achieve the location privacy preservation of data requesters and data uploading vehicles, and protect the confidentiality of the individual on-board sensed data.

The proposed range query scheme can preserve the location privacy of data requester and data uploading vehicles. In the proposed scheme, the data requester first identifies a large region to construct the grid cell structure, which contains the industrial district. Based on the constructed grid cell structure, nothing could be learned besides the fact that the data query occupies one or a few grid cells in the defined grid cell structure. Since the location of the data requester and the vehicles are mapped to the constructed grid cell architecture, protecting the grid cell scalar of the data requester and the vehicles means preserving the location privacy. The following two paragraphs illustrate how the proposed scheme can protect the grid cell scalar during the secure scalar product computation process.

In the proposed scheme, the data requester's multi-dimensional grid cell scalar $\vec{u}$ is structured into one dimension m_u , and then encrypted by the public key of the data server pk to generate $(C_{u,1}, C_{u,2})$. After receiving $(C_{u,1}, C_{u,2})$, the data server decrypts $C_{u,1}$ with its private key sk , and obtains $m_u + H(\alpha||1||TS)$. To prevent the data server from recovering the value of m_u , $H(\alpha||1||TS)$ and $H(\alpha||2||TS)$ are introduced in $(C_{u,1}, C_{u,2})$. Since the secret α is only shared between the data requester and data uploading vehicles and the defined security model does not take collusion attack into consideration, the data server cannot recover the value of m_u . As $(C_{u,1}, C_{u,2})$ are valid ciphertexts of the Paillier cryptosystem, which is known to be semantically secure against the chosen plaintext attack, and the composite grid cell scalar contained in $(C_{u,1}, C_{u,2})$ is also semantically secure. Meanwhile, the locations of data uploading vehicles are also encrypted by the public key of the data server and protected by the secret α . Since only the data server with the private key sk can decrypt $(C_{u,1}, C_{u,2})$ and $(C_{v,1}, C_{v,2})$, and the data server cannot recover the value of m_u and m_v with the decrypted results, the composite grid cell scalars of the data requester and vehicles can be protected during transmission.

Due to the homomorphic property of the Paillier cryptosystem, the data server first computes $(C_{v,1})^{m_u + H(\alpha||1||TS)}$ and then recovers the value of $s_{u,v} = m_u \cdot m_v + \eta_v + 1$ through decrypting $\frac{(C_{v,1})^{m_u + H(\alpha||1||TS)}}{C_{u,2} \cdot C_{v,2}}$. To avoid the exhaustive attack against $m_u \cdot m_v$, a random number η_v is also included in $s_{u,v}$, but the data server can still obtain the scalar product of $\vec{u} \cdot \vec{v}$ by computing $\frac{(s_{u,v} \bmod 2^{(k+2)\beta}) - (s_{u,v} \bmod 2^{(k+1)\beta})}{2^{(k+1)\beta}}$. In this way, the content of each composite grid cell scalar can

still be protected. Meanwhile, the location privacy of the data requester and data uploading vehicles can be preserved during the secure scalar product computation.

The individual on-board sensed data report is confidential in the proposed scheme. In the proposed scheme, each vehicle's on-board sensed data report is formed as $C_{v,0} = g^{sum_v + H(\alpha||0||TS)} \cdot (r_{v,0})^n \bmod n^2$. Since $C_{v,0}$ is a valid ciphertext of the Paillier cryptosystem, the sensed data sum_v in $C_{v,0}$ is also semantically secure. After decrypting $C_{v,0}$ with sk , the data server still cannot recover the value of $sum_v + H(\alpha||0||TS)$, since α is a secret shared between the data requester and vehicles. In addition, the data server aggregates all the sensed data within the group $\mathcal{N}_{in}$ and the group $\mathcal{N}_{out}$, respectively, and delivers the aggregated results S_{in} and S_{out} to the data requester. Without α , the data server also cannot learn the aggregation of the sensed data reports sum_v , and the confidentiality of the aggregation results can be achieved. Furthermore, the data requester can only obtain the aggregated results contained in S_{in} and S_{out} , but it cannot recover the value of each individual sensed data. Therefore, the confidentiality of individual sensed data can also be protected.

3.5 Performance Evaluation

In this subsection, we describe the parameter setup, and evaluate the performance of the proposed scheme in terms of computation complexity and communication overhead.

3.5.1 Parameter Setup

We conduct the simulations with the Java Paillier Library [149] on a desktop with 3.40 GHz processor and 8.00 GB memory to study the operation costs. The experimental result shows that the cost of a single exponentiation operation in $\mathbb{Z}_{n^2}^*$ ($|n^2| = 2048$) is $C_e = 8.82$ ms. The proposed scheme integrates all the elements in a scalar to one composite value, which requires us to choose the proper parameter β and the length of the scalar k . Given $|n| = 1024$, the

maximum value of β_{max} is 7 and the maximum length of the scalar is $k_{max} = 63$, which satisfies the condition that $2 \cdot k_{max} < 2^{\beta_{max}}$ and $(2 \cdot k_{max} + 1) \cdot \beta_{max} < |n|$.

We compare the proposed scheme with the traditional homomorphic secure scalar product when Paillier encryption is exploited, i.e., each element in a scalar is encrypted separately at the data requester side, and then transmitted to the data server, i.e., the ciphertext of the data requester are generated as $C_{u,1,i}, C_{u,2,i}, i = 1, 2, \dots, k$. At the vehicle side, each element in a scalar is also encrypted and delivered individually, i.e., the ciphertext of each vehicle are generated as $C_{v,1,i}, C_{v,2,i}, i = 1, 2, \dots, k$. At the data server side, the product of two corresponding elements in the two scalars are also computed individually, $C_{u,v,i} = \frac{(C_{v,1,i})^{u_i + H(\alpha||1||TS)}}{C_{u,2,i} \cdot C_{v,2,i}}$, and then aggregated to $C_{u,v} = \prod_{i=1}^k C_{u,v,i}$. Finally, the aggregated results are decrypted to derive the final results.

3.5.2 Computation Complexity

In the proposed scheme, when a data requester generates a data query, it requires 4 exponentiation operations in $\mathbb{Z}_{n^2}^*$ to generate $C_{u,1}||C_{u,2}$. Note that the computation of $2^{i \cdot \beta}, i \in \{1, 2, \dots, k\}$ can be conducted in the setup phase, and multiplication operation in $\mathbb{Z}_{n^2}^*$ is negligible in comparison with the exponentiation operation in $\mathbb{Z}_{n^2}^*$. In the traditional scheme, the data requester needs to generate the ciphertext $C_{u,i,1}||C_{u,i,2}$ for each element in the scalar separately. For a scalar with the length of k , the data requester needs to consume $4 \times k$ exponentiation operations in $\mathbb{Z}_{n^2}^*$ to generate the ciphertext. In the proposed scheme, when a vehicle receives a data query, it generates an encrypted location-based data report, which requires 6 exponentiation operations in $\mathbb{Z}_{n^2}^*$. However, for the traditional scheme, each vehicle needs to consume $2 + 4 \times k$ exponentiation operations to generate one data report.

After receiving all the data reports from w vehicles, the data server should compute the scalar product of the data requester and each vehicle, to identify whether the on-board data report is harvested within the industrial district. The data server takes one exponentiation operation in $\mathbb{Z}_{n^2}^*$ to recover the value of $m_u + H(\alpha||1||TS)$ by Paillier decryption, and it also consumes two

exponentiation operations in $\mathbb{Z}_{n^2}^*$ to obtain the value of $m_u \cdot m_v + 1$. Since the multiplication operation in $\mathbb{Z}_{n^2}^*$ is considered negligible in comparison to the exponentiation operation in $\mathbb{Z}_{n^2}^*$, the computation cost of aggregation is negligible, and it takes two exponentiation operations in $\mathbb{Z}_{n^2}^*$ for the Paillier decryption to recover the aggregated results. For the traditional scheme, it takes $3 \times k$ exponentiation operations in $\mathbb{Z}_{n^2}^*$ to recover the scalar product and costs two exponentiation operations to obtain the aggregated results.

Thus, in the proposed scheme, totally for the data requester, vehicle, and the data server, the computation cost will be $4 \times C_e$, $6 \times C_e$, and $(3 \times w + 2) \times C_e$. For the traditional scheme, the computation cost for the data requester, vehicle, and data server are $4 \times k \times C_e$, $(2 + 4 \times k) \times C_e$, and $(3 \times k \times w + 2) \times C_e$, respectively, since each element in a scalar are encrypted and processed independently and the computation complexity is proportional to the length of the scalar.

Fig. 3.4 show the computation complexity of the data server with both schemes in terms of the scalar length (i.e., data dimension in the figures) and the number of vehicles. Simulation results show that the proposed scheme greatly reduces the computation complexity of the data server in comparison with the traditional scheme. As shown in Fig. 3.4a, the computation complexity increases with both scalar length and the number of vehicles in the traditional scheme, while in the proposed scheme, the computation complexity only increases with the number of vehicles as shown in Fig. 3.4b. Meanwhile, Fig. 3.5 present the computation complexity of the data requester and vehicle in terms of the scalar length, and compare the proposed scheme with the traditional scheme, respectively. Fig. 3.4 and Fig. 3.5 indicate that the proposed scheme can achieve lower computation complexity compared to the traditional scheme.

Since the modern vehicles are powered by the alternators, by charging the on-board batteries and powering the electrical system when its engine is running, the vehicles are normally not constraint by the power consumption of the computation process. Thus, we do not take the energy consumption into consideration.

FIGURE 3.4. Computation complexity of data server. **(a)** Computation complexity of the data server with the traditional scheme; **(b)** Computation complexity of the data server with the proposed scheme.

FIGURE 3.5. Comparison of the computation complexity of the data requester and vehicle. **(a)** Computation complexity of the data requester; **(b)** Computation complexity of one data uploading vehicle.

3.5.3 Communication Overhead

We evaluate the communication overhead that introduced the data query, which is sent from the data query towards the data server, and the data report sent from the vehicles towards the data server, since the data query sent from the data server towards the RSU and the data query response sent from the data server do not involve the transmission of ciphertexts in the proposed scheme.

FIGURE 3.6. Vehicles to data server communication overhead. **(a)** Communication overhead of vehicles during data uploading with the traditional scheme; **(b)** Communication overhead of vehicles during data uploading with the proposed scheme.

We first consider the data requester to data server communication, where the data requester generates a location-based data query and delivers the data query to the data server. The location-based query is in the format of $C_{u,1} || C_{u,2}$, and its size is 2048×2 bits, if we choose 1024-bit n . If the traditional homomorphic scalar product protocol is adopted, the corresponding communication overhead is $2048 \times 2 \times k$ bits, when there are k partitioned grid cells. In Fig. 3.6, we plot the data requester to data server communication overhead in terms of the scalar length, and compare the proposed scheme to the traditional scheme, which shows that the proposed scheme reduces the communication overhead.

During the vehicle to data server communication, each vehicle generates a location-based data report and delivers the data report to the data server via the RSU. The data report is in the format of $C_{u,0} || C_{u,1} || C_{u,2}$, and its size is 2048×3 bits. The data server collects w data reports from the total w vehicles, the communication overhead between the vehicles and the data server is $2048 \times 3 \times w$ bits. For the traditional homomorphic scalar product protocol, the communication between the vehicles and the data server is $2048 \times 3 \times w \times k$ bits, when the scalar length is k . Fig. 3.7a,b plot the communication overhead introduced by vehicles of both schemes in terms of the scalar length and the number of vehicles. Simulation results show that the proposed scheme greatly reduces the vehicle to data server communication overhead in comparison with the traditional scheme. As shown in Fig. 3.7a, the vehicle to data server communication overhead

FIGURE 3.7. Communication overhead from the data requester to the data server.

increases with both scalar length and the number of vehicles in the traditional scheme, while in the proposed scheme, vehicle to data server communication overhead only increases with the number of vehicles as shown in Fig. 3.7b.

3.6 Summary

In this chapter, we have proposed a privacy-preserving range query scheme from the distributed on-board storage in VANETs, which achieves the secure scalar product computation with the homomorphic Paillier cryptosystem. The proposed scheme structures the multi-dimensional

scalar into one dimension, identifies the data harvested within the industrial district, and computes the aggregation results of the identified sensed data. Security analysis have been conducted to demonstrate its security properties, in terms of location privacy preservation and confidentiality. Performance evaluations have also been done, which indicates that the proposed scheme can significantly reduce the computation complexity and communication overhead. However, during each data querying process, each responding vehicle needs to generate a location privacy-preserving data report, and the data server needs to perform the privacy-preserving data filtering to identify the target data report-by-report, which causes the problem of the scarce bandwidth in VANETs being occupied and the computational resource of the data server being consumed.

Chapter 4

A privacy-preserving sensory data sharing scheme in Internet of Vehicles

In this chapter, a novel efficient and location privacy-preserving data sharing scheme in Internet of Vehicles (IoV) is proposed, in which the users can directly query the target data from the network edge without the involvement of the core entity, which greatly improves the computation and communication costs. During the vehicular sensory data collection phase, the proposed scheme structures the multi-dimensional sensory data captured at different locations and exploits the modified Paillier cryptosystem to achieve the location privacy-preserving sensory data aggregation, in which the aggregation result can be fully recovered by the trusted central entity. During the vehicular sensory data acquisition phase, the proposed scheme exploits the proxy re-encryption technique to achieve the location privacy-preserving aggregated sensory data querying at the network edge, i.e., without the involvement of the trusted central entity. Numerical analysis are performed to validate the effectiveness of the proposed scheme, i.e., low data querying failure probability. Security analysis and performance evaluation are also conducted to validate the security properties and show the computation and communication efficiency of the proposed scheme, respectively.

4.1 Introduction

As one of the most important branches of IoT, IoV, which is an extension of VANETs, mainly focuses on the information interaction between vehicles, roadside infrastructures, and surrounding environments without human intervention [150]. By coordinating the data captured by vehicular sensors (e.g., chemical spill detectors, vibration sensors, and acoustic detectors [135]), which can record a myriad of characteristics reflecting the physical phenomena, IoV enables the data sharing among vehicles and the deployed roadside infrastructures to further improve traffic safety and on-board experience in the intelligent transportation system (ITS); meanwhile, it can also support various value-added services, such as autonomous driving [151], fuel optimal cruising [152], and early warning [153].

Although IoV can improve the ITS performance through information interaction, there exist several challenges for the wide application of vehicular sensory data sharing in IoV, including location privacy preservation, vehicular sensory data collection, and vehicular sensory data acquisition. During the sensory data collection phase, since the captured sensory data are highly location-dependent, the location-based sensory data may disclose the moving trajectory and violate the location privacy of vehicles [154]. Recently, privacy-preserving data aggregation has received much attention in the era of IoT [63, 64], which are designed to achieve the data aggregation result while protecting the content of each individual data report. In [63], to achieve efficient privacy-preserving data aggregation, each user structures the small-sized multi-dimensional sensory data into one composite data report and encrypts each composite data report with the homomorphic encryption technique, and then the gateway aggregates the encrypted composite data reports. However, in IoV, vehicles are dynamically moving and the scale of sensory data reports in each data dimension (if we take different locations as different data dimensions) are constantly fluctuating in the temporal domain. The existing schemes are not able to count how many pieces of sensory data are generated in each data dimension, and further fail to calculate the average of the sensory data in each data dimension. Thus, to fully exploit sensory data generated by vehicles, the design of an efficient and location privacy-preserving data collection scheme in IoV is essential.

During the sensory data acquisition phase, since vehicles are normally interested in acquiring the sensory data in accordance with the potential moving path towards the destination, the location-based data queries may violate the location privacy of vehicles [99, 116]. Location-based data querying schemes with privacy preservation has attracted considerable attention [110, 114], however, the existing schemes are mainly designed for the outsourced cloud environment with fixed cloud storage, which cannot be directly applied towards the real-time services in IoV. To preserve location privacy, the querying location needs to be hid and the data captured at the unqueried locations also needs to be protected, and the oblivious transfer (OT) [117] can be employed to achieve both during the data querying process. However, in the existing OT protocols [115, 117], the communication overhead is proportional to the data dimension, which are not applicable for the IoV environment with short-termed and intermittent wireless connection. Thus, to achieve the efficient sensory data sharing in IoV, an efficient and location privacy-preserving data acquisition scheme in IoV is also essential.

In this chapter, aiming at the above challenges, we propose a novel efficient and location privacy-preserving data sharing scheme in IoV with the modified Paillier Cryptosystem, which enables the RSU to collect and distribute the sensory data generated by vehicles. Specifically, the contributions of this chapter are threefold as follows.

- First, we propose a privacy-preserving on-board sensory data collection scheme in IoV. The proposed scheme first enables a vehicle to structure the multi-dimensional sensory data captured at different locations into one composite data report, which greatly saves the communication and computation resources. The proposed scheme can also support the RSU to perform the privacy-preserving composite data reports aggregation and count the number of data reports contained in each data dimension. Meanwhile, the sensory data aggregation result can be retrieved and re-encrypted by the trusted traffic management authority.
- Second, we propose a privacy-preserving sensory data acquisition scheme in IoV. The proposed scheme enables a data querying vehicle to retrieve the sensory data captured

by other vehicles at the network edge, i.e., without the involvement of the trusted traffic management authority. Specifically, the RSU re-encrypts the aggregated sensory data into ciphertexts that can only be fully recovered by the data querying vehicle, while without the disclosure of the data querying location.

- Third, we offer numerical analysis to compute the probability of data querying failure, and perform detailed security analysis to validate the security property of the proposed scheme in terms of location privacy preservation, data integrity protection and collusion attack resistance. Through performance evaluation, we demonstrate that the proposed scheme can significantly reduce the computation cost and communication overhead in comparison with the existing secure data aggregation and OT protocols.

The remainder of this chapter is organized as follows. We introduce system model, show security requirements, and identify design goals in Section 4.2. In Section 4.3, we recall the modified Paillier cryptosystem and Chinese Remainder Theorem as the preliminaries. Then we present our proposed privacy-preserving data sharing scheme in Section 4.4, followed by our security analysis and performance evaluation in Section 4.5 and in Section 4.6, respectively. We finally conclude our work in Section 4.7.

4.2 System Model, Security Requirements, and Design Goals

In this section, we introduce system model, show security requirements, and identify design goals.

4.2.1 System Model

In the system model, we consider a practical scenario in an urban area, in which the data querying vehicle (V_a) aims to drive towards the destination (a hospital) in an efficient and safe

FIGURE 4.1. Vehicular Sensory Data Sharing System

manner, as shown in Figure 4.1. By querying the sensory information captured by other vehicles' inductive rotational-speed sensors [155] (the input data for the anti-lock braking system [156], which can detect the slippery area), V_a discover the hidden road conditions. If V_a notices the existence of the slippery area (the shadow area of a building blockage after snow in winter), V_a will avoid it and take an alternative path. The proposed system consists of four roles: a traffic management authority, an RSU, a group of data uploading vehicles ($V_i, i = 1, 2, \dots, w$), and a data querying vehicle (V_a), as shown in Figure 4.1.

- Traffic management authority is the centralized entity, which is responsible for the registration and key distribution of the RSU and vehicles. Traffic management authority divides the urban area into districts (e.g., with the coverage area of approximately 3 km^2), such that each district is large enough for the vehicles to be comfortable with revealing fact that they are locating somewhere within the district. In each district, the traffic management separates the roads into segments, and the sensory data generation locations

are denoted in the level of segments. For segment l , it is denoted with a unique two-dimension segment identifier (u_l, v_l) , which is derived from the segment coordinates.

- Each sensory data uploading vehicle ($V_i, i = 1, 2, \dots, w$) is equipped with the required inductive rotational-speed sensors, it periodically formulates a data report with the captured data, and uploads the data report towards the RSU, as shown in Figure 4.1 ((1) *Data Report*).
- The RSU collects and aggregates the on-board sensory data reports uploaded by the data uploading vehicles, and transmits the data aggregation result towards the traffic management authority, as shown in Figure 4.1 ((2) *Data Aggregation Result*). Meanwhile, it shares the aggregated on-board sensory data report with other vehicles when requested, as shown in Figure 4.1 ((4) *Data Query Response*). We set the transmission distance of the RSU to 1 km, such that it can provide signal coverage of a whole district [157].
- When the data querying vehicle (V_a) travels in vicinity to the potential slippery area, i.e., several segments ahead, it sends a data query towards the RSU for the sensory data captured at the slippery area, as shown in Figure 4.1 ((3) *Data Query*).

Communication Model. The wireless connections between the RSU and vehicles are realized through the IEEE 802.11p standard [145], which is a short to medium range communication technology operating at 5.9 GHz band designed for WAVE. The connection between the RSU and the traffic management authority is realized through either wired links or any other links with high bandwidth and low transmission delay.

4.2.2 Security Requirements

In the security model, we consider the traffic management authority is fully trusted. The RSU and vehicles are honest-but-curious (i.e., they will follow the defined protocols, but they are curious about the location information of other vehicles). Meanwhile, we assume the vehicular sensory data are correctly generated and delivered towards the RSU. Moreover, we assume

there is a compromised RSU, which may collude with a malicious vehicle, i.e., through mutually sharing their secret keys. We also make an assumption that there exists an adversary $\mathcal{A}$ located on the road, which could launch some active attacks to modify the data transmission and threaten the data integrity. Therefore, the following security requirements should be met.

Location Privacy Preservation. In the proposed scheme, since the locations of vehicles are denoted with segments, preserving location privacy means preventing the segments from being uncovered. During the sensory data collection phase, preserving the location privacy of data uploading vehicles indicates that, the sensory data generation segments contained in an individual data report should be protected, i.e., even if the RSU obtains all the possible data reports generated by a data uploading vehicle, it still cannot identify the data generation segments contained in an individual data report. During the sensory data acquisition phase, since a data querying vehicle only interests in sensory data in accordance with the potential trajectory towards its destination, the segment identifier contained in a data query should also be protected. In addition, the location privacy preservation requirement also should include that a data querying vehicle can only obtain the sensory data captured in the segment it is querying about.

Collusion resistance. The proposed scheme should resist the collusion between a compromised RSU and a malicious vehicle by sharing their secrets. Given the secret key of the compromised RSU, the malicious vehicle should not derive the content contained in any individual sensory data report during the data collection phase and should not obtain the location identifier contained in any data query during the data acquisition phase. Meanwhile, when the malicious vehicle receives a data query response from an uncompromised RSU, it can only learn the sensory data generated at the segment it is querying about. Given the session key of the malicious vehicle, the compromised RSU should also not obtain the content contained in any individual data report and the location identifiers contained in any individual data query.

Data Integrity. The proposed scheme should guarantee that any sensory data report, data query and data query response, should not be modified during the data transmission process, i.e., even

if the adversary $\mathcal{A}$ alters the data transmission, the malicious operation should be detected. In this way, the security requirement of data integrity could be achieved.

4.2.3 Design Goals

Under the aforementioned system model and security requirements, our design goal is to develop an efficient and privacy-preserving sensory data sharing scheme, which enables the vehicles to upload the captured sensory data towards the RSU and allows the vehicles to retrieve the sensory data captured by other vehicles. Specifically, the following three requirements should be achieved.

The proposed scheme should achieve the defined security requirements. If the security requirements are not taken into consideration, the location privacy of the involved vehicles could be violated, and the data transmission between the RSU and vehicles could be modified. Then the vehicles are not willing to join in the sensory data sharing process, and the sensory data could not be properly shared and fully exploited.

The proposed scheme should achieve high effectiveness during the on-board sensory data sharing process. Since the data captured by the on-board sensors are location-dependent and safety-critical, the accuracy of the collected on-board sensory data is mandatory. Meanwhile, in order to achieve the effective sensory data acquisition, the RSU should be able to provide near real-time sensory data with low probability of data querying failure.

The proposed scheme should achieve high efficiency in terms of computation complexity and communication overhead. The dynamically moving characteristic of vehicles and the scarce transmission bandwidth in vehicular network result in the requirement of short transmission interval, low processing delay and low communication overhead. Thus, the proposed scheme should achieve high computation and communication efficiency.

4.3 Preliminaries

In this section, we briefly review the modified Paillier cryptosystem [158] and the Chinese Remainder Theorem [159], which will serve as the basis of the proposed vehicular sensory data sharing scheme.

4.3.1 Modified Paillier Cryptosystem

The modified Paillier cryptosystem is exploited due to its homomorphic addition property and homomorphic multiplication property between one ciphertext and one plaintext [158], which consists of four components: key generation, encryption, decryption, and proxy re-encryption.

- *Key Generation:* Given security parameter κ , two large safe prime numbers p_1, p_2 in the form of $p_1 = 2p'_1 + 1$ and $p_2 = 2p'_2 + 1$ are first chosen, where $|p_1| = |p_2| = \kappa$, and p'_1, p'_2 are also primes. Then, the RSA modulus $n = p_1 p_2$ and $\lambda = \text{lcm}(p_1 - 1, p_2 - 1)$ are computed. We consider $\mathbb{G} = QR_{n^2}$ the cyclic group of quadratic residues modulo n^2 , whose order is $\text{ord}(\mathbb{G}) = \lambda(n^2)/2 = p_1 p'_1 p_2 p'_2 = n\lambda/2$, and the maximal order of an element in this group is $n\lambda/2$. Choose a random number $\mu \in \mathbb{Z}_{n^2}^*$ and a random value $x \in [1, \text{ord}(\mathbb{G})]$, then set $g = \mu^2 \bmod n^2$ and $h = g^x \bmod n^2$. The public key is represented as $pk = (n, g, h = g^x)$, and the corresponding secret key is x .

- *Encryption:* Given message $m \in \mathbb{Z}_n$, choose a random number $r \in \mathbb{Z}_{n^2}$ and generate the ciphertext pair (c_1, c_2) as $c_1 = g^r \bmod n^2$ and $c_2 = h^r(1 + mn) \bmod n^2$.

- *Decryption:* Given ciphertext pair (c_1, c_2) , message m can be decrypted as $m = \frac{c_2/(c_1)^x - 1 \bmod n^2}{n}$.

- *Proxy Re-encryption:* Randomly splits the secret key x into two shares x_1 and x_2 , such that $x = x_1 + x_2$. The modified Paillier cryptosystem enables an encrypted message (c_1, c_2) to be partially decrypted to a ciphertext pair as $(\tilde{c}_1, \tilde{c}_2)$ using x_1 as $\tilde{c}_1 = c_1$ and $\tilde{c}_2 = c_2/(c_1)^{x_1} \bmod n^2$. Then $(\tilde{c}_1, \tilde{c}_2)$ can be decrypted with x_2 to recover the original message m .

TABLE 4.1. Notations of Proposed Sensory Data Sharing Scheme

Notation	Definition
p_1, q_1	Parameters of Modified Paillier Cryptosystem
$pk = (n, g, h = g^x)$	Public Key of Modified Paillier Cryptosystem
$sk = (x = x_1 + x_2)$	Secret Key x of Modified Paillier Cryptosystem (Split into Two Shares)
(u_l, v_l)	Location Identifiers
α_l	Chinese Remainder Theorem Parameters
d	Random Number related to Sensory Data Size
$(d_{i,l}, e_{i,l})$	Sensory Data and Indicator

4.3.2 Chinese Remainder Theorem

The Chinese Remainder Theorem can solve any pair of congruences with relatively prime moduli [159], which can be exploited to structure the multi-dimension data into one structured data and is described as follows.

Chinese Remainder Theorem: Suppose that $q_1, q_2, \dots, q_k$ are relatively prime integers, and let $d_1, d_2, \dots, d_k$ be integers. Then, the system of congruences $s \equiv d_l \pmod{q_l}$ for $1 \leq l \leq k$, has a unique solution modulo $Q = q_1 \times q_2 \times \dots \times q_k$, which is

$$s \equiv d_1 Q_1 y_1 + d_2 Q_2 y_2 + \dots + d_k Q_k y_k \pmod{Q} \quad (4.1)$$

where $Q_l = \frac{Q}{q_l}$ and $y_l \equiv \frac{1}{Q_l} \pmod{q_l}$, for $1 \leq l \leq k$.

4.4 Proposed Sensory Data Sharing Scheme

In this section, we present our privacy-preserving vehicular sensory data sharing scheme, which mainly consists of three phases: system initialization, sensory data collection, and sensory data acquisition. Meanwhile, the definitions of some important notations are described in Table 4.1.

4.4.1 System Initialization

Given security parameter κ , the traffic management authority, which functions as a trusted authority (TA), selects two large safe prime numbers p_1 and p_2 , where $|p_1| = |p_2| = \kappa$, then it generates the modified Paillier Cryptosystem's public key $pk = (n = p_1 \cdot p_2, g, h = g^x \bmod n^2)$ and the private key $sk = x$. Then TA randomly splits the secret key x into three shares x_0, x_1 and x_2 , such that $x = x_0 + x_1 + x_2$, and it also calculates the value of $h_1 = g^{x_1+x_2} \bmod n^2$. Meanwhile, TA keeps the secret share x_0 to itself. During the registration of each RSU, TA securely delivers the secret value x_1 towards the RSU; meanwhile, during the registration of each vehicle, TA securely distributes the secret value x_2 towards the vehicle.

For a district with k segments, TA selects k equal-length prime numbers: $(q_1, q_2, \dots, q_k)$, where $|q_1| = |q_2| = \dots = |q_k| = \kappa_1$, and it satisfies the condition that $k \cdot \kappa_1 < |n|$. For segment l with the segment identifier (u_l, v_l) , TA assigns a prime number q_l towards it. Then TA computes

$$\begin{cases} Q = \prod_{l=1}^k q_l \\ Q_l = \frac{Q}{q_l}, y_l \equiv \frac{1}{Q_l} \bmod q_l \\ \alpha_l = Q_l \cdot y_l. \end{cases} \quad (4.2)$$

In addition, TA selects a constant number d with $|d| = \kappa_1 - 2$. Meanwhile, it chooses two secure cryptographic functions, $H_0 : \{0, 1\}^* \rightarrow \mathbb{Z}_n$ and $H : \{0, 1\}^* \rightarrow \{0, 1\}^{\kappa_1-2}$, such that $|H(\cdot) + d| < \kappa_1$. Finally, TA publishes the parameters: $params = \{pk, h_1, H_0, H, q_l : l = 1, 2, \dots, k\}$.

4.4.2 Sensory Data Collection

We describe the data report collection process as below and illustrate the process in Figure 4.2 (*Phase I*). For vehicle V_i , we let $d_{i,l}$ denote the data captured by the inductive rotational-speed sensor at segment l , where $|d_{i,1}| = |d_{i,2}| = \dots = |d_{i,k}|$ and $d_{i,l} < d, l = 1, 2, \dots, k$. Meanwhile, $e_{i,l}$ represents whether there is sensory data captured at segment l : (i) if there is sensory

data captured at segment l , i.e., $d_{i,l} > 0$, then $e_{i,l} = 1$; (ii) if there is no sensory data captured at segment l , i.e., $d_{i,l} = 0$, then $e_{i,l} = 0$.

FIGURE 4.2. Message flow during data sharing

In order to achieve the near real-time sensory data uploading, each vehicle V_i gathers the sensory data captured at all the segments within a district during a short period T , e.g., every ten minutes, and performs the following steps to generate a sensory data report:

- V_i structures $(d_{i,1}, d_{i,2}, \dots, d_{i,k})$ and $(e_{i,1}, e_{i,2}, \dots, e_{i,k})$ into two composite values $s_{i,d}$ and $s_{i,e}$, i.e., $s_{i,d} = \sum_{l=1}^k d_{i,l} \cdot \alpha_l$ and $s_{i,e} = \sum_{l=1}^k e_{i,l} \cdot \alpha_l$.
- V_i selects two random numbers $(r_{i,d}, r_{i,e}) \in \mathbb{Z}_{n^2}$, and generates the following ciphertexts,

$$\begin{cases} C_{i,1} = g^{r_{i,d}} \bmod n^2, C_{i,2} = h^{r_{i,d}} \cdot (1 + s_{i,d} \cdot n) \bmod n^2, \\ C_{i,3} = g^{r_{i,e}} \bmod n^2, C_{i,4} = h^{r_{i,e}} \cdot (1 + s_{i,e} \cdot n) \bmod n^2. \end{cases} \quad (4.3)$$

where $(C_{i,1}, C_{i,2})$ and $(C_{i,3}, C_{i,4})$ are the ciphertext pairs of $s_{i,d}$ and $s_{i,e}$, respectively.

- To protect data integrity, V_i also calculates the message authentication code of the ciphertext,

$$MAC_i = H_0(C_{i,1} || C_{i,2} || C_{i,3} || C_{i,4} || TS), \quad (4.4)$$

where TS is the current timestamp. Note that the identity-based signature can also be exploited here to protect the data integrity [27], and authenticate the origins of the sensory data reports if necessary.

Finally, V_i transmits $Data_Report_i = V_i || C_{i,1} || C_{i,2} || C_{i,3} || C_{i,4} || MAC_i || TS$ towards the RSU, as shown in Figure 4.2 (*Message (1)*).

After receiving the data report $Data_Report_i$, the RSU first verifies its correctness by checking $MAC_i \stackrel{?}{=} H_0(C_{i,1} || C_{i,2} || C_{i,3} || C_{i,4} || TS)$. If it does hold, the RSU performs the following steps to aggregate the data reports received from w data uploading vehicles, where the number of data uploading vehicles w satisfies the condition that $\lceil \log_2 w \rceil + |d_{i,l}| < |d|, l = 1, 2, \dots, k$.

$$\left\{ \begin{array}{l} C_{d,1} = \prod_{i=1}^w C_{i,1} \bmod n^2 = \prod_{i=1}^w g^{r_{i,d}} \bmod n^2 \\ C_{d,2} = \prod_{i=1}^w C_{i,2} \bmod n^2 = \left(\prod_{i=1}^w h^{r_{i,d}} \right) \cdot \left(1 + \sum_{i=1}^w s_{i,d} \cdot n \right) \bmod n^2 \\ C_{e,1} = \prod_{i=1}^w C_{i,3} \bmod n^2 = \prod_{i=1}^w g^{r_{i,e}} \bmod n^2 \\ C_{e,2} = \prod_{i=1}^w C_{i,4} \bmod n^2 = \left(\prod_{i=1}^w h^{r_{i,e}} \right) \cdot \left(1 + \sum_{i=1}^w s_{i,e} \cdot n \right) \bmod n^2 \end{array} \right. \quad (4.5)$$

where $(C_{d,1}, C_{d,2})$ and $(C_{e,1}, C_{e,2})$ are the ciphertext pairs of the multi-dimensional sensory data aggregation $\sum_{i=1}^w s_{i,d}$, and the corresponding scale $\sum_{i=1}^w s_{i,e}$. Then the RSU calculates the message authentication code of the aggregated ciphertext pairs $MAC_{ag} = H_0(C_{d,1} || C_{d,2} || C_{e,1} || C_{e,2} || TS)$, formulates the sensory data aggregation message $Data_Aggregation = C_{d,1} || C_{d,2} || C_{e,1} || C_{e,2} || MAC_{ag} || TS$, and delivers it towards the traffic management authority, as shown in Figure 4.2 (*Message (2)*).

After receiving $Data_Aggregation$, the traffic management authority verifies its correctness by checking whether $MAC_{ag} \stackrel{?}{=} H_0(C_{d,1} || C_{d,2} || C_{e,1} || C_{e,2} || TS)$, and re-encrypts the ciphertext $(C_{d,1}, C_{d,2})$ and $(C_{e,1}, C_{e,2})$ with the secret key x_0 , which is shown as follows

$$\left\{ \begin{array}{l} C_{t,1} = \frac{C_{d,2}}{(C_{d,1})^{x_0}} \bmod n^2 \\ C_{t,2} = \frac{C_{e,2}}{(C_{e,1})^{x_0}} \bmod n^2 \end{array} \right. \quad (4.6)$$

Then the traffic management authority generates the message authentication code $MAC_t = H_0(C_{t,1}||C_{t,2}||TS)$, and delivers $C_{t,1}||C_{t,2}||MAC_t||TS$ towards the RSU. In addition, the traffic management authority decrypts the received ciphertext pairs $(C_{d,1}, C_{d,2})$ and $(C_{e,1}, C_{e,2})$ with the secret key x , which is

$$\begin{cases} \sum_{i=1}^w s_{i,d} = \frac{C_{d,2}/(C_{d,1})^x - 1 \bmod n^2}{n}, \\ \sum_{i=1}^w s_{i,e} = \frac{C_{e,2}/(C_{e,1})^x - 1 \bmod n^2}{n}. \end{cases} \quad (4.7)$$

Finally, the traffic management authority can achieve the average of the vehicular sensory data $\bar{d}_l$ captured at segment l . According to the Chinese Remainder Theorem, $\bar{d}_l$ can be computed as

$$\bar{d}_l = \frac{\sum_{i=1}^w s_{i,d} \bmod q_l}{\sum_{i=1}^w s_{i,e} \bmod q_l} = \frac{\sum_{i=1}^w d_{i,l} \bmod q_l}{\sum_{i=1}^w e_{i,l} \bmod q_l}. \quad (4.8)$$

4.4.3 Sensory Data Acquisition

The data querying vehicle V_a intends to acquire the data captured by the inductive rotational-speed sensors at segment c with segment identifier (u_c, v_c) , and we summarize the data acquisition process as below and illustrate the process in Figure 4.2 (*Phase 2*). V_a selects three random numbers $(s_a, r_{a,1}, r_{a,2}) \in \mathbb{Z}_{n^2}$, and performs the following steps

- V_a calculates the encrypted data query

$$\begin{cases} C_{a,1} = g^{r_{a,1}} \bmod n^2, C_{a,2} = h_1^{r_{a,1}} \cdot g^{s_a \cdot r_{a,1}} \cdot (1 + u_c \cdot n) \bmod n^2, \\ C_{a,3} = g^{r_{a,2}} \bmod n^2, C_{a,4} = h_1^{r_{a,2}} \cdot g^{s_a \cdot r_{a,2}} \cdot (1 + v_c \cdot n) \bmod n^2. \end{cases} \quad (4.9)$$

where $(C_{a,1}, C_{a,2})$ and $(C_{a,3}, C_{a,4})$ are the ciphertext pairs of u_c and v_c respectively.

- To protect the data integrity, V_a generates the corresponding message authentication code MAC_a , which is

$$MAC_a = H_0(C_{a,1}||C_{a,2}||C_{a,3}||C_{a,4}||TS). \quad (4.10)$$

Then V_a transmits the data query towards the RSU, which is

$Data_Query_a = V_a || C_{a,1} || C_{a,2} || C_{a,3} || C_{a,4} || MAC_a || TS$, as shown in Figure 4.2 (*Message (3)*).

After receiving $Data_Query_a$, the RSU verifies its correctness by checking

$MAC_a \stackrel{?}{=} H_0(C_{a,1} || C_{a,2} || C_{a,3} || C_{a,4} || TS)$. If it does hold, the RSU selects two random numbers $(t_1, t_2) \in \mathbb{Z}_{n^2}$, and performs the following steps:

- The RSU calculates a secret key $\beta_l = H(m_l), l = 1, 2, \dots, k$, for each segment l with location identifier (u_l, v_l) , where $m_l = u_l \cdot t_1 + v_l \cdot t_2 \bmod n$. By computing $s_k = \sum_{l=1}^k \beta_l \cdot \alpha_l$, the RSU structures the multi-dimensional secret key $(\beta_1, \beta_2, \dots, \beta_k)$ into one composite secret key s_k .
- The RSU generates the ciphertext pair of $m_c = u_c \cdot t_1 + v_c \cdot t_2 \bmod n$ without learning the value of (u_c, v_c) , which is

$$\begin{cases} C_{r,1} = (C_{a,1})^{t_1} \cdot (C_{a,3})^{t_2} \bmod n^2, \\ C_{r,2} = \frac{(C_{a,2})^{t_1} \cdot (C_{a,4})^{t_2}}{(C_{r,1})^{x_1}} \bmod n^2, \end{cases} \quad (4.11)$$

such that m_c can only be recovered by V_a with the shared secret key x_2 and the privately-owned secret value s_a .

- The RSU also generates the ciphertext pairs of $\sum_{i=1}^w s_{i,d} + s_k$ and $\sum_{i=1}^w s_{i,e} + s_k$ respectively, and re-encrypts the derived ciphertext pairs with the secret key x_1 , such that they can be decrypted by V_a ,

$$\begin{cases} C_{r,3} = C_{d,1} \\ C_{r,4} = \frac{C_{t,1} \cdot (1 + s_k \cdot n)}{(C_{d,1})^{x_1}} \bmod n^2 \\ C_{r,5} = C_{e,1} \\ C_{r,6} = \frac{C_{t,2} \cdot (1 + s_k \cdot n)}{(C_{e,1})^{x_1}} \bmod n^2 \end{cases} \quad (4.12)$$

- To protect the data integrity, the RSU generates the corresponding message authentication code MAC_r , which is

$$MAC_r = H_0(C_{r,1}||C_{r,2}||C_{r,3}||C_{r,4}||C_{r,5}||C_{r,6}||TS). \quad (4.13)$$

Then the RSU transmits the data query response towards V_a , which is denoted as $Response_a = C_{r,1}||C_{r,2}||C_{r,3}||C_{r,4}||C_{r,5}||C_{r,6}||MAC_r||TS$, as shown in Figure 4.2 (*Message (4)*).

After receiving the data query response $Response_a$, V_a checks its validity by checking $MAC_r \stackrel{?}{=} H_0(C_{r,1}||C_{r,2}||C_{r,3}||C_{r,4}||C_{r,5}||C_{r,6}||TS)$. If it does hold, V_a performs the following steps to recover the queried vehicular sensory data:

- V_a recovers the value of m_c with the secret key x_2 and the secret value s_a , which is

$$m_c = u_c \cdot t_1 + v_c \cdot t_2 \bmod n = \frac{\frac{C_{r,2}}{(C_{r,1})^{x_2+s_a}} - 1 \bmod n^2}{n} \quad (4.14)$$

Then V_a computes the corresponding secret key $\beta_c = H(m_c)$.

- V_a also exploits the secret key x_2 to decrypt $(C_{r,3}, C_{r,4})$ and $(C_{r,5}, C_{r,6})$, and recovers the value of $\sum_{i=1}^w s_{i,d} + s_k$ and $\sum_{i=1}^w s_{i,e} + s_k$, respectively.
- V_a recovers the sum of the sensory data captured at segment c , i.e., $sum_{c,d}$ and the number of sensory data reports captured at segment c , i.e., $sum_{c,e}$, which is computed as $sum_{c,d} = \sum_{i=1}^w s_{i,d} + s_k \bmod q_c = \sum_{i=1}^w d_{i,c} + \beta_c$ and $sum_{c,e} = \sum_{i=1}^w s_{i,e} + s_k \bmod q_c = \sum_{i=1}^w e_{i,c} + \beta_c$, respectively. Then V_a derives the average of sensory data $\bar{d}_c$ captured at segment c , by computing $\bar{d}_c = \frac{sum_{c,d} - \beta_c}{sum_{c,e} - \beta_c}$.

4.4.4 Discussion

In this subsection, we analyze the probability when a data querying vehicle fails to retrieve data from the RSU, i.e., the probability of data querying failure. Here we define the probability of

FIGURE 4.3. Probability of data querying failure

data querying failure of segment l as the probability of the scenario when the number of data reports generated at segment l is less than a pre-defined threshold th , and the number of data querying vehicles is above zero. Meanwhile, the data uploading period is defined as T .

Let $X_{d,l}(T)$ denote the number of data reports which contains segment l within a finite time period $[0, T]$ and $X_{q,l}(T)$ be the number of data queries of segment l during $[T, 2T]$, which follows the Poisson $(\lambda_{d,l} \cdot T)$ and the Poisson $(\lambda_{q,l} \cdot T)$ distribution [160]. Then we obtain the probability of $X_{d,l}(T) = x_d$ data uploading vehicles and the probability of $X_{q,l}(T) = x_q$ data querying vehicles of segment l respectively, which are

$$\begin{cases} P\{X_{d,l}(T) = x_d\} = e^{-\lambda_{d,l} \cdot T} \cdot \frac{(\lambda_{d,l} \cdot T)^{x_d}}{x_d!} \\ P\{X_{q,l}(T) = x_q\} = e^{-\lambda_{q,l} \cdot T} \cdot \frac{(\lambda_{q,l} \cdot T)^{x_q}}{x_q!} \end{cases} \quad (4.15)$$

Let N_d and N_q be the average number of data reports and data queries in the given district, respectively. Meanwhile, $p_{d,l}$ and $p_{q,l}$ denote the probability of data uploading and the probability of data querying on segment l , respectively, which satisfy the condition that $\sum_{l=1}^k p_{d,l} = 1$ and $\sum_{l=1}^k p_{q,l} = 1$. Suppose all the vehicles are moving in the district with a constant speed of V_0 , and all the street segments are with the length of L m. Then the value of $\lambda_{d,l}$ and $\lambda_{q,l}$ can be calculated as

$$\lambda_{d,l} = \frac{N_d \cdot p_{d,l}}{L/V_0}, \lambda_{q,l} = \frac{N_q \cdot p_{q,l}}{L/V_0}. \quad (4.16)$$

The probability of data querying failure in segment l can be calculated as follows

$$\begin{aligned}
P\{X_{d,l}(T) \leq th \mid X_{q,l}(T) > 0\} &= \frac{P\{X_{d,l}(T) \leq th\}}{P\{X_{q,l}(T) > 0\}} \\
&= \frac{\sum_{x_d=0}^{th} P\{X_{d,l}(T) = x_d\}}{1 - P\{X_{q,l}(T) = 0\}} \\
&= \frac{\sum_{x_d=0}^{th} e^{-\lambda_{d,l} \cdot T} \cdot \frac{(\lambda_{d,l} \cdot T)^{x_d}}{x_d!}}{1 - e^{-\lambda_{q,l} \cdot T}}.
\end{aligned} \tag{4.17}$$

Then we calculate the probability of data querying failure in the given district, which is

$$P_f = \sum_{l=1}^k p_{q,l} \cdot P\{X_{d,l}(T) \leq th \mid X_{q,l}(T) > 0\}. \tag{4.18}$$

In Figure 4.3(a) and Figure 4.3(b), we show the probability of data querying failure in terms of the average number of data reports N_d and the data uploading period T , when the threshold are set to be $th = 5$ and $th = 0$ respectively. In Figure 4.3(a) and Figure 4.3(b), the number of data reports N_d ranges from 50 to 150, and the data uploading period T varies from 300 to 600 seconds. Meanwhile, we set the average number of the data querying vehicles to be $N_q = 50$, the length of each street segment is set to be $L = 500$ m, and the velocity of each vehicle is defined to be $V_0 = 10$ m/s. In addition, we assume that the number of segments in the given district are $k = 40$, and all the segments in the given district have an equal data uploading and data querying probability of $p_{d,l} = p_{q,l} = 0.025, l = 1, 2, \dots, k$. As shown in Figure 4.3(a) and Figure 4.3(b), the probability of data querying failure decreases with respect to the increase of the number of data reports N_d and the increase of the data uploading period T .

4.5 Security Analysis

In this section, we discuss the security properties of the proposed vehicular sensory data sharing scheme. Specifically, following the security requirements discussed earlier, our analysis will focus how the proposed scheme can achieve the security requirement of location privacy

preservation while taking the collusion between a compromised RSU and a malicious vehicle into consideration, and achieve the security goal of data integrity protection of the data transmission.

The proposed sensory data sharing scheme is location privacy-preserving during the sensory data collection phase. During the sensory data collection phase, since the locations are denoted in the level of segments, preserving the location privacy of data uploading vehicles is realized in terms of protecting the data collection segments. According to the Chinese Remainder Theorem, the sensory data collection segments of V_i can be achieved from $(s_{i,d}, s_{i,e})$, which needs to be protected. In the proposed scheme, $s_{i,d}$ and $s_{i,e}$ are encrypted by the public key $(n, g, h = g^x)$ to generate the ciphertext pairs $(C_{i,1}, C_{i,2})$ and $(C_{i,3}, C_{i,4})$. Since we exploit the modified Paillier cryptosystem, which is proved to be semantically secure against the adaptive chosen-ciphertext attack [161], the messages contained in $(C_{i,1}, C_{i,2})$ and $(C_{i,3}, C_{i,4})$ are also semantic secure. In case of the collusion between a compromised RSU and a malicious vehicle, as the colluding entities can only recover the sum of the secret values $x_1 + x_2$, they cannot obtain the individual data report $s_{i,d}$ and $s_{i,e}$ without x (or x_0), which preserves the location privacy of each individual data uploading vehicle.

After receiving w data reports, the RSU aggregates the encrypted data reports by exploiting the homomorphic addition property of the modified Paillier cryptosystem, i.e., $(C_{d,1} = \prod_{i=1}^w C_{i,1} \bmod n^2, C_{d,2} = \prod_{i=1}^w C_{i,2} \bmod n^2)$ and $(C_{e,1} = \prod_{i=1}^w C_{i,3} \bmod n^2, C_{e,2} = \prod_{i=1}^w C_{i,4} \bmod n^2)$, without learning the content of the data aggregation results. Given the ciphertext pairs $(C_{d,1}, C_{d,2})$ and $(C_{e,1}, C_{e,2})$, the traffic management authority can recover the content of $\sum_{i=1}^w s_{i,d}$ and $\sum_{i=1}^w s_{i,e}$ with the private key x , and it can re-encrypt the ciphertext pairs with secret value x_0 to generate the ciphertext pairs $(C_{d,1}, C_{t,1})$ and $(C_{e,1}, C_{t,2})$. For the traffic management authority, it can only achieve the aggregated sensory data, while the individual sensory data report is kept from the traffic management authority, and the segments contained in an individual data report are protected. Thus, the location privacy of each data uploading vehicle can be preserved during the sensory data collection phase.

The proposed sensory data sharing scheme is location privacy-preserving during the sensory

data acquisition phase. During the sensory data acquisition phase, a data querying vehicle V_a encrypts the segment identifier (u_c, v_c) with the public key (n, g, h_1) of the modified Paillier cryptosystem and a random number s_a , before sending towards the RSU. The segment identifier (u_c, v_c) contained in $(C_{a,1}, C_{a,2}, C_{a,3}, C_{a,4})$ is also semantic secure since $(C_{a,1}, C_{a,2})$ and $(C_{a,3}, C_{a,4})$ are also valid ciphertext pairs of the modified Paillier cryptosystem. Without the private key $x_1 + x_2$, the RSU cannot recover u_c and v_c contained in $(C_{a,1}, C_{a,2})$ and $(C_{a,3}, C_{a,4})$. In case of collusion, even though the colluding entities can obtain the value of $(x_1 + x_2)$, without the secret random value s_a , (u_c, v_c) cannot be fully recovered. Due to the homomorphic multiplicative property between one ciphertext and one plaintext, and the homomorphic additive property of the modified Paillier cryptosystem, the RSU calculates the ciphertext pair of $m_c = u_c \cdot t_1 + v_c \cdot t_2 \bmod n$, which is $((C_{a,1})^{t_1} \cdot (C_{a,3})^{t_2} \bmod n^2, (C_{a,2})^{t_1} \cdot (C_{a,4})^{t_2} \bmod n^2)$. The RSU also re-encrypts the derived ciphertext pair with the secret key x_1 and generates the ciphertext $C_{r,1} = (C_{a,1})^{t_1} \cdot (C_{a,3})^{t_2} \bmod n^2$ and $C_{r,2} = \frac{(C_{a,2})^{t_1} \cdot (C_{a,4})^{t_2}}{(C_{r,1})^{x_1}} \bmod n^2$, such that $(C_{r,1}, C_{r,2})$ can only be recovered by the data querying vehicle with x_2 and s_a . Meanwhile, the RSU generates a secret key β_l for each segment $l, l = 1, 2, \dots, k$ through computing $\beta_l = H(u_l \cdot t_1 + v_l \cdot t_2)$, and then it structures $(\beta_1, \beta_2, \dots, \beta_k)$ with $(\alpha_1, \alpha_2, \dots, \alpha_k)$ to construct $s_k = \sum_{i=1}^k \beta_i \cdot \alpha_i$. By exploiting the homomorphic additive property of the modified Paillier cryptosystem, the RSU generates the ciphertext pairs of $\sum_{i=1}^w s_{i,d} + s_k$ and $\sum_{i=1}^w s_{i,e} + s_k$ respectively. The RSU also re-encrypts the derived ciphertext pairs with the secret key x_1 to generate $(C_{r,3}, C_{r,4})$ and $(C_{r,5}, C_{r,6})$, such that they can be decrypted with x_2 .

After receiving $(C_{r,1}, C_{r,2}, C_{r,3}, C_{r,4}, C_{r,5}, C_{r,6})$, only V_a can recover the value of $\sum_{i=1}^w s_{i,d} + s_k$ and $\sum_{i=1}^w s_{i,e} + s_k$. Since $(C_{r,1}, C_{r,2})$ is a valid ciphertext pair of the modified Paillier cryptosystem, V_a can obtain the value of $m_c = u_c \cdot t_1 + v_c \cdot t_2 \bmod n$ with $x_2 + s_a$. Although V_a can recover the value of $m_c = u_c \cdot t_1 + v_c \cdot t_2 \bmod n$, it still cannot obtain the value of (t_1, t_2) . Without (t_1, t_2) , V_a cannot recover the secret keys of the unqueried segments, which prevents V_a from learning the sensory data generated by the unqueried sensory data reports. Given $(C_{r,3}, C_{r,4}, C_{r,5}, C_{r,6})$, even though all the vehicles can decrypt $(\sum_{i=1}^w s_{i,d} + s_k, \sum_{i=1}^w s_{i,e} + s_k)$ with private key x_2 , without s_k , they still cannot learn the content of the sensory data in each

FIGURE 4.4. Computation complexity of data server. (a) Computation complexity of the data server with the traditional scheme; (b) Computation complexity of the data server with the proposed scheme.

data dimension. Even when V_a is one of the colluding entities and has obtained the value of x_1 , it still cannot obtain the content of the unqueried locations without (t_1, t_2) . Thus, the security goal of location privacy preservation can be achieved during the sensory data acquisition phase.

The proposed data sharing scheme can achieve the security goal of data integrity protection. To protect data integrity, each data transmission flow between an RSU and a vehicle is attached with a message authentication code, which can be validated to guarantee the correctness of the data transmission. Thus, the security goal of data integrity protection can be achieved in the proposed scheme.

4.6 Performance Evaluation

In this section, we evaluate the performance of the proposed privacy-preserving data sharing scheme, in terms of the computation costs at the RSU and the vehicles, and the communication overhead between the RSU and the vehicles.

4.6.1 Parameter Setup

We evaluate the computation costs of the proposed scheme on a desktop with Intel i7-3770 3.4 GHz processor, 8 GB RAM, and Window 7 platform. Meanwhile, we evaluate the performance by implementing the proposed scheme with Java, and the detailed parameter settings are shown in Table 4.2. Given the parameter settings in Table 4.2, when $\kappa_1 = 24$, it has $|d| = 22$, and the maximum allowable number of data uploading vehicles in a district is 2^{13} , which fully satisfies the scalability of the data uploading vehicles.

TABLE 4.2. Parameters Settings

Simulation Parameters	Value
κ, κ_1	$\kappa = 512, \kappa_1 = 24$
$p_1, p_2, n = p_1 p_2$	$ p_1 = p_2 = 512, n = 1024$
$q_l, l = 1, 2, \dots, k$	$ q_l = \kappa_1 = 24$
k	$k = 40$: Number of Segments
$d_{i,l}, i = 1, 2, \dots, w, l = 1, 2, \dots, k$	$ d_{i,l} = 8$: Sensory Data Report

During the data collection phase, we compare the proposed scheme with the traditional homomorphic data aggregation scheme, when the modified Paillier cryptosystem is exploited. In the traditional data aggregation scheme, each dimension of data is encrypted separately, i.e., during the data report generation phase, the ciphertexts of V_i can be derived as $(C_{i,1,l}, C_{i,2,l}, C_{i,3,l}, C_{i,4,l}), l = 1, 2, \dots, k$, where $C_{i,1,l} = g^{r_{i,d,l}} \bmod n^2$, $C_{i,2,l} = h^{r_{i,d,l}} \cdot (1 + d_{i,l} \cdot n) \bmod n^2$, $C_{i,3,l} = g^{r_{i,e,l}} \bmod n^2$, $C_{i,4,l} = h^{r_{i,e,l}} \cdot (1 + e_{i,l} \cdot n) \bmod n^2$ and $(r_{i,d,l}, r_{i,e,l}) \in \mathbb{Z}_{n^2}$. Meanwhile, at the RSU and the traffic management authority side, each data report dimension is aggregated and decrypted individually.

During the data querying phase, we compare the proposed scheme with the traditional oblivious transfer scheme, when the modified Paillier cryptosystem is exploited. In the traditional oblivious transfer scheme, the data query sent by the data requester is still in the format of $(C_{a,1}, C_{a,2}, C_{a,3}, C_{a,4}, C_{a,5})$, while the data query response is in the format of $(C_{r,0}, C_{r,1}, C_{r,2})$ and $(C_{r,3,l}, C_{r,4,l}, C_{r,5,l}, C_{r,6,l}), l = 1, 2, \dots, k$.

4.6.2 Computation Complexity

For the proposed data sharing scheme, during the data collection phase, to generate an encrypted data report $C_{i,1}||C_{i,2}||C_{i,3}||C_{i,4}$, V_i needs to perform 4 exponentiation operations in $\mathbb{Z}_{n^2}$ and 4 multiplication operations in $\mathbb{Z}_{n^2}$. Then the traffic management authority performs 2 exponentiation operations in $\mathbb{Z}_{n^2}$ and 4 multiplication operations in $\mathbb{Z}_{n^2}$ to decrypt $\sum_{i=1}^w s_{i,d}$ and $\sum_{i=1}^w s_{i,e}$. During the data acquisition phase, V_a performs 5 exponentiation operations in $\mathbb{Z}_{n^2}$ and 4 multiplication operations in $\mathbb{Z}_{n^2}$ to generate $C_{a,0}||C_{a,1}||C_{a,2}||C_{a,3}||C_{a,4}$. After receiving $C_{a,0}||C_{a,1}||C_{a,2}||C_{a,3}||C_{a,4}$, the RSU takes 7 exponentiation operations in $\mathbb{Z}_{n^2}$ and 4 multiplication operations in $\mathbb{Z}_{n^2}$ to generate $C_{r,0}||C_{r,1}||C_{r,2}$, and it takes 6 exponentiation operations in $\mathbb{Z}_{n^2}$ and 10 multiplication operations in $\mathbb{Z}_{n^2}$ to generate $C_{r,3}||C_{r,4}||C_{r,5}||C_{r,6}$. After receiving $C_{r,0}||C_{r,1}||C_{r,2}||C_{r,3}||C_{r,4}||C_{r,5}||C_{r,6}$, the vehicle takes 4 exponentiation operations in $\mathbb{Z}_{n^2}$ and 7 multiplication operations in $\mathbb{Z}_{n^2}$ to retrieve m_c , $\sum_{i=1}^w s_{i,d} + s_k$, and $\sum_{i=1}^w s_{i,e} + s_k$.

We compare the proposed scheme with the traditional data aggregation scheme during the data collection phase, in which each vehicle generates a ciphertext for each data dimension during the data collection phase. Specifically, under this setting, for a k -dimension data report, it takes $4 * k$ exponentiation operations in $\mathbb{Z}_{n^2}$ and $4 * k$ multiplication operations in $\mathbb{Z}_{n^2}$ to generate $(C_{i,1,l}, C_{i,2,l}, C_{i,3,l}, C_{i,4,l})$, $l = 1, 2, \dots, k$. Then the traffic management authority performs $2 * k$ exponentiation operations in $\mathbb{Z}_{n^2}$ and $4 * k$ multiplication operations in $\mathbb{Z}_{n^2}$ for the sensory data decryption. During the data acquisition phase, we compare the proposed scheme with the traditional oblivious transfer protocol. V_a also performs 5 exponentiation operations in $\mathbb{Z}_{n^2}$ and 4 multiplication operations in $\mathbb{Z}_{n^2}$ to generate $C_{a,0}||C_{a,1}||C_{a,2}||C_{a,3}||C_{a,4}$. After receiving $C_{a,0}||C_{a,1}||C_{a,2}||C_{a,3}||C_{a,4}$, the RSU performs the $7 + 6 * k$ exponentiation operations in $\mathbb{Z}_{n^2}$ and $4 + 10 * k$ multiplication operations in $\mathbb{Z}_{n^2}$ to generate the queried data response $C_{r,0}||C_{r,1}||C_{r,2}$ and $C_{r,3,l}||C_{r,4,l}||C_{r,5,l}||C_{r,6,l}$, $l = 1, 2, \dots, k$. Then the data querying vehicle spends 4 exponentiation operations in $\mathbb{Z}_{n^2}$ and 7 multiplication operations in $\mathbb{Z}_{n^2}$ to derive the data query results.

We run our experiments 1000 times, in the proposed scheme, during the data collection phase, the average computation cost of a vehicle is 67.41 ms and the average computation cost of the

FIGURE 4.5. Communication cost of Vehicle-to-RSU during data collection. (a) Cost of the proposed scheme; (b) Cost of the traditional scheme.

FIGURE 4.6. Communication cost of Vehicle-to-RSU during data acquisition. (a) Cost of the proposed scheme; (b) Cost of the traditional scheme.

traffic management authority is 34.59 ms. In the proposed scheme, during the data acquisition phase, the average computation cost of a vehicle for data query generation is 84.25 ms and for data query response recovery is 68.23 ms; and the average computation cost of the RSU for the generation time of $C_{r,0}||C_{r,1}||C_{r,2}$ is 60.17 ms and for the generation time $C_{r,3}||C_{r,4}||C_{r,5}||C_{r,6}$ is 69.02 ms. In Figure. 4.4a, we plot and compare the computation cost of the vehicle and the traffic management authority in both schemes in terms of the data dimension, which varies from 4 to 40, during the data collection phase.

4.6.3 Communication Overhead

In the proposed scheme, during the sensory data report collection phase, the vehicle-to-RSU communication overhead of $C_{i,1}||C_{i,2}||C_{i,3}||C_{i,4}||\sigma_i$ is $2048 * 4 + 1024$ bits, if we choose 1024-bit n . If the RSU collects the w data reports from w data uploading vehicles, and the overall vehicle-to-RSU communication overhead is $2048 * 4 * w + 1024 * w$ bits. During the sensory data acquisition phase, the vehicle-to-RSU communication overhead $C_{a,1}||C_{a,2}||C_{a,3}||C_{a,4}||C_{a,5}||\sigma_a$ is $2048 * 5 + 1024$ bits; meanwhile, the size of the RSU-to-vehicle communication overhead $C_{r,0}||C_{r,1}||C_{r,2}||C_{r,3}||C_{r,4}||C_{r,5}||C_{r,6}||\sigma_r$ is $2048 * 7 + 1024$ bits.

For the traditional scheme, the overall vehicle-to-RSU communication overhead during the data report collection phase is $2048 * 4 * w * k + 1024 * w$ bits if we choose 1024-bit n . During the data acquisition phase, the vehicle-to-RSU communication overhead of data query is $2048 * 5 + 1024$ bits, and the RSU-to-vehicle communication overhead of the data query in the traditional scheme is $3 * 2048 + 2048 * 4 * k + 1024$ bits.

As shown in Figure 4.5, during the sensory data collection phase, the communication overhead (including both data query and data query response) of the proposed scheme only increases in terms of the number of data reports, while the communication overhead of the traditional scheme increases in terms of both the number of data dimension and the number of data reports. Thus, the proposed scheme can greatly reduce the communication overhead during the sensory data collection phase.

As shown in Figure 4.6, during the sensory data acquisition phase, the communication overhead (both data query and data query response) of the proposed scheme does not increase with the number of data dimension, while the communication overhead of the traditional scheme increases with the number of data dimension. Thus, the proposed scheme can greatly reduce the communication overhead during the sensory data acquisition phase.

4.7 Summary

In this chapter, we have proposed an efficient and location privacy-preserving sensory data sharing scheme in IoV. By exploiting the modified Paillier Cryptosystem, the proposed scheme can achieve the location privacy-preserving multi-dimensional sensory data aggregation, while counting the number of data reports in each data dimension. Meanwhile, the proposed scheme also achieves the location privacy-preserving data acquisition at the network edge, i.e., the RSU re-encrypts the sensory data aggregation results into ciphertexts which can only be decrypted by the data querying vehicle; meanwhile, the data querying location and the unqueried sensory data are still protected. Numerical analysis are performed to demonstrate the effectiveness of the proposed scheme in terms of the low data querying failure probability. Security analysis are conducted to validate the security properties, and experiments are also performed to show its efficiency in terms of the low computation cost and low communication overhead. However, in the proposed scheme, each data query can still be linked to a unique vehicle, which violates the identity privacy of the vehicle.

Chapter 5

A Privacy-preserving and Verifiable Querying Scheme in Vehicular Fog Data Dissemination

In this chapter, a secure vehicular data querying scheme in vehicular data dissemination is proposed, in comparison with the security schemes proposed in the previous chapters, each data request cannot be associated with a specific vehicle. Specifically, the proposed scheme introduces an invertible matrix to structure multiple data requests, and encrypts the data requests with the homomorphic Paillier Cryptosystem technique. For the data communication between the vehicles and the RSU, the data aggregation is performed on the ciphertext at the RSU without decryption. Based on the aggregation result, the RSU can recover the data requests while protecting the unlinkability between the data query and its origin. In addition, the RSU can verify the correctness of the recovered data requests by exploiting batch verification, without violating the privacy of the data requesting vehicles. Through security analysis, we demonstrate that the proposed scheme can achieve the security goals of unlinkability, verifiability, and confidentiality. Through performance evaluation, we demonstrate that the proposed scheme can adapt to the fluctuating number of the data querying vehicles, and significantly reduce the computation complexity and communication overhead.

5.1 Introduction

With the recent development of automotive industry and ITS, various data dissemination applications have been designed in VANETs for the purposes safety enhancement, driving assistance and entertainment services [162]. In vehicular fog, the RSUs can act as fog storage devices to maintain and distribute a plethora of local information, which greatly reduces the content delivery latency and saves the transmission bandwidth [163–165]. In the existing schemes, the RSUs can disseminate information towards vehicles in either a proactive manner, or a request-response manner [10, 166, 167]. For the request-response manner, vehicles send requests towards the RSUs to acquire the intended contents, and the RSUs disseminate the required data towards vehicles as a response. Although the request-response based approach can help vehicles to retrieve information at the network edge with high efficiency, there still exists some challenges in the aspect of security and privacy preservation.

For the request-response based data dissemination, the data requests uploaded by vehicles are tightly connected to the drivers and passengers on-board. If a data request can be uniquely identified, some privacy-related information could be disclosed from the data request (such as hospitals or churches) and these privacy-sensitive information must be strictly prohibited from unauthorized access. By exploiting the pseudonym-based schemes [63, 95], provably good privacy in VANETs can be achieved, however, an adversary can still associate a data request with a particular vehicle, and further deduce the trajectory and behavioral patterns through observation.

When there exist a group of vehicles sending data requests simultaneously, it is highly possible that the data request generated by one specific vehicle is indistinguishable from the rest of the data requests, i.e., disrupting the association between a vehicle and a data request. However, if multiple data requests are processed independently, it could bring heavy computation complexity towards the RSU, especially when there involves a large number of data querying vehicles. Thus, the data requests should be processed aggregately and then recovered; meanwhile, the processed scheme should also be adaptive to the fluctuation of the data requesting vehicles in

the spatio-temporal domain. A privacy-preserving multi-dimensional data aggregation scheme with the homomorphic Paillier cryptosystem is proposed in [63], which structures the multi-dimensional data report by exploiting a super-increasing sequence. A privacy-preserving route reporting aggregation scheme in VANET is proposed in [168], which utilizes the homomorphic encryption technique to calculate the total number of vehicles in each route segment. However, the above-mentioned schemes can only calculate the aggregation result, but they are not able to recover each individual data report.

In addition, the correctness of the recovered data requests should also be verified with privacy preservation. A verifiable data sharing scheme is proposed in [169], which exploits the identity-based signature (IBS) technique to achieve the distributed verification of the involved identities. Moreover, an efficient verifiable data aggregation scheme exploiting the certificate-less aggregate signature is also proposed in [170], which verifies the correctness of the big data obtained from IoT terminals and effectively achieves the computation reduction at the IoT data center. However, in our scenario, when the plaintext space of the data requests is limited, given the signature, the content of each individual data request can still be violated under the exhaustive attack. Therefore, how to verify the correctness of the recovered data request while preserving the privacy of the data querying vehicle, have become an emerging challenge in the request-response based data dissemination in vehicular fog.

In this chapter, to solve the above problems, we propose a privacy-preserving data dissemination scheme in vehicular fog, which enables the vehicles (end users) to acquire local data from the RSU (fog storage device) in a request-response paradigm. Specifically, the main contributions of the chapter are three folds.

- Firstly, the proposed scheme exploits the an invertible matrix to structure multiple data requests, such that each individual data request can be recovered by the RSU without being associated with its origin vehicle. In addition, the proposed scheme owns the benefit of scalability, which can be adaptive to the scenario when there involves a large number of data querying vehicles.

- Secondly, the proposed scheme achieves the efficient batch verification of the recovered data requests. With an identity-based verification scheme, the RSU can verify the correctness of the recovered data requests, without the disclosure the origin of each individual data request.
- Thirdly, we demonstrate the security properties of the proposed data dissemination scheme, in terms of unlinkability, confidentiality and verifiability. Meanwhile, we perform extensive simulation to demonstrate the efficiency of the proposed scheme, i.e., computation complexity and communication overhead.

The rest of this chapter is organized as follows. We introduction our system model, security requirements, and identify our design goal in Section 5.2. We propose our scheme in Section 5.4, followed by security analysis and performance evaluation in Section 5.5 and Section 5.6, respectively. Finally, we draw our conclusion in Section 5.7.

5.2 Model, Requirements and Design Goal

In this section, we describe system model, show security requirements, and identify design goals.

5.2.1 System Model

In the system model, we consider a data dissemination system with the request-response paradigm in vehicular fog, in which vehicles located on the parking lot send data requests towards the RSU (fog storage device). After receiving the uploaded data requests, the RSU distribute the queried data as a response. Specifically, the proposed system model consists of three entities: data querying vehicles, an RSU and an information server, as shown in Fig. 5.1.

Each data querying vehicle acts as an end-user, which is equipped with an OBU. It generates the data requests, analyzes the data responses and communicates with other involved entities, i.e., RSUs and other vehicles.

The RSU under consideration acts as the fog storage device, which stores a catalog of local information (such as local emergency contacts, local hospitals, public libraries, bank services, and so on), and it distributes the local information towards the vehicles under its coverage upon requesting. Meanwhile, the data items kept in the local storage of the RSU are periodically updated by the backbone information server.

The information server is a trustable central entity, which initializes the whole system and is responsible for the registration of the RSUs and vehicles. It also collects the local information and periodically updates the information items stored in RSUs.

FIGURE 5.1. Vehicular data querying architecture under consideration

Communication Model. The V2V communication and the V2I communications are realized through the IEEE 802.11p WAVE standard [145], which is a short to medium range communication technology operating at 5.9 GHz band. The connection between the RSU and the information server is realized through secure wired link with high bandwidth and low transmission delay.

5.2.2 Security Requirements

In our security model, we assume the RSU is honest-but-curious, i.e., it will correctly execute the operations defined in the protocol and disseminate the queried data as responses, but it may try to violate the privacy of the data querying vehicles through linking a particular recovered data query with its origin vehicle; meanwhile, we assume neither the RSU nor the data querying vehicles will collude with each other in the proposed scheme. Moreover, there exists an

adversary, which may eavesdrop the data transmission and launch some active attacks to threaten the data integrity. In addition, some vehicles may send data requests towards the RSU without registration. Therefore, to preserve the privacy of the data querying vehicles and to protect the confidentiality and correctness of the data requests, the following security requirements should be satisfied in the proposed scheme.

Unlinkability. At the RSU side, the RSU can only learn the contents of the received data requests, but it cannot link a particular data request to the corresponding origin vehicle. Meanwhile, each involved vehicle should not be able to associate a given data request with a particular vehicle. In other words, the proposed scheme should not disclose any information about the association between a vehicle and a data request [171]. The unlinkability arises primarily from the need to protect the vehicle's anonymity, i.e., preserving vehicle's privacy by removing the vehicle's identity from its data request [172].

Confidentiality. In the proposed scheme, the content of the data requests and the data dissemination response should be protected from the adversary. Even if the adversary eavesdrops the data transmission, it cannot recover the contents of the data requests and the data dissemination. In addition, the confidentiality requirement should also include that the data dissemination can only be learnt by the involved vehicles which have already uploaded data requests.

Verifiability. The RSU should be able to verify that the recovered data request has not been modified during data processing and transmission, and authenticate that each received data request is actually generated and signed by a vehicle with registration. Meanwhile, the coordinating vehicle should also be able to verify that the received data request has not been modified during data processing.

5.2.3 Design Goals

Under the aforementioned system model and security requirements, our design goal is to develop an efficient and privacy-preserving data dissemination scheme with request-response paradigm in vehicular fog. Specifically, the following three objectives should be achieved.

The proposed scheme should achieve the above security requirements. If the proposed scheme does not take the security requirements into consideration, the privacy of the data querying vehicles could be violated, the confidentiality and correctness of the data requests could be threatened, and the correctness of the recovered data requests cannot be guaranteed. Then the vehicles may not be willing to join in the data dissemination process, and data dissemination system cannot be properly functioned.

The proposed scheme should achieve the goal of scalability. Since the number of data querying vehicles on the parking lot varies in the spatial temporal domain, the proposed scheme should be adaptive to the fluctuating number of vehicles, especially be capable of functioning in the scenario when there involves a large number of data querying vehicles.

The proposed scheme should achieve high efficiency in terms of computation and computation cost. Even though the fog storage device RSU possesses high processing capability, the proposed scheme should also take the computation overhead into consideration, especially when there are a large number of data querying vehicles. Due to scarce and limited transmission bandwidth, the proposed scheme should also take the communication overhead into consideration.

5.3 Preliminaries

In this section, we briefly review the security techniques of Paillier cryptosystem and bilinear maps, which function as the basis of the proposed scheme in the aspects of encryption and verification respectively.

5.3.1 Paillier Cryptosystem

In our proposed scheme, the Paillier Cryptosystem is the building block [143], which has been widely exploited due to its homomorphic additive properties. Specifically, the Paillier Cryptosystem consists of three algorithms: key generation, encryption and decryption.

- *Key Generation:* Given the security parameter κ , two large prime numbers p_1, q_1 are first chosen, where $|p_1| = |q_1| = \kappa$; meanwhile, the RSA modulus $n = p_1 q_1$ and $\lambda = \text{lcm}(p_1 - 1, q_1 - 1)$ are also calculated. After choosing a generator $g \in \mathbb{Z}_{n^2}^*$, a function $L(u) = \frac{u-1}{n}$ is defined, and $\mu = (L(g^\lambda \bmod n^2))^{-1} \bmod n$ is further computed. Then, the public key $pk = (n, g)$, and the corresponding private key $sk = (\lambda, \mu)$ are derived.
- *Encryption:* Given a message $m \in \mathbb{Z}_n$, choose a random number $r \in \mathbb{Z}_n^*$, and the ciphertext can be generated as $c = E(m) = g^m \cdot r^n \bmod n^2$.
- *Decryption:* Given the ciphertext $c \in \mathbb{Z}_{n^2}^*$, the corresponding message can be recovered as $m = D(c) = L(c^\lambda \bmod n^2) \cdot \mu \bmod n$.

5.3.2 Bilinear Maps

Let $\mathbb{G}_1$ and $\mathbb{G}_2$ be two cyclic groups with the same prime order q , i.e., $|\mathbb{G}_1| = |\mathbb{G}_2| = q$. A bilinear map $e: \mathbb{G}_1 \times \mathbb{G}_1 \rightarrow \mathbb{G}_2$ has the following properties:

- 1) Bilinearity: $\forall P, Q \in \mathbb{G}_1$, and $\forall a, b \in \mathbb{Z}_q^*$, we can derive $e(aP, bQ) = e(P, Q)^{ab}$;
- 2) Nondegeneracy: There exist $P \in \mathbb{G}_1$, which satisfies the condition that $e(P, P) \neq 1_{\mathbb{G}_2}$.
- 3) Computable: $\forall P, Q \in \mathbb{G}_1$, there is an efficient algorithm to compute $e(P, Q)$.

A bilinear parameter generator $\mathcal{Gen}$ denotes a probabilistic algorithm that takes a parameter κ_1 as input, and outputs a 5-tuple $(q, P, \mathbb{G}_1, \mathbb{G}_2, e)$ as the output, where q is κ_1 -bit prime, $\mathbb{G}_1$ and $\mathbb{G}_2$ are two cyclic groups with order q , $P \in \mathbb{G}$ is a generator, and $e: \mathbb{G}_1 \times \mathbb{G}_1 \rightarrow \mathbb{G}_2$ is non-degenerated and computable bilinear map.

5.4 Proposed Verifiable Data Querying and Dissemination Scheme

In this section, we propose the privacy-preserving vehicular data query scheme, which consists of four parts: system initialization, group formulation, data query generation, and data query

TABLE 5.1. Notations of Proposed Querying and Dissemination Scheme

Notation	Definition
p_1, q_1	Parameters of Paillier Cryptosystem
$pk = (n, g)$	Public Key of Paillier Cryptosystem
$sk = (\lambda, \mu)$	Secret Key of Paillier Cryptosystem
$q, g, \mathbb{G}, \mathbb{G}_T, e, h$	Parameters of Bilinear Groups
$\mathbf{A}$	Invertible Matrix for Data Requests Structuring
p_2	Length of each Matrix Element

aggregation and recovery. Meanwhile, the definitions of some important notations are described in Table 5.1.

5.4.1 System Initialization

For the vehicular data querying system under consideration, the information server, which also acts as the TA, will bootstrap the whole system. Given a security parameter κ , TA selects two large prime numbers p_1 and q_1 , where $|p_1| = |q_1| = \kappa$. Then TA calculates the Paillier Cryptosystem's public key $(n = p_1 q_1, g)$, and the corresponding private key (λ, μ) . Meanwhile, given security parameter κ_1 , the TA generates the bilinear parameters $(q, P, \mathbb{G}_1, \mathbb{G}_2, e(\cdot, \cdot))$ by running $\mathcal{Gen}(\kappa_1)$. The TA selects a random number $x \in \mathbb{Z}_q^*$ as the secret key, and computes the public key $P_0 = x \cdot P$. In addition, the TA selects a secure encryption algorithm $\mathcal{Enc}(\cdot)$, and chooses two secure cryptographic hash functions $H(\cdot), H_1(\cdot)$, where $H : \{0, 1\}^* \rightarrow \mathbb{Z}_n^*, H_1 : \{0, 1\}^* \rightarrow \mathbb{Z}_q^*$. The TA also selects a large prime number α with the length of $|\alpha| = \kappa$, and chooses a constant number d with $|d| < \kappa$. Then the TA publishes the system parameter as $params = \{n, g, q, P, P_0, \mathbb{G}_1, \mathbb{G}_2, e(\cdot, \cdot), h = e(P, P), \mathcal{Enc}(\cdot), H(\cdot), H_1(\cdot), \alpha\}$.

FIGURE 5.2. Key distribution

As shown in Fig. 5.2, during the registration of each RSU, the TA delivers the private key (λ, μ) towards the it. Meanwhile, during the registration of a vehicle with identity $ID_i \in \{0, 1\}^*$, the

TA first computes the secret key $SK_{ID_i} = \frac{1}{h_1(ID_i)+x} \cdot P$, and securely distributes SK_{ID_i} towards the vehicle.

5.4.2 Group Formulation

When a vehicle V_a intends to query the RSU for local information, it acts as a coordinating and broadcasts a cooperation request towards all the vehicles on the parking lot to send data queries together, so as to prevent itself from being associated with its data query. If a vehicle is interested in joining the collaborative data querying process, it accepts the cooperation request, and establishes a secure session key with V_a . Thus, a cooperative data querying group with k vehicles can be formulated.

FIGURE 5.3. Secret distribution among vehicles

V_a selects a large prime p_2 , where $k \cdot p_2 < d$, and generates a random invertible matrix $\mathbf{A}$ with the size of $k \times k$, where each element $a_{i,j} \in GF(p_2), i = 1, \dots, k, j = 1, \dots, k$. V_a also selects a session key k_d for the queried data encryption. Then V_a performs the secret distribution process, which can be illustrated in Fig. 5.3. V_a assigns a row vector $\mathbf{a}_i$ ($i \neq a$) towards a collaborating vehicle V_i , and allocates a random number $s_i \in \mathbb{Z}_n^*$ for V_i . (Note that one row vector can only be allocated once. Meanwhile, $\mathbf{a}_i$ only denotes the row vector allocated to V_i , and it does not necessarily be the i th column vector of $\mathbf{A}$.) Then V_a securely transmits $s_i || \mathbf{a}_i || p_2 || k_d || TS$ towards V_i , as shown in Fig. 5.3 (*Message (1)*), where TS is the current timestamp.

V_a also randomly selects a subcoordinating vehicle within the collaborating vehicles, which is denoted as V_b , and acknowledges the rest of vehicles to establish a secure session key with V_b .

After key establishment, V_b chooses a random number $t_i \in \mathbb{Z}_n^*$ ($i \neq a, b$) for each vehicle V_i , and securely delivers $t_i || TS$ towards V_i , as shown in Fig. 5.3 (Message (2)).

5.4.3 Data Query Generation

5.4.3.1 Data Query Generation of V_i

FIGURE 5.4. Vehicular data querying message flow

for each vehicle V_i ($i \neq a$), it first chooses the needed data query identifier Q_i ($|Q_i| = l_q$ bits) corresponding to the queried data item, from the predefined data query space. Then V_i performs the following steps to generate a data query.

Step-1: To prevent the exhaustive attack, V_i selects a random number x_i with the fixed length of l_x bits ($l_q + l_x < |p_2|$), and concatenates Q_i with x_i to generate $m_i = Q_i || x_i$. Then V_i multiplies m_i with each element in the received row vector $\mathbf{a}_i = (a_{i,1}, a_{i,2}, \dots, a_{i,k})$, and generates $\mathbf{d}_i = (d_{i,1}, d_{i,2}, \dots, d_{i,k})$, where $d_{i,j} = m_i \cdot a_{i,j} \bmod p_2, j = 1, \dots, k$.

Step-2: V_i chooses a random number $r_{i,1} \in \mathbb{Z}_n^*$, and computes the ciphertext

$$C_{i,1} = g^{d_{i,1} + \alpha \cdot d_{i,2}} \cdot r_{i,1}^n \cdot H(TS)^{s_i} \bmod n^2, \quad (5.1)$$

which prevents the RSU from recovering the value of $d_{i,1} + \alpha \cdot d_{i,2}$. Then V_i generates the message authentication code $MAC_{i,1} = H(C_{i,1} || TS)$ corresponds to $C_{i,1}$. In addition, V_i ($i \neq a, b$) generates a verifiable signature $\sigma_i = (\sigma_{i,1}, \sigma_{i,2})$ corresponds to m_i , by choosing a random

number $r_{i,2} \in \mathbb{Z}_q^*$ and calculating

$$\begin{cases} \sigma_{i,1} = h^{r_{i,2}}, \\ \sigma_{i,2} = \frac{H_1(m_i||TS) + H_1(s_i||TS) + H_1(t_i||TS) + r_{i,2}}{x + H_1(ID_i)} \cdot P \end{cases} \quad (5.2)$$

For V_b , it also selects a random number $r_{b,2} \in \mathbb{Z}_q^*$, and generates the signature $\sigma_b = (\sigma_{b,1}, \sigma_{b,2})$, which is

$$\begin{cases} \sigma_{b,1} = h^{r_{b,2}}, \\ \sigma_{b,2} = \frac{H_1(m_b||TS) + H_1(s_b||TS) - \sum_{i \neq b} H_1(t_i||TS) + r_{b,2}}{x + H_1(ID_b)} \cdot P \end{cases} \quad (5.3)$$

Finally, V_i transmits $ID_i||C_{i,1}||MAC_{i,1}||\sigma_i||TS$ towards the RSU, as shown in Fig. 5.4 (*Message (1)*).

Step-3: For V_i ($i \neq a, b$), it calculates $e_{i,j} = d_{i,j} + H_1(t_i||j||TS) \bmod p_2, j = 3, \dots, k$ and generates a vector $\vec{e}_i = (e_{i,3}, \dots, e_{i,k})$. Then it encrypts $\vec{e}_i$ with the secret value s_i to generate the ciphertext, which is

$$C_{i,2} = \mathcal{Enc}_{H(s_i||TS)}(e_{i,3}||\dots||e_{i,k}) \quad (5.4)$$

and transmits $C_{i,2}$ towards V_a , which can prevent V_a from directly recovering the value of $d_{i,j}$. Then V_i generates the message authentication code corresponds to $C_{i,2}$, which is

$$MAC_{i,2} = H(C_{i,2}||TS) \quad (5.5)$$

and transmits $C_{i,2}||MAC_{i,2}||TS$ towards V_a , as shown in Fig. 5.4 (*Message (2)*). For V_b , it also calculates $e_{b,j} = d_{b,j} - \sum_{i=1, i \neq a, b}^k H(t_i||j||TS) \bmod p_2, j = 3, \dots, k$ and generates a vector $\vec{e}_b = (e_{b,3}, \dots, e_{b,k})$. Then it generates the ciphertext with the secret value s_b , which is

$$C_{b,2} = \mathcal{Enc}_{H(s_b||TS)}(e_{b,3}||\dots||e_{b,k}) \quad (5.6)$$

V_b also generates the message authentication code corresponds to $C_{b,2}$, which is

$$MAC_{b,2} = H(C_{b,2}||TS) \quad (5.7)$$

and transmits $C_{b,2}||MAC_{b,2}||TS$ towards V_a , as shown in Fig. 5.4 (*Message (2)*).

5.4.3.2 Data Query Generation of V_a

for V_a , it also formulates the data query $m_a = Q_a||x_a$, and generates a row vector $\mathbf{d}_a = (d_{a,1}, d_{a,2}, \dots, d_{a,k})$ with *Step-1* defined above, and performs the following steps.

Step-1: V_a selects a random number $r_{a,1} \in \mathbb{Z}_n^*$, and generates the ciphertext

$$C_{a,1} = g^{d_{a,1} + \alpha \cdot d_{a,2}} \cdot r_{a,1}^n \cdot H(TS)^{s_a} \bmod n^2. \quad (5.8)$$

where $s_a = -\sum_{i \neq a} s_i \bmod n$. Meanwhile, V_a also chooses a random number $r_{a,2} \in \mathbb{Z}_q^*$, and generates the verifiable signature σ_a corresponds to m_a , which is

$$\begin{cases} \sigma_{a,1} = h^{r_{a,2}}, \\ \sigma_{a,2} = \frac{H_1(m_a||TS) - \sum_{i \neq a} H_1(s_i||TS) + r_{a,2}}{x + H_1(ID_a)} \cdot P. \end{cases} \quad (5.9)$$

Step-2: After receiving $C_{i,2}$ ($i \neq a$), V_a checks the correctness of the received message authentication codes, by checking whether $MAC_{i,2} \stackrel{?}{=} H(C_{i,2}||TS)$. If it verifies to be correct, V_a decrypts $C_{i,2}$ ($i \neq a$) with $H(s_i||TS)$, and obtains $e_{i,j}$, $j = 3, \dots, k$. Then V_a aggregates $e_{i,j}$ to achieve the value of $\sum_{i=1, i \neq a}^k d_{i,j}$, and updates the matrix elements $a_{a,j}$, $j = 3, \dots, k$, which is

$$a_{a,j} = \left(-\sum_{i=1, i \neq a}^k d_{i,j}\right) \cdot m_a^{-1} \bmod p_2. \quad (5.10)$$

Step-3: V_a encrypts the session key k_d for the data dissemination protection with the public key (n, g) , which is denoted as $C_k = g^{k_d} \cdot r_k^n \bmod n^2$ and $r_k \in \mathbb{Z}_n^*$ is a randomly chosen number. Then V_a generates the corresponding MAC, which is $MAC_a = H(C_{a,1}||C_k||\mathbf{A}||TS)$. Then V_1 transmits the $ID_a||C_{a,1}||\sigma_a||\mathbf{A}||C_k||MAC_a||TS$ towards the RSU, as shown in Fig. 5.4 (*Message (3)*).

5.4.4 Data Report Aggregation and Reading

After receiving $ID_a||C_{a,1}||\sigma_a||\mathbf{A}||C_k||MAC_a||TS$, the RSU first checks the correctness of MAC_a by computing $MAC_a \stackrel{?}{=} H(C_{a,1}||C_k||\mathbf{A}||TS)$. The RSU also verifies the correctness of $MAC_{i,1}$ by computing $MAC_{i,1} \stackrel{?}{=} H(C_{i,1}||TS)$. If they are verified to be correct, the RSU aggregates all the received data queries $C_{i,1}, i = 1, \dots, k$ and obtains the data aggregation result C_1 , which is

$$\begin{aligned}
 C_1 &= \prod_{i=1}^k C_{i,1} \bmod n^2 \\
 &= \prod_{i=1}^k g^{d_{i,1} + \alpha \cdot d_{i,2}} \cdot r_{i,1}^n \cdot H(TS)^{s_i} \bmod n^2 \\
 &= g^{\sum_{i=1}^k d_{i,1} + \alpha \cdot d_{i,2}} \cdot \left(\prod_{i=1}^k r_{i,1} \right)^n \cdot H(TS)^{\sum_{i=1}^k s_i} \bmod n^2 \\
 &= g^{\sum_{i=1}^k d_{i,1} + \alpha \cdot d_{i,2}} \cdot \left(\prod_{i=1}^k r_{i,1} \right)^n \bmod n^2
 \end{aligned} \tag{5.11}$$

By taking $M = \sum_{i=1}^k d_{i,1} + \alpha \cdot d_{i,2}$ and $R = \prod_{i=1}^k r_{i,1}$, the aggregated ciphertext $C = g^M \cdot R^n \bmod n^2$ is still a ciphertext of the Paillier Cryptosystem, and the RSU uses the private key (λ, μ) to recover $M = \sum_{i=1}^k d_{i,1} + \alpha \cdot d_{i,2}$. Then the RSU generates $\mathbf{y} = (y_1, y_2, \dots, y_k)$, in which

$$\begin{cases} y_1 = (M \bmod \alpha) \bmod p_2 = (\sum_{i=1}^k d_{i,1}) \bmod p_2, \\ y_2 = (\frac{M - y_1}{\alpha}) \bmod p_2 = (\sum_{i=1}^k d_{i,2}) \bmod p_2. \end{cases} \tag{5.12}$$

Meanwhile, the RSU sets $y_3 = \dots = y_k = 0$, computes the inverse of $\mathbf{A}$ in the field $GF(p_2)$ (denoted as $\mathbf{A}^{-1}$), and calculates $\mathbf{y} \cdot \mathbf{A}^{-1}$ to recover $\mathbf{m} = (m_1, \dots, m_k)$, without disclosing the actual association between each vehicle and its data query. Furthermore, the RSU abandons the last l_z bits in each m_i to obtain the data query identifiers $(Q_1, Q_2, \dots, Q_k)$.

5.4.5 Data Query Verification

To verify the correctness of $(m_1, \dots, m_k)$, the RSU exploits the received verifiable signatures $(\sigma_1, \dots, \sigma_k)$, and performs the batch verification process shown in Eq. (5.13). With the given verification process, the RSU only performs 2 pairing operations in $\mathbb{Z}_q^*$. If Eq. (5.13) verifies to be correct, the recovered data queries are verified to be correctly processed by V_a and the RSU. Meanwhile, the RSU decrypts C_k with its private key (λ, μ) , and obtains the session key k_d for the data transmission protection. Then the RSU identifies the data content queried by the vehicles m and formulates a ciphertext $C_m = \mathcal{Enc}_{H(k_d||TS)}(m)$ and the corresponding MAC: $MAC_m = H(C_m||TS)$. Moreover, the RSU broadcasts the data query response $C_m||MAC_m$ towards the data querying vehicles. After receiving $C_m||MAC_m$, each vehicle first verifies the correctness of MAC_m by checking $MAC_m \stackrel{?}{=} H(C_m||TS)$. If it verifies, each vehicle recovers the value of m with k_d . Based on the queried data m , each vehicle can identify the information queried by itself.

$$\begin{aligned}
& h^{\sum_{i=1}^k H_1(m_i||TS)} \cdot \prod_{i=1}^k \sigma_{i,1} \\
&= e(P, P)^{\sum_{i=1}^k H_1(m_i||TS) + r_{i,2}} \\
&= \prod_{i=1}^k e(P, \sigma_{i,2} \cdot (x + H_1(ID_i))) \\
&= \prod_{i=1}^k e(x \cdot P, \sigma_{i,2}) \cdot \prod_{i=1}^k e(P, H_1(ID_i) \cdot \sigma_{i,2}) \\
&= e(P_0, \sum_{i=1}^k \sigma_{i,2}) \cdot e(P, \sum_{i=1}^k H_1(ID_i) \cdot \sigma_{i,2})
\end{aligned} \tag{5.13}$$

Correctness. Now we prove that the updated matrix is still an invertible matrix. Let $\mathbf{B}$ be the inverse of the matrix $\mathbf{A}$ in the field $GF(p_2)$, $\mathbf{A}$ and $\mathbf{B}$ satisfy the following equation:

$$\begin{bmatrix} a_{1,1} & \dots & a_{1,k} \\ \vdots & \ddots & \vdots \\ a_{k,1} & \dots & a_{k,k} \end{bmatrix} \begin{bmatrix} b_{1,1} & \dots & b_{1,k} \\ \vdots & \ddots & \vdots \\ b_{k,1} & \dots & b_{k,k} \end{bmatrix} = \begin{bmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{bmatrix} \quad (5.14)$$

Suppose V_a occupies the x th row vector of $\mathbf{A}$. Then we calculate $e_j = -a_{x,j} - (\sum_{i=1, i \neq a}^k m_i a_{i,j}) \cdot m_a^{-1} \bmod p_2, j = 3, 4, \dots, k$, and update $a_{x,j}$ with $a_{x,j} + e_j \bmod p_2$, which is

$$\mathbf{A}' = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \dots & a_{1,x} & \dots & a_{1,k} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{x,1} & a_{x,2} & a_{x,3} + e_3 & \dots & a_{x,x} + e_x & \dots & a_{x,k} + e_k \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{k,1} & a_{k,2} & a_{k,3} & \dots & a_{k,x} & \dots & a_{k,k} \end{bmatrix} \quad (5.15)$$

Then we multiply the updated matrix $\mathbf{A}'$ with $\mathbf{B}$, and obtains

$$\mathbf{A}' \times \mathbf{B} = \begin{bmatrix} 1 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \sum_{i=3}^k e_i b_{i,1} & \dots & \sum_{i=3}^k e_i b_{i,x} + 1 & \dots & \sum_{i=3}^k e_i b_{i,k} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 1 \end{bmatrix} \quad (5.16)$$

As the product $\mathbf{A}' \times \mathbf{B}$ is an invertible matrix, $\mathbf{A}'$ is also an invertible matrix.

5.5 Security Analysis

In this section, we analyze the security properties of the proposed scheme in vehicular fog. In particular, following the security requirements discussed earlier, our analysis will focus on how the proposed scheme can prevent data requests from being associated with their origin vehicles, protect the confidentiality of data transmission, and achieve the verifiability of recovered data requests.

- *The individual vehicle's data request is unlinkable in the proposed scheme.* Since $C_{i,1}$ is formulated as $g^{d_{i,1} + \alpha \cdot d_{i,2}} \cdot r_i^n \cdot H(TS)^{s_i} \bmod n^2$ and we do not take the collusion attack into consideration, the RSU cannot directly recover the value of $d_{i,1} + \alpha d_{i,2}$ without learning the value of s_i . At the RSU side, by aggregating $C_1 = \prod_{i=1}^k C_{i,1} = g^{\sum_{i=1}^k d_{i,1} + \alpha \cdot d_{i,2}}$ and decrypting C_1 , the RSU can derive the values of $y_1 = \sum_{i=1}^k d_{i,1} \bmod p_2$ and $y_2 = \sum_{i=1}^k d_{i,2} \bmod p_2$. Then the RSU calculates $\mathbf{m} = \mathbf{y} \cdot \mathbf{A}^{-1}$ to recover the uploaded data requests $\mathbf{m} = (m_1, \dots, m_k)$. Since the RSU cannot achieve the information about which row vector in $\mathbf{A}$ corresponds to which data querying vehicle and we do not take the collusion attack into consideration, given the recovered data requests, it is impossible for the RSU to associate a particular data request with its origin vehicle. In addition, given σ_i and m_i , it is impossible to determine whether the data query m_i is contained in the signature σ_i by calculating $e(\sigma_{i,2}, P_0 + H_1(ID_i) \cdot P) \stackrel{?}{=} \sigma_{i,1} \cdot h^{H_1(m_i || TS)}$. This is because σ_i is protected with $H_1(s_i || TS) + H_1(t_i || TS)$, which are secrets shared with V_a and V_b respectively. Thus, the security goal of unlinkability can be achieved.

- *The confidentiality of the data transmission can be achieved in the proposed scheme.* In the proposed scheme, the ciphertext of V_i 's data request m_i is formulated as $(C_{i,1}, C_{i,2})$. Since the Paillier Cryptosystem is proved to be semantic secure against the chosen plaintext attack, the data request m_i contained in $C_{i,1}$ is also semantically secure and privacy-preserving. For the data request contained in $C_{i,2}$, it is encrypted with s_i , which is a secret value shared between V_a and V_i , and we do not take the collusion attack into consideration. As a result,

the adversary cannot recover the value of the data request m_i even if it eavesdrop the ciphertext $(C_{i,1}, C_{i,2})$. For V_a , after decrypting $C_{i,2}$ with $H(s_i||TS)$, it can obtain the value of $e_{i,j} = d_{i,j} + H(t_i||j||TS)$, $j = 3, \dots, k$. However, V_a still cannot recover the individual value of $d_{i,j}$, this is because t_i is a secret value shared between V_i and the coordinating vehicle V_b , and we do not take the collusion attack into consideration. In addition, the data query response from the RSU m is encrypted by a session key k_d , which is a session key shared among the RSU and the data querying vehicles, and the ciphertext C_m can only be decrypted by the data querying vehicles. Even if the adversary eavesdrops C_m , it still cannot learn the value of the data query response m . Thus, the security goal of confidentiality can be achieved in the proposed vehicular data dissemination scheme.

- *The verifiability of recovered data requests and the data integrity of data transmission can be achieved in the proposed scheme.* In the proposed scheme, the signature σ_i can be considered as a commitment of the data request m_i , i.e., σ_i preserves the correctness of m_i while hiding the real value of m_i . Meanwhile, the data request m_i is signed by the identity-based signature (IBS) scheme proposed in [173], which can also support batch verification. Since the security of the IBS scheme depends on the k -collision attack algorithm (k -CCA) complexity assumption, both the source authentication of the data querying vehicle V_i and the correctness verification of recovered data request m_i can be achieved. In addition, $MAC_{i,1}$ and $MAC_{i,2}$ are also generated to guarantee data integrity of $C_{i,1}$ and $C_{i,2}$ by checking the correctness of $MAC_{i,l} \stackrel{?}{=} H(C_{i,l}||TS)$, $l = 1, 2$. Thus, the security goal of verifiability and data integrity can be achieved.

Based on the security analysis, we can conclude that the proposed data dissemination scheme in vehicular fog is secure, which can successfully achieve our security goals.

5.6 Performance Evaluation

In this section, we evaluate the performance of the proposed secure vehicular data querying scheme, in terms of the maximum allowable number of vehicles, computational cost and communication overhead.

FIGURE 5.5. Computation complexity of the proposed vehicular data querying scheme. (a) Computation complexity of the RSU; (b) Computation complexity of each vehicle.

5.6.1 Analysis on the Number of Data Querying Vehicles

In performance evaluation, we choose $|n| = 1024$ bits for the plaintext of the Paillier's cryptosystem, and the size of the ciphertext in $\mathbb{Z}_{n^2}^*$ is 2048 bits. Meanwhile, we choose the Type-A pairing from the JPBC Library [174], and each signature generated in $\mathbb{Z}_q^*$ is 1024 bits. The length of the prime number α is 512 bits, and the length of the value d is 511 bits. Meanwhile, we set the length of the prime number p_2 to be 384 bits, while it satisfies the condition that $l_q + l_x < 384$ bits. (The length of l_q and l_x are selected according to the size of the predefined data query space.)

For comparison, we consider a traditional approach, which can also achieve the security requirements of unlinkability, confidentiality, and verifiability. Specifically, for each vehicle V_i in the traditional scheme, it structures the data query m_i with the row vector $\mathbf{a}_i = (a_{i,1}, a_{i,2}, \dots, a_{i,k})$ to generate the row vector $\mathbf{d}_i = (d_{i,1}, d_{i,2}, \dots, d_{i,k}) = (a_{i,1}m_i, a_{i,2}m_i, \dots, a_{i,k}m_i)$. Then V_i generates the ciphertexts of all the elements in $\mathbf{d}_i$ with the public key of the RSU (n, g) , such that the ciphertexts $C_{i,j} = g^{d_{i,j}} \cdot r_{i,1}^n \cdot H(TS)^{s_i} \bmod n^2$ ($j = 1, 2, \dots, k$) can be obtained; meanwhile, V_i also calculates the identity-based signature σ_i , which corresponds to m_i , and generates the message authentication code $MAC_i = H(C_{i,1} || \dots || C_{i,k} || TS)$. In addition, V_i formulates and transmits the message $C_{i,1} || \dots || C_{i,k} || \sigma_i || MAC_i || TS$ towards the RSU.

FIGURE 5.6. Computation complexity of the proposed vehicular data querying scheme. (a) Computation complexity of the RSU; (b) Computation complexity of each vehicle.

5.6.2 Computation Complexity

In the proposed vehicular data querying scheme, to generate an encrypted data query $(C_{i,1}, C_{i,2})$, $i \neq a$, each vehicle performs three exponentiation operations in $\mathbb{Z}_{n^2}^*$ to generate the Paillier cryptosystem's ciphertext $C_{i,1}$, and it takes one symmetric encryption to generate $C_{i,2}$. To generate the signature $\sigma_i = (\sigma_{i,1}, \sigma_{i,2})$, it takes one exponentiation operation in $\mathbb{Z}_q^*$ to generate $\sigma_{i,1}$ and it spends one exponentiation operation $\mathbb{Z}_q^*$ to generate $\sigma_{i,2}$. After receiving $C_{i,2}$, $i \neq a$, V_a takes $k - 1$ symmetric decryption operations to decrypt $C_{i,2}$. Meanwhile, it also takes three exponentiation operations and two exponentiation operations in $\mathbb{Z}_{n^2}^*$ to generate the Paillier cryptosystem's ciphertext $C_{a,1}$ and C_k respectively. For the generation of signature $\sigma_a = (\sigma_{a,1}, \sigma_{a,2})$, it takes two exponentiation operations in $\mathbb{Z}_q^*$. In comparison with computational costs of the exponentiation operations in $\mathbb{Z}_{n^2}^*$ and $\mathbb{Z}_q^*$, the computational cost of the symmetric encryption and decryption are negligible. Denote the computation costs of an exponentiation operation in $\mathbb{Z}_{n^2}^*$ and an exponentiation operation in $\mathbb{G}_1$ by C_{pe} and C_{be} , respectively. Then, totally for V_i and V_a , the computational costs will be $3 * C_{pe} + 2 * C_{be}$, $5 * C_{pe} + 2 * C_{be}$ respectively. For the RSU, it takes one exponentiation operation in $\mathbb{Z}_{n^2}^*$ to decrypt $C = \prod_{i=1}^k C_{i,1}$, and it enables the RSU to aggregate the data queries submitted by multiple vehicles into one compressed data query. With the proposed scheme, it can largely reduce the computational complexity of the

RSU. After decryption, the RSU spends two bilinear pairing operations in $\mathbb{G}$ and $k + 1$ exponentiation operations in $\mathbb{G}$ to verify the correctness of recovered data queries, which largely saves the computational resource of the RSU. We denote the bilinear pairing operation in $\mathbb{G}$ as C_{bp} , and the computational cost of the RSU is $C_{pe} + 2 * C_{bp} + (k + 1) * C_{be}$.

For comparison, we consider a traditional approach, which can also achieve the goal of securely deliver the data queries towards the RSU and verify the correctness of the recovered data queries, while without associating each data query with its origin vehicle. In the traditional data querying approach, each vehicle $V_i, i = 1, \dots, k$ encrypts each $d_{i,j}, j = 1, \dots, k$ with the public key of the RSU (n, g) and generates k ciphertexts in $\mathbb{Z}_{n^2}^*$. Under this setting, each vehicle needs to consume $3 * k$ exponentiation operations in $\mathbb{Z}_{n^2}^*$ to generate the ciphertexts $C_{i,1,l}, l = 1, \dots, k$. After receiving the encrypted data queries, the RSU aggregates the data queries and spends k exponentiation operations in $\mathbb{Z}_{n^2}^*$ for the decryption of the Paillier cryptosystem's ciphertext. Thus, the total computational cost of each vehicle is $3 * k * C_{pe} + 2 * C_{be}$, and the computational cost of the RSU is $k * C_{pe} + 2 * C_{bp} + (k + 1) * C_{be}$.

We compare the computation complexity of the proposed data query scheme with the traditional scheme. We conduct the experiments with the Java Paillier Library [149] and the Type-A pairing from the JPBC Library [174] on a desktop with 3.40GHz processor and 8.00GB memory. The experimental results indicate that the cost on a single exponentiation operation in $\mathbb{Z}_{n^2}^*$ ($|n^2| = 2048$) is $C_{pe} = 9.18$ ms, a single exponentiation operation in $\mathbb{G}$ is $C_{be} = 0.75$ ms, and a single bilinear pairing operation in $\mathbb{G}$ is $C_{bp} = 5.86$ ms. Based on the derived operation costs, we derive the computation costs of the RSU and each data querying vehicle, and compares with the traditional scheme, when the value of k varies between 5 to 100. As shown in Fig. 5.5, the proposed scheme outperforms the traditional scheme in terms of the computational cost of the RSU and each vehicle respectively. Thus, we can conclude that the proposed vehicular data querying scheme largely reduces the computation delay of all the involved entities.

5.6.3 Communication Overhead

The communication overhead of the proposed scheme can be divided into two parts, vehicle-to-RSU communication and vehicle-to-vehicle communication. We first consider the vehicle-to-RSU communication, for $V_i, i \neq a$, the data query sent towards the RSU is in the format of $ID_i || C_{i,1} || MAC_{i,1} || \sigma_i || TS$, and the corresponding communication overhead should be $|ID_i| + 2048 + 1024 + 1024 + |TS|$ bits. For V_a , the data query sent towards the RSU is in the form of $ID_a || C_{a,1} || \sigma_a || \mathbf{A} || C_k || MAC_a || TS$, and the length is $|ID_a| + 2048 + 1024 + 384 * k * k + 2048 + 1024 + |TS|$ bits. Thus, when we set $|ID_i| + |TS|$ as 64 bits, the overall vehicle-to-RSU communication overhead is $64 * k + 2048 * (k + 1) + 1024 * 2 * k + 384 * k * k$ bits. For the traditional scheme, each vehicle V_i generates k ciphertexts $C_{i,1,l}, l = 1, \dots, k$ with the length of 2048 bits. The overall communication overhead is $2048 * k * k + 64 * k + 384 * k * k$ bits.

Next, we consider the communication overhead of the vehicle-to-vehicle communication. The data query from V_i to V_1 is in the format of $ID_i || C_{i,2} || MAC_{i,2} || TS$, and the corresponding length is $|ID_i| + 384 * (k - 2) + 1024 + |TS|$ bits. Thus, the overall vehicle-to-vehicle communication overhead is $384 * (k - 2) * (k - 1) + 1024 * (k - 1) + 64 * (k - 1)$ bits. Fig. 5.6 depicts the communication overhead between the RSU and vehicles, and the total communication overhead, respectively, when the number of data querying vehicles varies from 5 to 100. As shown in Fig. 5.6, we can observe that the proposed scheme can achieve lower communication overhead than the traditional scheme.

5.7 Summary

In this chapter, we have proposed a secure request-response based vehicular data dissemination scheme in the parking lot scenario. By exploiting the homomorphic Paillier cryptosystem technique and structuring the data queries with an invertible matrix, the proposed scheme can prevent the RSU from associating each received data request with its origin vehicle, and can

support batch verification of the recovered data requests. Through security analysis, we demonstrate that the proposed scheme can achieve the security goals of unlinkability and confidentiality. Through performance evaluation, we test the performance of the proposed scheme and demonstrate its scalability, computation and communication efficiency. For the future work, we will study the possible collusion attacks and extend the proposed scheme to resist such attacks.

Chapter 6

Secure handover key management protocols in LTE-A Vehicular Networks

For VANETs, it is more applicable towards the scenario with low mobility scenario, this chapter relates to the secure issues existed in LTE-A vehicular network, which is a network supports high-mobility and with QoS guarantee. Two novel secure handover key management schemes in LTE-A vehicular networks are proposed in chapter. Specifically, a secure handover session key management scheme via mobile relay, and a novel secure CoMP joint transmission handover key management scheme in LTE-A vehicular networks are designed in the following two sections.

6.1 Achieve secure handover session key management via mobile relay

6.1.1 Introduction

With the proliferation of various enabling technologies, many small objects that surround us can be integrated into the Internet in one form or another, which is called the IoT paradigm [175].

The 3GPP has identified machine type communication (MTC) to be one of the facilitators for IoT in LTE-A networks [176, 177], and the highly penetrated portable electronic devices (smart phones, tablets, etc.) are regarded as common contact points or gateways for the large-scale deployment of IoT devices [178, 179]. Advanced transportation (high-speed trains, buses, trams, cars, etc.) which are moving along the rails/roads and equipped with various sensors and enough processing power, are perfect candidates for the real-time environmental monitoring and mapping, which is an indispensable part of IoT [180, 181]. For the information and data collected by the on-board sensors, they can be aggregated and processed by the on-board units, and then transmitted to the Internet by on-board mobile devices. In the meanwhile, the mobile devices which belong to the passengers are also responsible for managing the small scale autonomous networks, and the constant wireless broadband network access is mandatory for these on-board mobile devices [182, 183]. To increase the handover success rate, the mobile relay node (MRN) represents all the UEs to conduct handover, as a result, the heavy traffic loads caused by the handover can be released [123, 184].

From the perspective of mobile operators, the deployment of MRNs may introduce heavy capital expenditure (CAPEX) due to the large quantity of public transportation, and bring high operating expenditure (OPEX) owing to their moving characteristics [185]. Meanwhile, the public transportation companies may have a high motivation to install their own MRNs on the top of their public transportation to improve the on-board experiences of their passengers and the quality of the IoT-related services, and it can further gain more reputation and economic profits. Despite the countless advantages brought by the MRNs, there still exists many practical challenges in real applications. One crucial challenge is the emerging threats and difficulties in session key establishment between the on-board UE and the DeNB during the handover process [186].

In the 3GPP LTE-Advanced Networks, the session key established between one UE and one DeNB, aims to protect the integrity and confidentiality of the control-plane signaling, and the confidentiality of the user-plane data transmission [187, 188]. In the Evolved Packet System (EPS) of the LTE-Advanced Networks, the session key generation is based on the simultaneous

key update of both the UE and the DeNB during the handover process according to a key chaining counter [118]. However, there exist many problems related to this key management scheme. One of the potential threats is the desynchronization attack caused by manipulating the key chaining counter, and this type of attack may leave the subsequent session keys vulnerable to be compromised [1]. Especially in our MRN scenario, on the one hand, the on-board UEs may experience frequent handover, and the probability of the desynchronization attack may increase. On the other hand, the untrusted third-party owned MRN can store all the messages it received-and-forwarded, and the corruption of a chain of session keys can potentially lead to the devastating effect of the leakage of all the transmitted messages. Furthermore, the frequent handover may also bring in heavy computational pressure to the mobility management entity (MME) in the core network, due to the frequent fresh keying material generation.

To protect the secure data transmission between the on-board UEs and the DeNBs, a new handover key management scheme needs to be devised. For the on-board UEs, the proposed scheme should be able to protect the confidentiality and integrity of the data transmission. From the perspective of the DeNBs, if one DeNB is compromised by a rogue base station attack, the corrupted DeNB should be isolated from the network. And the establishment of the large amount of session keys should also not introduce heavy computational load and severe handover latency. From the perspective of the MME, the key establishment process should not introduce heavy computational burden. For the MRN, the ultimate goal of the deployment of the MRN is to bring high quality wireless access service to the on-board UEs, and the MRNs can also be involved in the session key establishment process without revealing the content of the session keys.

Motivated by the above-mentioned analysis, in this section, we devise a novel secure handover key management scheme in LTE-A networks, when the MRNs are owned by third parties. With the proposed handover key management scheme, each on-board UE can successfully establish its session key with the DeNB during the handover process while guaranteeing the forward and backward session key separations. Specifically, the main contributions of this scheme are threefold.

1) Firstly, instead of updating the session key at both the on-board UE and the DeNB sides, the session key is generated by the on-board UE itself and delivered to the DeNB. Thus, the DeNB can achieve perfect one-hop forward/backward key separation and resist against the desynchronization attack. While in the EPS key management scheme, the forward separation is achieved on a two-hop basis and the key management process is vulnerable to the desynchronization attack [118].

2) Secondly, the session key generated by the on-board UE is encrypted by a universal public key, i.e., the public key of the MME, and there requires no public key transmission of the connected DeNB. The location information of the target DeNB is also exploited in the session key encryption to avoid collusion between the mobile relay and the compromised DeNB. Since the target DeNB cannot decrypt the received message encrypted by the public key of the MME, the MRN re-encrypts the received messages into messages that can be decrypted by the target DeNB without revealing the contents of the established session keys. Furthermore, as the MME is not involved during the session key establishment process, the proposed scheme introduces no computational load to the MME.

3) Finally, to validate the effectiveness of our proposed handover key management scheme, we conduct numerical experiments. We first implement the cryptographical results using Java, then we examine the computational delay with respect to (w.r.t.) the number of MRNs and the number of UEs, and the communication overhead and storage load are also analyzed.

The remainder of this section is organized as follows. We introduce the system model and security requirements, and identify the design goals of our work in Subsection 6.1.2. In Subsection 6.1.3, we briefly review the LTE EPS handover key management, and recall the bilinear pairing and proxy re-encryption techniques as preliminaries, which will be utilized in subsequent subsections. We propose our handover key management in MRN scenario in Subsection 6.1.4, followed by security analysis and performance evaluation in Subsection 6.1.5 and in Subsection 6.1.6, respectively.

6.1.2 System Model, Security Requirement and Design Goal

In this subsection, we introduce the system model, present the security requirements, and identify our design goals.

6.1.2.1 System Model

The mobile relay network scenario under the 3GPP LTE architecture is composed of the access domain (E-UTRAN) and the core domain (EPC) [118]. The access domain consists of on-board UEs, MRNs installed on public transportation, and mobile relay capable DeNBs, as shown in Fig. 6.1. The DeNBs are connected to each other through the X2 link, while the DeNB and the MME is connected by the S1-C link [118]. When public vehicles or trains are moving forward, the UEs located inside the carriage maintain their connections to the network by communicating with the MRNs which are moving along with the on-board UEs. While at the same time, the MRNs communicate with the fixed DeNBs located near the roads or railways. In the core network domain, we only take the MME into consideration, and the functionalities of an MME includes performing mutual authentication with the UEs on behalf of the EPC and taking control of mobility management of the UEs.

When an on-board UE communicates with one DeNB through the cooperation of a MRN, it may roam out of the coverage area of the current DeNB, and then connect to another DeNB. The 3GPP LTE network supports two types of handover, i.e., inter-MME handover and intra-MME handover. Since the inter-MME handover occurs between the UE and MME by running full EPS-AKA procedure, without the involvement of any signaling transmission between DeNBs, we only consider the intra-MME handover process and focus on the wireless links between the DeNB and the on-board UEs. Under this mobile relay network scenario, the on-board UEs conduct intra-MME handover between two neighboring DeNBs with the help of a MRN which is moving along. When a group of on-board UEs roam to the coverage of a new DeNB, new session keys need to be generated and established to protect the security of the radio access links between the on-board UEs and the new DeNB.

FIGURE 6.1. System model under consideration

6.1.2.2 Security Requirements

A secure handover key management mechanism is one of the most important processes for the successful practice of a MRN. In our security model, we regard all the entities in the EPC and all the on-board UEs as trustable. Since the MRNs are owned and deployed by the third parties instead of the operators in this network scenario, the MRNs are considered to be untrusted. All the messages transmitted between the DeNBs and the on-board UEs are received-and-forwarded by the MRNs, the MRNs can eavesdrop, modify and replay the received messages. Furthermore, the DeNBs are also viewed as untrusted since they reside in the open environment, and the adversaries can launch some active attacks to corrupt the DeNBs. Therefore, the following security requirements should be satisfied in the secure handover key management.

Secure Session Key Establishment [189]. During the handover process, a secure session key between each on-board UE and the DeNB should be derived and established to protect the confidentiality and integrity of the transmission in the radio access link. The MRN should not disrupt the secure session key generation and association or achieve the session keys shared between the on-board UEs and the DeNB, even though it is capable of eavesdropping, modifying or replaying the messages passed through.

Backward Key Separation. For any connected DeNB, it is computationally infeasible to derive the previous session keys based on the current session keys shared with the on-board UEs. If one DeNB is corrupted, without the previous session keys, the messages transmitted between the previous DeNBs and the on-board UEs cannot be decrypted, and the attack can only compromise the transmission of the corrupted DeNB.

Forward Key Separation. Given the session keys shared between the on-board UEs and the connected DeNB, the DeNB is not computationally feasible to know the future session keys after the subsequent handover. Forward key separation prevents a corrupted DeNB from compromising other future DeNBs, and keeps the security breaches as local as possible.

Collusion Resistance [190]. For any compromised DeNB, even though the malicious MRNs owned by third-parties possess sufficient storage and computation capability, it is computationally infeasible for a collusion coalition between a corrupted DeNB and a malicious MRN to recover any of the previous and afterwards transmitted messages.

6.1.2.3 Design Goals

Based on the aforementioned system model and security requirements, our design goal is to develop an efficient and robust key management scheme during the handover process. Specifically, the following objectives should be achieved for the key management scheme during handover.

Security: The above-mentioned security requirements should be satisfied in the devised handover key management scheme. According to the previous statement and analysis, without taking the secure handover key management scheme into consideration, the real application of the MRNs is not even possible. Thus, the proposed handover key management scheme should simultaneously guarantee the secure session key derivation and association, achieve forward and backward key separation, at the same time, resist collusion attacks in the mobile relay handover scenario.

Efficiency: The security handover mechanism should not introduce heavy computational load to the MME, due to the keying material generation and update caused by the simultaneous frequent handover of the large amount of on-board UEs. For the on-board UEs and the target DeNB, the proposed handover key management scheme should not introduce heavy computational and traffic load, and the latency caused by the session key establishment should not degrade the user experience of the on-board UEs.

6.1.3 Preliminaries

In this subsection, we recall the cryptographic techniques: bilinear pairing and proxy re-encryption, which function as the basis of our proposed handover key management mechanism.

6.1.3.1 Bilinear Pairing

Let $\mathbb{G}$ and $\mathbb{G}_{\mathbb{T}}$ be two multiplicative groups with the same prime order q . The bilinear map $e: \mathbb{G} \times \mathbb{G} \rightarrow \mathbb{G}_{\mathbb{T}}$ has the following properties:

- 1) Bilinearity: $\forall g, h \in \mathbb{G}$, and $\forall a, b \in \mathbb{Z}_q^*$, we can derive the equation $e(g^a, h^b) = e(g, h)^{ab}$;
- 2) Nondegeneracy: There exists at least one g , where $g \in \mathbb{G}$ and one h , where $h \in \mathbb{G}$, which satisfies the condition that $e(g, h) \neq 1_{\mathbb{G}_{\mathbb{T}}}$;
- 3) Computable: $\forall g, h \in \mathbb{G}$, there is an efficient algorithm to compute $e(g, h)$.

In group $\mathbb{G}$, the Computational Deffie-Hellman (CDH) problem is computationally intractable, i.e., given (g, g^a, g^b) for $g \in \mathbb{G}$ and unknown $a, b \in \mathbb{Z}_q^*$, it is computationally infeasible to calculate g^{ab} in polynomial time [84]. While the Decisional Diffie-Hellman (DDH) problem is easy to solve, i.e., given (g, g^a, g^b, g^c) for $g \in \mathbb{G}$ and unknown $a, b, c \in \mathbb{Z}_q^*$, it is easy to determine whether $c = ab \bmod q$ by checking whether $e(g^a, g^b) \stackrel{?}{=} e(g^c, g)$.

A bilinear parameter generator $\mathcal{Gen}$ denotes a probabilistic algorithm that takes a parameter related to κ as the input, and gives a 5-tuple $(q, g, \mathbb{G}, \mathbb{G}_{\mathbb{T}}, e)$ as the output, where q is κ -bit prime,

$\mathbb{G}$, and $\mathbb{G}_T$ are two multiplicative groups with order q , $g \in \mathbb{G}$ is a generator, and $e : \mathbb{G} \times \mathbb{G} \rightarrow \mathbb{G}_T$ is non-degenerated and computable bilinear map.

6.1.3.2 Proxy Re-encryption

Proxy re-encryption technique is a cryptographic technique which allows a proxy to alter a ciphertext encrypted for one entity into a new ciphertext which can be decrypted by another entity. Assume *User B* encrypts a message with the public key of *User A*, and it can only be decrypted by the private key of *User A*. If *User A* permits *User C* to reveal the content of the message encrypted by the public key of *User A*, *User A* distributes the corresponding re-encryption key $rk_{A \rightarrow C}$ to the proxy. And the proxy transforms the ciphertext encrypted by the public key of *User A* into a new ciphertext which could be decrypted by the private key of *User C*. Meanwhile, the proxy server is not able to read or reveal the contents of the underlying messages [191].

6.1.4 Proposed Handover Key Management Scheme via Mobile Relay

In this subsection, we introduce our proposed secure handover key management scheme in the MRN scenario, as shown in Fig. 6.2. The session keys are first generated and encrypted by the on-board UEs, and then delivered to the MRNs. The MRNs re-encrypt the received messages, so that the target DeNB can successfully decrypt the received messages to obtain the session keys.

6.1.4.1 System Initialization

Given the security parameter κ , the MME first generates the bilinear parameters $(q, g, \mathbb{G}, \mathbb{G}_T, e)$ by running $\mathcal{Gen}(\kappa)$, then chooses a symmetric encryption algorithm $\mathcal{Enc}()$ and two secure one-way Hash functions H_1 and H_2 , where $H_1 : \{0, 1\}^* \rightarrow \mathbb{G}$ and $H_2 : \{0, 1\}^* \rightarrow \mathbb{G}$. In addition, the MME owns a private key x_M , and computes $y_M = g^{x_M}$ as its public key, the notations are

TABLE 6.1. Notations of Secure Session Key Establishment in MRN handover

Notation	Definition
eNB_ID_j	Identifier of the target DeNB
$H_1(L_j)$	Location-based Generator of $DeNB_j$
x_j	Private Key of $DeNB_j$
y_j	Public Key of $DeNB_j$
x_M	Private Key of MME
y_M	Public Key of MME
$H_1(L_j)^{x_j/x_M}$	Re-encryption Key of $DeNB_j$
x_i	Private Key of UE_i
y_i	Public Key of UE_i
K_i^j	Session Key Shared between UE_i and $DeNB_j$

shown in TABLE 6.1. Finally, the MME keeps the private key x_M in confidential, and publishes the system parameter $params = (q, g, \mathbb{G}, \mathbb{G}_T, e, H_1, H_2, y_M, \mathcal{Enc}())$. For each on-board UE i , it owns a private key x_i , and the corresponding public key y_i , where $y_i = g^{x_i}$. The public key of a UE is stored in the MME during the registration phase, and it can be achieved by the DeNBs under the coverage of the MME.

For DeNB j , it also possesses a pair of private and public keys, which are denoted as $x_j \in \mathbb{Z}_q^*$ and $y_j = H_1(L_j)^{x_j}$. $L_j = (TAC || eNB_ID_j)$ is a unique parameter which is a combination of the location information and its unique identity. The location of DeNB j is indicated by the tracking area code (TAC), which is a unique code that is assigned by the operator to each of its tracking area, including a group of neighboring DeNBs. The unique identity of the DeNB j (eNB_ID_j) is also assigned by the operator during the infrastructure deployment phase. When DeNB j registers itself to the MME, DeNB j delivers the public key $y_j = H_1(L_j)^{x_j}$ to the MME, and proves the authenticity of y_j using the following zero-knowledge proof process [192]. The MME first picks a random number r_z and a random message M_z to form the ciphertext ($a_j = H_1(L_j)^{r_z}, b_j = y_j^{r_z} M_z$), and then delivers (a_j, b_j) to DeNB j . Then DeNB j decrypts the received ciphertext and replies the random message $M_z = b_j / (a_j)^{x_j}$ to the MME. Thus, the DeNB j can prove that it does deliver its public key to the MME and it does use the unique generator $H_1(L_j)$ in the public key y_j . After the validation of DeNB j , the MME generates a re-encryption key $H_1(L_j)^{x_j/x_M}$ and distributes it to DeNB j . Since

FIGURE 6.2. Message flow in MRN scenario during UE handover

the DeNB j periodically broadcasts the pilot signal containing the TAC and eNB_ID_j in its coverage area, the MRN can also produce the generator $H_1(L_j)$ of DeNB j .

6.1.4.2 Handover Key Management Scheme in MRN Scenario

The message flows between an on-board UE and the target DeNB with the involvement of the source DeNB and MRN in the secure handover key management process is depicted in Fig. 6.2.

Step 1: Before the handover process, the MRN conducts the communication channel measurement on behalf of the on-board UEs under its coverage area. After conducting the measurement process, the MRN generates and sends a channel measurement report to the source DeNB as a request for handover (*message (1)*).

Step 2: While at the same time, the MRN generates $H_1(L_j) = H_1(TAC || eNB_ID_j)$ and delivers $H_1(L_j)$ to the on-board UEs under its coverage. The MRN also acknowledges the on-board UEs to prepare for the generation of the session keys that will be utilized during the next hop transmission (*message (a)*).

Step 3: According to the channel measurement report, the source DeNB determines whether the MRN representing the on-board UEs under its coverage begins to conduct the handover process or not. If the source DeNB decides to handover to the target DeNB, it initiates a handover request to the target DeNB with the public keys of the on-board UEs attached, which will be further utilized by the target DeNB for signature validation (*message (2)*).

Algorithm 2: Session Key Generation and Encryption

Data: A group of n on-board UEs obtain the location-based generator $H_1(L_j)$ of the target DeNB j from the MRN;

Compute:

for $i = 1 : n$ **do**

1. On-board UE i chooses a random number $r_i \in \mathbb{Z}_q^*$;

2. Then on-board UE i generates c_i^1 with public key of the MME y_M :

$$c_i^1 = (y_M)^{r_i} = (g^{x_M})^{r_i} = (g^{r_i \cdot x_M});$$

3. On-board UE i generates c_i^2 with the session key K_i^j :

$$c_i^2 = K_i^j \cdot e(H_1(L_j), g)^{r_i};$$

4. On-board UE i generates the signature s_i^j with the current time identifier $Timestamp$ and its private key x_i :

$$s_i^j = H_2(K_i^j || Timestamp)^{x_i}$$

end for

Output: $C_i^j = (c_i^1, c_i^2, s_i^j)$, for $i = 1, 2, \dots, n$

Step 4: After receiving the handover request from the source DeNB, the target DeNB checks its current traffic load. If there is enough spare resource for the on-board UEs, the target DeNB accepts the handover request and replies its re-encryption key $H_1(L_j)^{x_j/x_M}$ back to the source DeNB (message (3)). Then the source DeNB helps to deliver $H_1(L_j)^{x_j/x_M}$ to the MRN and send a handover command to the MRN (message (3)).

Step 5: After achieving (message (a)), on-board UE i randomly generates a session key, K_i^j , which later will be shared between on-board UE i and the target DeNB j . Then all the on-board UEs begin to compute their encrypted session keys and the corresponding signatures using Algorithm 2, where s_i^j is the signature of the session key K_i^j , aiming to validate the correctness and integrity of the session key sent to the target DeNB after the decryption process. Then the on-board UE transmits C_i^j to the MRN. Regardless of which DeNB it is connected to, all the on-board UEs utilize the universal public key of the MME y_M for the session key encryption instead of exploiting the public key of the target DeNB, which saves the computational and communication resource spent on update of the public key. Furthermore, the generation of c_i^1 and c_i^2 can be done offline in Algorithm 2, which can effectively reduce the computational delay.

Step 6: Since the session key K_i^j shared between the on-board UE i and the target DeNB j is encrypted by the public key of the MME y_M , without the private key x_M , the target DeNB cannot decrypt K_i^j directly. Even though the MRN is owned and deployed by the third party

Algorithm 3: Session Key Re-encryption**Data:** The MRN obtains the re-encryption key $H_1(L_j)^{x_j/x_M}$ from the target DeNB j ;**Compute:****for** $i = 1 : n$ **do**The MRN derives $\hat{c}_i^1$ from c_i^1 and $H_1(L_j)^{x_j/x_M}$: $\hat{c}_i^1 = e(c_i^1, H_1(L_j)^{x_j/x_M}) = e(g, H_1(L_j))^{r \cdot x_j}$;Update $\hat{C}_i^j$: $\hat{C}_i^j = (\hat{c}_i^1, c_i^2, s_i^j)$;**end for****Output:** $\hat{C}_i^j$, for $i = 1, 2, \dots, n$

public transportation company, the ultimate goal of the deployment of the MRN is to provide high-quality wireless access services to the on-board UEs. Besides, the MRN possesses certain computational capability. Thus, the MRN has the motivation and ability to join in the key establishment process and functions as a proxy. In this situation, the MRN acts as a delegate to transfer the messages encrypted with the public key of MME into messages, that can be decrypted with the private key of the target DeNB without revealing the contents of the messages. During the re-encryption process, the MRN converts c_i^1 into $\hat{c}_i^1$ by performing Algorithm 3, then the converted message $\hat{C}_i^j = (\hat{c}_i^1, c_i^2, s_i^j)$ is delivered to the target DeNB, as shown in (message (4)).

Step 7: When the re-encrypted message $\hat{C}_i^j = (\hat{c}_i^1, c_i^2, s_i^j)$ arrives at the target DeNB, the target DeNB j first decrypts the session key generated by the on-board UE i with its own private key x_j , and validates the accuracy of the session key with the public key y_i of the on-board UE i , as shown in Algorithm ??.

Correctness. We first validate the correctness of the decryption process of K_i^j at the target DeNB j :

$$\begin{aligned}
 K_i^j &= \frac{c_i^2}{(\hat{c}_i^1)^{x_j^{-1}}} = \frac{K_i^j \cdot e(H_1(L_j), g)^r}{(e(g, H_1(L_j))^{r \cdot x_j})^{x_j^{-1}}} \\
 &= \frac{K_i^j \cdot e(g, H_1(L_j))^r}{(e(g, H_1(L_j))^{r \cdot x_j})^{x_j^{-1}}} = \frac{K_i^j \cdot e(g, H_1(L_j))^r}{e(g, H_1(L_j))^r}
 \end{aligned} \tag{6.1}$$

Then we check the correctness of the signature verification process by exploiting the following mathematical process:

$$e(s_i^j, g) \stackrel{?}{=} e(H_2(K_i^j || \text{Timestamp}), y_i) \quad (6.2)$$

Step 8: After the signature verification, the symmetric session key K_i^j is successfully established between the target DeNB j and the on-board UE i , and it is used to protect the control-plane and data-plane transmission.

Thus, we can achieve the objective of the secure key establishment between the target DeNB and the on board UEs under the help of the MRN.

Algorithm 4: Session Key Decryption and Signature Verification

Data: The DeNB j gets $\hat{C}_i^j, i \in \{1, 2, \dots, n\}$;

Compute:

for $i = 1 : n$ **do**

1. The DeNB decrypts the session key K_i^j with x_j :

$$K_i^j = \frac{c_i^2}{(c_i^1)^{x_j^{-1}}} = \frac{K_i^j \cdot e(g, H_1(L_j))^r}{e(g, H_1(L_j))^r};$$

2. The DeNB conducts the following computation:

$$e(H_2(K_i^j || \text{Timestamp}), y_i);$$

3. The DeNB checks whether:

$$e(s_i^j, g) \stackrel{?}{=} e(H_2(K_i^j || \text{Timestamp}), y_i);$$

end for

Output: K_i^j , for $i = 1, 2, \dots, n$

6.1.5 Security Analysis

In this subsection, detailed security analysis of the proposed handover key management scheme is given to demonstrate that it satisfies the requirements of the efficient secure session key establishment, backward and forward key separation, and collusion-resistance.

- *The proposed handover key management scheme can efficiently achieve the goal of secure session key establishment between the on-board UEs and the target DeNB.* When each on-board UE uses the public key y_M of the MME for the session key encryption, the ciphertext (c_i^1, c_i^2) for the session key K_i^j shared between the on-board UE i and the

target DeNB j can only be decrypted by the MME with the private key x_M , which is shown below

$$K_i^j = \frac{c_i^2}{e(c_i^1, H_1(L_j))^{x_M^{-1}}} = \frac{K_i^j \cdot e(H_1(L_j), g)^r}{e(g^{r \cdot x_M}, H_1(L_j))^{x_M^{-1}}}. \quad (6.3)$$

Therefore, no one else can obtain the session key K_i^j at this stage. Because the MRN possesses computational capability and has the motivation to conduct the re-encryption, the MRN converts the messages encrypted by the public key y_M of the MME into messages which could be decrypted by the private key x_j of the target DeNB. After the re-encryption, only the target DeNB can decrypt the session key K_i^j with its private key x_j , as indicated in Eq. (6.1). To guarantee the correctness and integrity of the decrypted session key, a signature s_i^j is introduced to further prevent the MRN from distorting the session key. Thus, the goal of the secure session key establishment can be achieved.

- *The proposed handover key management mechanism can achieve the backward and forward key separation of the session keys during the handover process.* During each handover, the session key K_i^j shared between the target DeNB j and the on-board UE i is generated randomly by the on-board UE i , and transmitted to the target DeNB via the MRN without revealing the content of the session key. Since one session key is randomly generated and can only be utilized once, the compromise of one session key cannot further corrupt the other sessions keys shared between the on-board UE and the other DeNBs. Furthermore, during each session key establishment process, the corresponding re-encryption key $H_1(L_j)^{x_j/x_M}$ for the session key are distinctly different at different L_j , the forward and afterward DeNB cannot decrypt and obtain the session key. As a result, the backward and forward key separation can be achieved during the handover key management in the MRN scenario.
- *The proposed handover key management mechanism can resist the collusion attack between the corrupted DeNB and the MRN.* Suppose DeNB h is corrupted by an attacker and colludes with a malicious MRN, and the goal of collusion is intended to reveal the

transmission content of DeNB j . In our scheme, if we do not introduce the location-based generator $H_1(L_j)$ in the re-encryption key, i.e., $H_1(L_j)^{x_j/x_M}$ becomes g^{x_j/x_M} , and c_i^2 will be computed as $c_i^2 = K_i^j \cdot e(g, g)^r$. Therefore, when DeNB h and the malicious MRN intend to reveal the session key K_i^j , the malicious MRN can transform c_i^1 with the re-encryption key g^{x_h/x_M} of DeNB h and obtain the following

$$\hat{c}_i^{1h} = e(g^{x_M r}, g^{x_h/x_M}) = e(g, g)^{x_h r} \quad (6.4)$$

Then, the session key can be gained by the DeNB h with its private key x_h

$$K_i^j = \frac{c_i^2}{(\hat{c}_i^{1h})^{x_h^{-1}}} = \frac{K_i^j \cdot e(g, g)^r}{e(g, g)^r} \quad (6.5)$$

Since our proposed scheme introduces $H_1(L_j)$ in the re-encryption key $H_1(L_j)^{x_j/x_M}$, obviously, the above collusion does not work. Therefore, the goal of the collusion attack resistance can be achieved.

6.1.6 Performance Evaluation

In this subsection, we evaluate the performance of the proposed handover key management scheme in the MRN scenario, in terms of the computational delay of DeNB, MRN, and on-board UE, the communication overhead of the DeNB and on-board UE transmission, and the storage costs of DeNB, MRN, and on-board UE.

6.1.6.1 Computational Cost

T_{UE} , T_{MRN} , and T_{DeNB} denote the average computational cost of an on-board UE, an MRN, and a DeNB, respectively. The experimental environment is a desktop with dual-core 3.10GHz processor and 8.00GB installed memory, using the Type A pairing in the Java Pairing-Based Cryptography Library [174]. We test the value of T_{UE} , T_{MRN} , and T_{DeNB} 1000 times, and their average values are reported as 13.3ms, 10.1ms, and 21.2ms, respectively.

Assume there are n on-board UEs and m MRNs installed on a public transportation, and we also assume that the MRNs and the target DeNB have the ability of conducting parallel computation, i.e., four calculations can be carried out simultaneously. Thus, there are $\lfloor \frac{n}{4m} \rfloor + 1$ on-board UEs allocated to each MRN.

FIGURE 6.3. Computational delay of DeNB, when $T_{MRN} \leq T_{DeNB}$

As shown in Fig. 6.3, when $T_{MRN} \leq T_{DeNB}$, the bottleneck of the computational delay is T_{DeNB} , that is, the second on-board UE needs to wait a short time for the DeNB to conduct decryption and signature verification after the re-encryption of the MRN. In this case, the time consumption for the on-board UE with the maximum computational delay is

$$T_{max-UE} = T_{UE} + T_{MRN} + (\lfloor \frac{n}{4m} \rfloor + 1)T_{DeNB}. \quad (6.6)$$

When $T_{MRN} > T_{DeNB}$, as shown in Fig. 6.4, the corresponding time consumption of the on-board UE with the maximum computational delay is denoted as

$$T_{max-UE} = T_{UE} + (\lfloor \frac{n}{4m} \rfloor + 1)T_{MRN} + T_{DeNB} \quad (6.7)$$

FIGURE 6.4. Computational delay of DeNB, when $T_{MRN} > T_{DeNB}$

Under both cases, the time consumption of the on-board UEs with the minimum delay keeps the same, i.e., the computational delay of the first on-board UE, which is $T_{min-UE} = T_{UE} + T_{MRN} + T_{DeNB}$. In our proposed handover key management scheme, we use Eq. (6.6) to calculate the maximum delay of the on-board UEs, since $T_{MRN} \leq T_{DeNB}$. And the total time consumption of the target DeNB during the session key establishment process is $T_{total-DeNB} = mT_{max-UE}$. Since there are m MRNs installed on the public transportation, and the time consumption of the target DeNB with one MRN equals the maximum transmission delay of the on-board UE.

FIGURE 6.5. Comparison of computational delay of DeNB w.r.t. the number of UE

To validate the efficiency of the proposed group handover key management scheme, we compare the proposed scheme with the sequential handover scheme, that is, on-board UEs conduct the handover separately. The time consumption of each on-board UE without exploiting the group handover is T_{min-UE} , and the total time consumption of the target DeNB is $T'_{total-DeNB} = nT_{min-UE}$.

In Fig. 6.5, we compare the computational delay of the target DeNB with and without exploiting the group handover scheme with varying number of the on-board UEs, which is from 20 to 80. As shown in Fig. 6.5, with more on-board UEs, the computational delay of the target DeNB becomes longer, when the number of the MRN is set to be 1 and 2, respectively. As shown in Fig. 6.5 and Fig. 6.6, the computational delay of the DeNB in our proposed group

FIGURE 6.6. Comparison of computational delay of DeNB w.r.t. the number of MRN

handover scheme is always shorter than that in the sequential handover scheme, which shows the efficiency of our proposed scheme in terms of the DeNB's computational delay.

Given the same parameter setup, we simulate and compare the computational delay of the target DeNB with and without exploiting the group handover scheme by choosing different number of MRNs, which ranges from 1 to 4. As shown in Fig. 6.6, the total computational delay of the target DeNB (with and without exploiting group signature) increases w.r.t. the increase of the number of MRNs, when the number of the on-board UEs are set to be 40 and 60, respectively. This is explained by the fact that for multiple MRNs, the computational delay of the on-board UE and the MRNs in $T_{total-DeNB}$ are calculated multiple times. In the sequential handover scheme, the computational delay of the target DeNB $T'_{total-DeNB}$ almost keep the uniform, this is explained by the fact that each on-board UE calculates the computational delay independently.

Since the proposed group handover key management scheme not only needs to consider the computational delay of the DeNB, but also needs to evaluate the computational delay of the on-board UEs, which is related to the quality of service of the on-board UEs. Thus we compare the maximum, average, and minimum computational delay of an on-board UE in the group handover scheme in Fig. 6.7, when the number of MRN is set to be 2 and 3. With the increase of the number of the on-board UEs, the maximum and average computational delay of the

FIGURE 6.7. Computational delay of UE w.r.t. the number of UE

on-board UE increases. This is because with more on-board UEs, the maximum and average waiting time of the on-board UEs are longer.

FIGURE 6.8. Computational delay of UE w.r.t. the number of MRN

The computational delay of the on-board UE is also related to the number of the MRN, so we compares the maximum, average, and minimum computational delay of the on-board UEs. We simulate the computational delay of the on-board UEs by choosing different number of MRN, which is from 1 to 5. As shown in Fig. 6.8, with more MRNs, the maximum and average computational delay of the on-board UE decrease. This is because the increase of the number of the on-board UEs lead to the decrease of the number of the on-board UEs attached to each MRN, the maximum and average waiting time of the on-board UE decrease.

FIGURE 6.9. Maximum computational delay of UE w.r.t. the number of MRN and UE

As shown in Fig. 6.9, we show the maximum computational delay of UE w.r.t. the increase of number of MRN and UE. The maximum computational delay of UE increases w.r.t. the increase of the number of UE under the same number of MRN. While with the increase of the number of MRN, the maximum computational delay of UE decreases under the same number of UE.

6.1.6.2 Communication and Storage Overhead

In order to analyze the communication overhead, we assume that both the MRN and the DeNB successfully re-encrypts and decrypts the received messages. For each MRN, since the length of the generator $H_1(L)$ is 1024 bits, the total throughput caused by the location-based generator is calculated as $Throughput_{LB} = 1024 * m$ bits. For the on-board UEs under the coverage of one MRN, the introduced throughput is $Throughput_{UE} = 3 * 1024 * n$ bits. The overall throughput taking the 1024 bits re-encryption key transmission into consideration is $1024 + 1024 * m + 3072 * n * m$ bits.

We analyze the storage overhead of the handover key management scheme, from the perspective of the three involved entities. For the on-board UEs, the occupied storage space for the public key of the MME is 1024 bits. The storage space of the MRN for the re-encryption key is

also 1024 bits. From the perspective of the DeNB, the storage space of the public keys belongs to the on-board UEs is $1024 * m * n$ bits.

6.2 Achieving secure CoMP joint transmission handover

6.2.1 Introduction

Due to the pervasive deployment and broadband transmission of LTE-A networks, it is particularly suitable for the category of vehicular applications with high-bandwidth demands and sensitive QoS requirements, such as intelligent traffic control and on-board infotainment services [5]. However, the high-speed mobility of vehicles results in frequent handovers, which causes the vehicles to constantly suffer from high inter-cell interference and low throughput at cell edges [132]. To solve this problem, CoMP joint processing handover schemes are proposed in LTE-A networks [132, 133], which provides multiple transmissions for each UE to increase the received signal strength and decrease the received signal spatial interference.

Although the CoMP joint processing handover schemes can improve the cell edge performance of UEs, there still exist several challenges for its real application, especially in the aspect of secure handover key management. Firstly, in LTE-A system, a new session key is derived from the keying material provided by the MME, and the key chaining counter exchanged between a UE and an eNB. However, in the existing scheme, the forward key separation can be compromised by the desynchronization attack through manipulating the key chaining counter. Even though EPS-AKA can prevent the desynchronization attack through mutual authentication and root key update [1], it requires the transmission of IMSI which may cause the UE identity disclosure [193]. In addition, it also brings additional resource consumption at the core network, increases the handover delay and introduces extra traffic overhead. Secondly, to achieve the diversity gain at the receiver side, the multiple data transmissions from multiple eNBs should be encrypted with an identical session key. Meanwhile, due to the bandwidth limitation of existing backhaul links [194] and the fact that the cryptographic operations may increase the data

size, a secure session key needs to be established among the vehicle and the cooperating eNBs. To address the above challenges, a new secure handover scheme needs to be devised for the CoMP joint transmission scenario.

Based on the above mentioned, in this section, we devise a new secure key management scheme for the CoMP joint transmission handover scenario in LTE-A vehicular networks. With the proposed scheme, each LTE-enabled vehicle can successfully establish its session key with the cooperating eNBs during the CoMP joint transmission handover. Specifically, the main contributions of this section are threefold:

- Firstly, the proposed scheme can establish a secure session key between the vehicle and the cooperating eNBs involved in the CoMP joint transmission handover, which can accommodate to the bandwidth limitation of the backhaul links and help to achieve the diversity gain at the receiver side.
- Secondly, to resist the desynchronization attack and achieve the backward/forward key separation, the established session key is generated by the vehicle and then delivered towards the cooperating eNBs after protection. To reduce the operational complexity of the key encryption process, the session key is encrypted by the product of the public keys of all cooperating eNBs, which then can be decrypted with the private keys of each cooperating eNB respectively.
- Thirdly, to validate the efficiency of the proposed scheme, we implement and test the computational delay in Java, and we also analyze the feasibility of the proposed scheme.

The rest of this section is organized as follows. We introduction our system model, show our security requirements, and identify our design goal in Subsection 6.2.2. We propose our scheme in Subsection 6.2.3, followed by security analysis and performance evaluation in Subsection 6.2.4 and Subsection 6.2.5, respectively.

6.2.2 Model, Requirements and Design Goal

In this section, we introduce our system model, security requirements, and identify our design goal.

6.2.2.1 System Model

In our system model, we consider a CoMP joint transmission handover scenario in LTE-A vehicular networks, when a vehicle located under the coverage of the serving eNB (SeNB) is moving towards the target eNB (TeNB). As shown in Fig. 6.10, the system consists of five entities: a vehicle, an SeNB, a TeNB, an MME, and an S-GW.

- The vehicle is equipped with a communication module (via either tethering a smartphone to the vehicle or owning a built-in on-board cellular module), which enables its bi-directional cellular communication with eNBs [195]. Note that, the two words: vehicle and UE, will be used interchangeably throughout this proposed scheme.
- MME is the key control node in LTE-A access network, whose functionalities include UE authentication and UE mobility management. Meanwhile, an interface *SI-MME* is set up between the MME and each eNB for the control-plane signaling transmission.
- S-GW is responsible for routing and forwarding user data packets. During an inter-eNB handover, it also acts as the mobility anchor for user-plane data. A path for the user-plane data traffic is established between each eNB and the S-GW, which is designated as *SI-U* interface. In addition, S-GW is connected to MME through *S11* interface.
- SeNB and TeNB are two neighboring base stations deployed along with the highway, and there exists a direct interface (*X2 interface*) established between them. Meanwhile, SeNB and TeNB are connected to the same MME, and the handover preparation is occurred through *X2 interface*. In addition, SeNB is currently responsible for making handover decisions for the vehicle, connecting to MME and S-GW for control- and user-plane data transmission, and sharing CoMP joint transmission related data and user data towards TeNB.

6.2.2.2 Security Requirements

In the duration of the CoMP joint transmission handover, the following two security requirements should be achieved.

FIGURE 6.10. CoMP joint transmission handover in LTE-A vehicular networks

- *The proposed scheme should achieve both forward and backward key separation between two consecutive eNBs.* In specific, based on the session key shared between a UE and a connected eNB, it should be impossible for the connected eNB to recover or infer the previous and future session keys shared between the UE and other eNBs, which prevents a compromised eNB from endangering the neighboring eNBs and keeps the security breaches locally.
- *The proposed scheme should successfully establish a session key between the vehicle and the cooperating eNBs in CoMP joint transmission.* To achieve the diversity gain brought by the multiple transmissions, and reduce the traffic load brought by the cryptographic operations in the backhaul, a session key should be established among the vehicle and the cooperating eNBs. After key establishment, the user-plane data of SeNB can be shared towards TeNB (the shared data is protected by the IPSec protocol in the wired backhaul), then encrypted by the cooperating eNBs with an identical session key, and transmitted towards the vehicle to achieve the diversity gain brought by the multiple transmissions.

6.2.2.3 Design Goal

Based on the above system model and security requirements, our design goal is to develop a secure CoMP joint transmission handover key management scheme in LTE-A vehicular network, which should 1) satisfy the above security requirements (achieves forward and backward key

TABLE 6.2. Notations of Proposed Key Establishment Scheme in CoMP Handover

Notation	Definition
$H_0(), H()$	Two Secure Cryptographic Hash Functions
sk_i	Secret Key of eNB_i , $sk_i = (r_i, X_i)$
pk_i	Public Key of eNB_i , $pk_i = (R_i, A_i)$
h_i	String Related to eNB_i
sk_v, pk_v	Secret and Public Key pair of Vehicle v
$q, g, \mathbb{G}, \mathbb{G}_T, e$	Parameters of Bilinear Groups
K_{ST}	Session Key Delivered to SeNB and TeNB
K_T	Session Key Delivered to TeNB

separation and secure session key establishment); 2) minimize the traffic load and signaling interactions during the secure handover process; 3) reduce the computational complexity brought to the eNBs and vehicle, due to the fast moving characteristic of the vehicles and the stringent handover interval.

6.2.3 Proposed Handover Key Management Scheme for CoMP

In this subsection, based on the asymmetric group key agreement technique [196], we propose our secure CoMP joint transmission handover key management scheme in LTE-A vehicular network, which consists of three parts: preliminaries, system initialization, and secure handover key management process. Meanwhile, the definitions of some used notations are described in Table 6.2.

6.2.3.1 Preliminaries

In our proposed scheme, the bilinear pairing technique is our building block. Let $\mathbb{G}$ and $\mathbb{G}_T$ be two multiplicative groups with the same prime order q . The bilinear map $e: \mathbb{G} \times \mathbb{G} \rightarrow \mathbb{G}_T$ has the following properties: 1) Bilinearity: $\forall g, h \in \mathbb{G}$, and $\forall a, b \in \mathbb{Z}_q^*$, we can derive $e(g^a, h^b) = e(g, h)^{ab}$; 2) Nondegeneracy: There exist $g \in \mathbb{G}$ and $h \in \mathbb{G}$, satisfy the condition that $e(g, h) \neq 1_{\mathbb{G}_T}$; 3) Computable: $\forall g, h \in \mathbb{G}$, there is an efficient algorithm to compute $e(g, h)$.

A bilinear parameter generator $\mathcal{Gen}$ denotes a probabilistic algorithm that takes a parameter related to κ as the input, and generates a 5-tuple $(q, g, \mathbb{G}, \mathbb{G}_T, e)$ as the output, where q is κ -bit prime, $\mathbb{G}$ and $\mathbb{G}_T$ are two multiplicative groups with order q , $g \in \mathbb{G}$ is a generator, and $e : \mathbb{G} \times \mathbb{G} \rightarrow \mathbb{G}_T$ is non-degenerated and computable bilinear map.

6.2.3.2 System Initialization

In the system initialization phase, we assume the MME will bootstrap the entire system. Specifically, given a security parameter κ , the MME generates a five-tuple bilinear parameters $(q, g, \mathbb{G}, \mathbb{G}_T, e)$ by running $\mathcal{Gen}(\kappa)$. In addition, the MME chooses a secure cryptographic hash functions $H()$, where $H : \{0, 1\}^* \rightarrow \mathbb{G}$. Finally, the MME publishes the system parameters $params = (q, g, \mathbb{G}, \mathbb{G}_T, e, H())$.

The SeNB, located at L_s , chooses two random elements $r_s \in \mathbb{Z}_q^*$ and $X_s \in \mathbb{G}$ as its secret key $sk_s = (r_s, X_s)$, and calculates the corresponding public key $pk_s = (R_s, A_s)$, where $R_s = g^{-r_s}$, $A_s = e(X_s, g)$. The SeNB uses the private key $sk_s = (r_s, X_s)$ to compute and keep the signature $\sigma_s = X_s \cdot h_s^{r_s}$ privately, where $h_s = H(SeNB||L_s)$. Similarly, the TeNB, located at L_t , chooses two random elements $r_t \in \mathbb{Z}_q^*$ and $X_t \in \mathbb{G}$ as its secret key $sk_t = (r_t, X_t)$, and calculates the corresponding public key $pk_t = (R_t, A_t)$, where $R_t = g^{-r_t}$, $A_t = e(X_t, g)$. The TeNB also uses the private key $sk_t = (r_t, X_t)$ to compute and keep the signature $\sigma_t = X_t \cdot h_t^{r_t}$ privately, where $h_t = H(TeNB||L_t)$.

Both SeNB and TeNB broadcast their locations L_s, L_t to their respective neighboring eNBs. Here, we consider SeNB and TeNB are neighbors. Then, SeNB computes $h_t = H(TeNB||L_t)$ and sends back the signature $\sigma_{ts} = X_s \cdot h_t^{r_s}$ to TeNB. Similarly, TeNB computes $h_s = H(SeNB||L_s)$ and sends back the signature $\sigma_{st} = X_t \cdot h_s^{r_t}$ to SeNB. SeNB can verify σ_{st} by checking $e(\sigma_{st}, g)e(h_s, R_t) \stackrel{?}{=} A_t$, and TeNB can also verify σ_{ts} by checking $e(\sigma_{ts}, g)e(h_t, R_s) \stackrel{?}{=} A_s$. The correctness are obvious.

For the vehicle, it first chooses a random number $r_v \in \mathbb{Z}_q^*$ as its private key $sk_v = r_v$, and then calculates the corresponding public key $pk_v = g^{r_v}$.

6.2.3.3 Secure Handover Key Management Process

During the secure CoMP joint transmission handover key management process, the message flows between the vehicle, SeNB, TeNB and the core network (in terms of MME and S-GW) are shown in Fig. 6.11. The detailed steps of secure handover process are as follows:

FIGURE 6.11. Message flows during CoMP joint transmission handover

Step 1: The vehicle conducts the channel measurement. If the reference signal receiving power of TeNB is above a certain threshold (i.e., $RSRP_t \geq R_{th}$), the vehicle sends a channel measurement report towards SeNB, as shown in Fig. 6.11. Then SeNB initiates a ‘Cooperating Request’ towards TeNB together with the vehicle’s public key pk_v , through X2 interface. After some evaluations, TeNB will reply to SeNB with a ‘Cooperating ACK’, if TeNB owns enough resources for the CoMP joint transmission. Finally, SeNB forwards the ‘Cooperating ACK’ together with the public keys $pk_s = (R_s, A_s)$, $pk_t = (R_t, A_t)$ towards the vehicle.

Step 2: The vehicle first selects a random session key k_{st} , which will be shared among the vehicle, TeNB and SeNB. Then, the vehicle encrypts k_{st} with the public keys of SeNB and TeNB, as shown in Algorithm 5. After that, it sends $(C_{st}^1, C_{st}^2, C_{st}^3, \hat{\sigma}, \text{time})$ towards TeNB and SeNB.

Algorithm 5: Cooperating Session Key Encryption

Data: The symmetric key k_{st} , public keys (R_s, A_s) , (R_t, A_t) , and private key sk_v

1. Calculates $A = A_s \cdot A_t$, $R = R_s \cdot R_t$;
2. Chooses a random number $r \in \mathbb{Z}_q^*$, and computes $C_{st}^1 = g^r$, $C_{st}^2 = R^r$, $C_{st}^3 = k_{st} \cdot A^r$;
3. Uses $sk_v = r_v$ to generate the signature $\hat{\sigma}$ on k_{st} and timestamp **time**, which is $\hat{\sigma} = H(k_{st} || \text{time})^{r_v}$.

Output: $C_{st}^1, C_{st}^2, C_{st}^3, \hat{\sigma}, \text{time}$

Step 3: After receiving $(C_{st}^1, C_{st}^2, C_{st}^3, \hat{\sigma}, \text{time})$, SeNB decrypts k_{st} with σ_s and σ_{st} ; meanwhile, TeNB decrypts k_{st} with σ_t and σ_{ts} . SeNB and TeNB also verify the received signature $\hat{\sigma}$ with the public key of the vehicle $pk_v = g^{r_v}$. Here, we take SeNB as an example to check the decryption process, as shown in Algorithm 6.

Algorithm 6: Cooperating Session Key Decryption and Signature Verification

Data: The received ciphertext $(C_{st}^1, C_{st}^2, C_{st}^3, \hat{\sigma}, \text{time})$, and the signatures σ_s, σ_{st}

1. SeNB decrypts k_{st} with σ_s and σ_{st} ,

$$k_{st} = \frac{C_{st}^3}{e(\sigma_s \cdot \sigma_{st}, C_{st}^1) e(h_s, C_{st}^2)};$$

2. Verifies the signature,

$$e(g, \hat{\sigma}) \stackrel{?}{=} e(pk_v, H(k_{st} || \text{time}))$$

Output: k_{st}

Correctness. We first validate the correctness of decryption process of k_{st} in Algorithm 6:

$$\begin{aligned} k_{st} &= \frac{C_{st}^3}{e(\sigma_s \cdot \sigma_{st}, C_{st}^1) \cdot e(h_s, C_{st}^2)} \\ &= \frac{k_{st} \cdot e(X_s \cdot X_t, g)^r}{e(X_s \cdot X_t \cdot h_s^{(r_s+r_t)}, g^r) \cdot e(h_s, g^{-(r_s+r_t)r})} \\ &= \frac{k_{st} \cdot e(X_s \cdot X_t, g)^r}{e(X_s \cdot X_t, g^r) \cdot e(h_s^{r_s+r_t}, g^r) \cdot e(h_s, g^{-(r_s+r_t)r})} \end{aligned} \quad (6.8)$$

Then we validate the correctness of the k_{st} through signature verification,

$$\begin{aligned} e(g, \hat{\sigma}) &= e(g, H(k_{st} || \text{time})^{sk_v}) \\ &= e(g^{sk_v}, H(k_{st} || \text{time})) \\ &= e(pk_v, H(k_{st} || \text{time})) \end{aligned} \quad (6.9)$$

After the signature verification is finished, a secure session key k_{st} can be established among the vehicle, TeNB and SeNB. Meanwhile, the control-plane signaling exchanged with the MME and the user-plane data traffic exchanged with the S-GW, are still connected with SeNB.

Step 4: The vehicle generates a new session key k_t , which is established between the vehicle and TeNB after the CoMP joint transmission. Similar to Algorithm 5, the vehicle encrypts k_t with the public key of TeNB (R_t, A_t) : $C_t^1 = g^{r_1}$, $C_t^2 = R_t^{r_1}$, $C_t^3 = k_t \cdot A_t^{r_1}$, where $r_1 \in \mathbb{Z}_q^*$ is

a randomly chosen number, and it generates the corresponding signature $\hat{\sigma}' = H(k_t || \text{time}')^{sk_v}$, where time' is a new timestamp. In addition, the vehicle sends $C_t^1 || C_t^2 || C_t^3 || \hat{\sigma}' || \text{time}'$ towards TeNB.

Step 5: After receiving $C_t^1 || C_t^2 || C_t^3 || \hat{\sigma}' || \text{time}'$, TeNB decrypts k_t with σ_t , which is $k_t = \frac{C_t^3}{e(\sigma_t, C_t^1) \cdot e(h_t, C_t^2)}$. In addition, it validates the correctness of k_t through the signature verification, which is $e(g, \hat{\sigma}') \stackrel{?}{=} e(pk_v, H(k_t || \text{time}'))$. After the signature verification, TeNB keeps k_t in its storage.

Step 6: The vehicle keeps the channel measurement. If the reference signal receiving power of SeNB is above a certain threshold (i.e., $RSRP_S \leq R_{th}$), SeNB terminates the CoMP joint transmission and sends a ‘Handover Request’ towards TeNB, as shown in Fig. 6.11. Then TeNB replies with a ‘Handover Request’, and k_t is established between the vehicle and TeNB. Meanwhile, the control-plane signaling exchanged with the MME and the user-plane data traffic exchanged with the S-GW, are redirected to be connected with TeNB, as shown in the ‘S1 Packet Switch Request’.

Finally, we complete the secure key management process during the CoMP joint transmission handover.

6.2.4 Security Analysis

In this subsection, we briefly analyze the security properties of the proposed secure CoMP joint transmission handover key management scheme.

- To achieve the security goal of backward and forward key separation, in the proposed key management scheme, the session key k_{st} is randomly chosen by the vehicle, encrypted with product of the public keys of SeNB and TeNB, and then delivered towards SeNB and TeNB. Since the exploited asymmetric group key agreement technique [196] is proved to be secure as an encryption scheme, its signature is proved to be secure against existential forgery under a chosen message attack, and its aggregatability is also proved to be secure against the non-adaptively chosen message attack, the security of the proposed handover key management can

be guaranteed; i.e., without knowing the signature σ_s (or σ_t), it is infeasible for an adversary (e.g., a compromised eNB) to recover the correct k_{st} from $(C_{st}^1, C_{st}^2, C_{st}^3)$. In addition, a new random session key k_t is also generated and encrypted by the public key of TeNB pk_t , and then delivered towards TeNB. After receiving (C_t^1, C_t^2, C_t^3) , TeNB recovers k_t with the message-signature pair (h_t, σ_t) . After the CoMP joint transmission, the previous session key k_{st} is no longer utilized and k_t is established between the vehicle and TeNB, and SeNB cannot recover k_t since it does not know the value of σ_t . Due to the fact that the session keys are randomly generated and can only be utilized once, based on k_{st} , it is impossible to infer k_t , or vice versa. Thus, the security goal of backward and forward key separation can be successfully achieved during the secure CoMP joint transmission handover key establishment process.

- In the proposed secure key management scheme, during the CoMP joint transmission, a secure session key can be established between the vehicle and the cooperating eNBs. Since k_{st} is encrypted with product of the public keys of SeNB and TeNB ($R = R_s \cdot R_t, A = A_s \cdot A_t$), each cooperating eNB can decrypt k_{st} with h_s (or h_t) and the corresponding signatures σ_s and σ_{st} (or σ_t and σ_{ts}), as shown in Eq. 6.8. After decryption, each cooperating eNB can verify the signature $\hat{\sigma}$ on k_{st} and timestamp time with the public key of the vehicle pk_v to ensure that the session key k_{st} is valid and generated by the vehicle, as shown in Eq. 6.9, and establish a secure session key with the vehicle. Thus, the vehicle can benefit from the diversity gain brought by the multiple transmissions.

By summarizing the above, our proposed scheme can achieve the desirable security properties.

6.2.5 Performance Evaluation

In this subsection, we evaluate the performance of the proposed secure handover key management scheme in the CoMP joint transmission handover scenario, in terms of the computational delay and the eNB valid serving ability.

The simulation parameters are listed in Table 6.3, referring to the network scenario setup in [132] and the traffic scenario setup in [197]. The experimental environment is a desktop

FIGURE 6.12. eNB valid serving ratio with different vehicle density and different average vehicle speed

TABLE 6.3. Parameters Settings

Simulation Parameters	Value
Cell Radius R_{radius}	3 km
Site-to-site Distance D	4.8 km
Vertical Distance between eNB and Road	30 m
Road Structure	Four-lane Two-way Highway
Vehicle Speed v	10 m/s to 40 m/s (36 km/h to 144 km/h)

with dual-core 3.10 GHz processor and 8.00 GB installed memory, based on the type A pairing in the Java Pairing-Based Cryptography Library [174].

We first calculate the communication overhead introduced by the proposed scheme, based on the assumption that both SeNB and TeNB can successfully decrypt the received messages. In our simulation setup, the length of each generator is 128 bytes and the traffic load introduced by the secure key management scheme is $128 * 8 = 1024$ bytes. To examine the execution performance, we test the time overhead introduced to the vehicle, SeNB and TeNB 1000 times, and the testing results are $t_v = 110.5$ ms, $t_s = 34.5$ ms and $t_t = 17.3$ ms. Since the generation and encryption of k_{st} and k_t can be conducted offline (before the handover process), t_v only influences the performance of the vehicle and does not influence the key establishment performance. To measure the serving capability of TeNB, we first estimate the maximal number of vehicles

that TeNB can serve. Given average travel speed v , coverage range R , and the processing time $2 * (t_t + t_s)$ (TeNB has two neighboring eNBs and there are vehicles leaving the coverage of TeNB since the highway is two-way), the maximal number of vehicles that TeNB can serve is $S_{max} = \frac{R}{2*v*(t_t+t_s)}$. Similar to [197], given requesting vehicle density S_{req} (including UEs located inside vehicles since they are also moving along the highway), we define TeNB serving ratio S_{ratio} as

$$S_{ratio} = \begin{cases} 1 & , \quad S_{req} \leq S_{max} \\ S_{max}/S_{req} & , \quad S_{req} > S_{max} \end{cases} \quad (6.10)$$

We test TeNB serving ratio 1000 times, and the valid CoMP joint transmission request range is set to be $R = (2 * R_{radius} - D)/2 = 600$ m (which is half of the overlapping area of TeNB and SeNB, and the chosen R can effectively help to achieve the diversity gain brought by the two cooperating eNBs). We assume the vehicle arrival density S_{req} follows the Poisson distribution with the mean parameter varies from 10 to 200. As shown in Fig. 6.12, the average valid TeNB serving ratio decreases with the increase of the average vehicle speed v and the vehicle density. Therefore, we can conclude that the proposed secure handover key management scheme is feasible for most of the network and traffic scenarios.

6.3 Summary

In this chapter, we have proposed two secure handover key management schemes in mobile relay networks and in CoMP joint transmission respectively. In the first mobile relay based scheme, secure session keys are established between the on-board UEs and the target DeNB while exploiting the third-party owned MRNs during key establishment, without disclosing the content of the session keys. While in the second CoMP joint transmission based scheme, the established session keys can help the vehicle to achieve the diversity gain brought by the CoMP joint transmission, and it can also be adaptive to the limited backhaul bandwidth.

Chapter 7

Conclusions and Future Work

In this chapter, a summary of contributions of this dissertation is performed, future research works are identified, and a final remark is made.

7.1 Contributions

The major contributions of this thesis are summarized as follows, which is also illustrated in Fig. 7.1:

- For the secure and privacy-preserving vehicular data processing, a novel privacy-preserving range query scheme has been first proposed, which enables the data requester to accurately retrieve the sensed data from the distributed on-board storage with location privacy preservation. The proposed scheme exploits structured scalars to denote the querying and sensing locations, and achieves the privacy-preserving data filtering with the homomorphic Paillier cryptosystem technique.
- Secondly, to improve the communication overhead caused by the situation when the sensory data are maintained on-board, an efficient and location privacy-preserving data sharing scheme is proposed, which enables the aggregation and acquisition of the vehicular

FIGURE 7.1. Structure of the proposed schemes in this thesis

sensors captured data at the network edge with the re-encryption technique, i.e., without the involvement of the trusted central entity.

- Thirdly, to disrupt the linkage between a data request and a specific vehicle, an invertible matrix is introduced to structure multiple data requests in a secure vehicular data dissemination scheme, such that the RSU can recover the content of the data queries while protecting the unlinkability between the data query and its origin.
- For the secure session key establishment schemes in LTE-A vehicular networks, two novel secure handover key management schemes in LTE-A vehicular networks are proposed. Specifically, in the first scheme, secure session keys are established between the on-board UEs and the target DeNB while exploiting the third-party owned mobile relays during key establishment process, without disclosing the content of the session keys. While in the second scheme, the established session keys can help the vehicle to achieve

the diversity gain brought by the CoMP joint transmission, and it can also be adaptive to the limited backhaul bandwidth.

7.2 Future Work

Although our research have brought great advancements in securing vehicular networks, security issues in vehicular networks are still promising to support the wide deployment of various applications and there exist several research directions need to be further investigated as an extension of my Ph.D. dissertation. Thus, works following two research topics will be performed as a continuation of my Ph.D. dissertation.

- *Security schemes considering collusion attacks in vehicular cloud.* In vehicular cloud, the secure data processing (including data aggregation, data queries, and data computation) are performed between the vehicles and roadside infrastructures. However, the vehicles are independently owned and the roadside infrastructures are deployed in an open environment, which make it possible for collusion attacks between the vehicle and the roadside infrastructure and that between vehicles. In the current schemes, we almost do not take the collusion attack into consideration, which may cause the location privacy disclosure and malfunctioning of the whole system. Therefore, as one of our future work, we propose to take the collusion attack into consideration for the security protocols proposed for the secure vehicular data processing.
- *Secure handover schemes between VANETs and LTE-A vehicular networks.* Although enabling a broad range of safety and non-safety applications for VANETs, the roadside infrastructure is still sparsely deployed and the defined transmission range is still limited, which causes the short-range and intermittent data transmission. In LTE-A vehicular networks, eNBs are widely and intensively deployed which can provide the high quality V2I communications even at high moving speed. However, it cannot support the V2V communications. In order to fully exploit the benefits of the LTE and 802.11p interfaces,

as a future work, the secure handover key management schemes need to be designed under the heterogeneous network.

Appendix

A. Author's Publications

A.1 Journal Papers

[J1] **Qinglei Kong**, Rongxing Lu, Shuo Chen, Hui Zhu: Achieve Secure Handover Session Key Management via Mobile Relay in LTE-Advanced Networks. *IEEE Internet of Things Journal* 4(1): 29-39 (2017). **Impact factor: 7.60**

[J2] **Qinglei Kong**, Rongxing Lu, Maode Ma, Haiyong Bao: Achieve Location Privacy-Preserving Range Query in Vehicular Sensing. *Sensors* 17(8): 1829 (2017). **Impact factor: 2.677**

[J3] **Qinglei Kong**, Rongxing Lu, Maode Ma, Haiyong Bao: A Privacy-Preserving Sensory Data Sharing Scheme in Internet of Vehicles. Accepted by *Future Generation Computer Systems*. **Impact factor: 3.997**

[J4] **Qinglei Kong**, Rongxing Lu, Hui Zhu, Maode Ma: Achieving Secure and Privacy-Preserving Incentive in Vehicular Cloud. Accepted by *IEEE Access*. **Impact factor: 3.224**

[J5] **Qinglei Kong**, Rongxing Lu, Maode Ma, Haiyong Bao: A Privacy-preserving and Verifiable Data Dissemination Scheme in Vehicular Cloud. Submitted to *IEEE Transactions on Vehicular Technology*

A.2 Conference Papers

[C1] **Qinglei Kong**, Rongxing Lu, Hui Zhu, Abdulrahman Alamer, Xiaodong Lin: A Secure and Privacy-Preserving Incentive Framework for Vehicular Cloud on the Road. 2016 IEEE Global Communications Conference, GLOBECOM 2016, Washington, DC, USA, December 4-8, 2016

[C2] **Qinglei Kong**, Maode Ma, Rongxing Lu: Achieving Secure CoMP Joint Transmission Handover in LTE-A Vehicular Networks. 2017 IEEE 86th Vehicular Technology Conference, VTC Fall 2017, Toronto, ON, Canada, September 24-27, 2017

[C3] **Qinglei Kong**, Maode Ma: Achieving Secure Vehicle-to-vehicle Key Establishment in LTE-A Networks, Accepted by *The 16th ITS Asia-Pacific Forum FUKUOKA 2018*.

Bibliography

- [1] C. Han and H. Choi, "Security analysis of handover key management in 4g LTE/SAE networks," *IEEE Trans. Mob. Comput.*, vol. 13, no. 2, pp. 457–468, 2014.
- [2] D. Jiang and L. Delgrossi, "Ieee 802.11p: Towards an international standard for wireless access in vehicular environments," in *VTC Spring 2008 - IEEE Vehicular Technology Conference*, 2008, pp. 2036–2040.
- [3] Z. H. Mir and F. Filali, "Lte and ieee 802.11 p for vehicular networking: a performance evaluation," *EURASIP Journal on Wireless Communications and Networking*, vol. 2014, no. 1, p. 89, 2014.
- [4] X. Wang, S. Mao, and M. X. Gong, "An overview of 3gpp cellular vehicle-to-everything standards," *GetMobile*, vol. 21, no. 3, pp. 19–25, 2017.
- [5] G. Araniti, C. Campolo, M. Condoluci, A. Iera, and A. Molinaro, "LTE for vehicular networking: a survey," *IEEE Communications Magazine*, vol. 51, no. 5, pp. 148–157, 2013.
- [6] A. S. Prasad and S. Rao, "A mechanism design approach to resource procurement in cloud computing," *IEEE Trans. Computers*, vol. 63, no. 1, pp. 17–30, 2014.
- [7] M. Whaiduzzaman, M. Sookhak, A. Gani, and R. Buyya, "A survey on vehicular cloud computing," *J. Network and Computer Applications*, vol. 40, pp. 325–344, 2014.
- [8] S. Olariu, M. Eltoweissy, and M. F. Younis, "Towards autonomous vehicular clouds," *EAI Endorsed Trans. Mobile Communications Applications*, vol. 1, no. 1, p. e2, 2011.

- [9] R. Yu, Y. Zhang, S. Gjessing, W. Xia, and K. Yang, "Toward cloud-based vehicular networks with efficient resource management," *IEEE Network*, vol. 27, no. 5, pp. 48–55, 2013.
- [10] E. Lee, E. Lee, M. Gerla, and S. Oh, "Vehicular cloud networking: architecture and design principles," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 148–155, 2014.
- [11] W. He, G. Yan, and L. D. Xu, "Developing vehicular data cloud services in the iot environment," *IEEE Trans. Industrial Informatics*, vol. 10, no. 2, pp. 1587–1595, 2014.
- [12] M. Gerla, "Vehicular cloud computing," in *The 11th Annual Mediterranean Ad Hoc Networking Workshop, Med-Hoc-Net 2012, Ayia Napa, Cyprus, June 19-22, 2012*, 2012, pp. 152–155.
- [13] S. Arif, S. Olariu, J. Wang, G. Yan, W. Yang, and I. Khalil, "Datacenter at the airport: Reasoning about time-dependent parking lot occupancy," *IEEE Trans. Parallel Distrib. Syst.*, vol. 23, no. 11, pp. 2067–2080, 2012.
- [14] W. Kim and M. Gerla, "NAVOPT: navigator assisted vehicular route optimizer," in *Proceedings of the Fifth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing, IMIS 2011, Seoul, Korea, June 30-July 02, 2011*, 2011, pp. 450–455.
- [15] A. Zubaida, A. Saleh, M. Rashid, and A. M. B., "Intelligent disaster management system based on cloud-enabled vehicular networks," in *ITS Telecommunications (ITST), 2011 11th International Conference on*, 2011, pp. 361–368.
- [16] U. Lee, B. Zhou, M. Gerla, E. Magistretti, P. Bellavista, and A. Corradi, "Mobeyes: smart mobs for urban monitoring with a vehicular sensor network," *IEEE Wireless Commun.*, vol. 13, no. 5, pp. 52–57, 2006.
- [17] S. Kumar, L. Shi, N. Ahmed, S. Gil, D. Katabi, and D. Rus, "Carspeak: a content-centric network for autonomous driving," in *ACM SIGCOMM 2012 Conference, SIGCOMM '12, Helsinki, Finland - August 13 - 17, 2012*, 2012, pp. 259–270.

-
- [18] R. Lu, X. Lin, X. Liang, and X. S. Shen, "Secure provenance: the essential of bread and butter of data forensics in cloud computing," in *Proceedings of the 5th ACM Symposium on Information, Computer and Communications Security, ASIACCS 2010, Beijing, China, April 13-16, 2010*, 2010, pp. 282–292.
- [19] G. Yan, D. Wen, S. Olariu, and M. C. Weigle, "Security challenges in vehicular cloud computing," *IEEE Trans. Intelligent Transportation Systems*, vol. 14, no. 1, pp. 284–294, 2013.
- [20] F. Ahmad, M. Kazim, A. Adnane, and A. Awad, "Vehicular cloud networks: Architecture, applications and security issues," in *8th IEEE/ACM International Conference on Utility and Cloud Computing, UCC 2015, Limassol, Cyprus, December 7-10, 2015*, 2015, pp. 571–576.
- [21] J. Yuan and S. Yu, "Flexible and publicly verifiable aggregation query for outsourced databases in cloud," in *IEEE Conference on Communications and Network Security, CNS 2013, National Harbor, MD, USA, October 14-16, 2013*, 2013, pp. 520–524.
- [22] X. Chen, J. Li, J. Weng, J. Ma, and W. Lou, "Verifiable computation over large database with incremental updates," *IEEE Trans. Computers*, vol. 65, no. 10, pp. 3184–3195, 2016.
- [23] "Ieee standard for wireless access in vehicular environments—security services for applications and management messages," *IEEE Std 1609.2-2016 (Revision of IEEE Std 1609.2-2013)*, pp. 1–240, 2016.
- [24] T. W. Chim, S. Yiu, L. C. K. Hui, and V. O. K. Li, "VSPN: vanet-based secure and privacy-preserving navigation," *IEEE Trans. Computers*, vol. 63, no. 2, pp. 510–524, 2014.
- [25] X. Lin, X. Sun, P. Ho, and X. Shen, "GSIS: A secure and privacy-preserving protocol for vehicular communications," *IEEE Trans. Vehicular Technology*, vol. 56, no. 6, pp. 3442–3456, 2007.

- [26] R. Lu, X. Lin, X. Liang, and X. S. Shen, "A dynamic privacy-preserving key management scheme for location-based services in vanets," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 1, pp. 127–139, 2012.
- [27] J. Sun, C. Zhang, Y. Zhang, and Y. Fang, "An identity-based security system for user privacy in vehicular ad hoc networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 21, no. 9, pp. 1227–1239, 2010.
- [28] R. Lu, Y. Rahulamathavan, H. Zhu, C. Xu, and M. Wang, "Security and privacy challenges in vehicular cloud computing," *Mobile Information Systems*, vol. 2016, pp. 6 812 816:1–6 812 816:2, 2016.
- [29] A. Daeinabi and A. G. Rahbar, "Detection of malicious vehicles (DMV) through monitoring in vehicular ad-hoc networks," *Multimedia Tools Appl.*, vol. 66, no. 2, pp. 325–338, 2013.
- [30] K. Shim, "CPAS: an efficient conditional privacy-preserving authentication scheme for vehicular sensor networks," *IEEE Trans. Vehicular Technology*, vol. 61, no. 4, pp. 1874–1883, 2012.
- [31] X. Jia, X. Yuan, L. Meng, and L. Wang, "EPAS: efficient privacy-preserving authentication scheme for vanets-based emergency communication," *JSW*, vol. 8, no. 8, pp. 1914–1922, 2013.
- [32] D. Huang, Z. Zhou, L. Xu, T. Xing, and Y. Zhong, "Secure data processing framework for mobile cloud computing," in *2011 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 2011, pp. 614–618.
- [33] R. Gummadi, H. Balakrishnan, P. Maniatis, and S. Ratnasamy, "Not-a-bot: Improving service availability in the face of botnet attacks." in *NSDI*, vol. 9, 2009, pp. 307–320.
- [34] P. Gilbert, L. P. Cox, J. Jung, and D. Wetherall, "Toward trustworthy mobile sensing," in *Eleventh Workshop on Mobile Computing Systems and Applications, HotMobile '10, Annapolis, Maryland, USA, February 22-23, 2010*, 2010, pp. 31–36.

- [35] C. Cornelius, A. Kapadia, D. Kotz, D. Peebles, M. Shin, and N. Triandopoulos, "Anonymense: privacy-aware people-centric sensing," in *Proceedings of the 6th International Conference on Mobile Systems, Applications, and Services (MobiSys 2008)*, Breckenridge, CO, USA, June 17-20, 2008, 2008, pp. 211–224.
- [36] D. Christin, C. Roßkopf, M. Hollick, L. A. Martucci, and S. S. Kanhere, "Incognisense: An anonymity-preserving reputation framework for participatory sensing applications," *Pervasive and Mobile Computing*, vol. 9, no. 3, pp. 353–371, 2013.
- [37] X. O. Wang, W. Cheng, P. Mohapatra, and T. F. Abdelzaher, "Enabling reputation and trust in privacy-preserving mobile sensing," *IEEE Trans. Mob. Comput.*, vol. 13, no. 12, pp. 2777–2790, 2014.
- [38] Y. Zhang, Q. Chen, and S. Zhong, "Efficient and privacy-preserving min and k th min computations in mobile sensing systems," *IEEE Trans. Dependable Sec. Comput.*, vol. 14, no. 1, pp. 9–21, 2017.
- [39] Q. Li and G. Cao, "Providing privacy-aware incentives for mobile sensing," in *2013 IEEE International Conference on Pervasive Computing and Communications, PerCom 2013, San Diego, CA, USA, March 18-22, 2013*, 2013, pp. 76–84.
- [40] B. Liu, Y. Jiang, F. Sha, and R. Govindan, "Cloud-enabled privacy-preserving collaborative learning for mobile sensing," in *The 10th ACM Conference on Embedded Network Sensor Systems, SenSys '12, Toronto, ON, Canada, November 6-9, 2012*, 2012, pp. 57–70.
- [41] M. Chaqfeh, A. Lakas, and I. Jawhar, "A survey on data dissemination in vehicular ad hoc networks," *Vehicular Communications*, vol. 1, no. 4, pp. 214–225, 2014.
- [42] J. Brønsted and L. M. Kristensen, "Specification and performance evaluation of two zone dissemination protocols for vehicular ad-hoc networks," in *Proceedings 39th Annual Simulation Symposium (ANSS-39 2006)*, 2-6 April 2006, Huntsville, Alabama, USA, 2006, pp. 68–79.

- [43] Y. Bi, H. Zhao, and X. S. Shen, "A directional broadcast protocol for emergency message exchange in inter-vehicle communications," in *Proceedings of IEEE International Conference on Communications, ICC 2009, Dresden, Germany, 14-18 June 2009*, 2009, pp. 1–5.
- [44] C. Huang and S. Y. Lin, "An advanced vehicle collision warning algorithm over the DSRC communication environment: An advanced vehicle collision warning algorithm," in *27th IEEE International Conference on Advanced Information Networking and Applications, AINA 2013, Barcelona, Spain, March 25-28, 2013*, 2013, pp. 696–702.
- [45] D. J. Sun, X. Ni, and L. Zhang, "A discriminated release strategy for parking variable message sign display problem using agent-based simulation," *IEEE Trans. Intelligent Transportation Systems*, vol. 17, no. 1, pp. 38–47, 2016.
- [46] A. Daniel, A. Paul, A. Ahmad, and S. Rho, "Cooperative intelligence of vehicles for intelligent transportation systems (ITS)," *Wireless Personal Communications*, vol. 87, no. 2, pp. 461–484, 2016.
- [47] A. Nandan, S. Das, G. Pau, M. Gerla, and M. Y. Sanadidi, "Co-operative downloading in vehicular ad-hoc wireless networks," in *2nd International Conference on Wireless on Demand Network Systems and Service (WONS 2005), 19-21 January 2005, St. Moritz, Switzerland*, 2005, pp. 32–41.
- [48] Y. Leng and L. Zhao, "Novel design of intelligent internet-of-vehicles management system based on cloud-computing and internet-of-things," in *International Conference on Electronic and Mechanical Engineering and Information Technology, EMEIT 2011, Harbin, Heilongjiang, China, 12-14 August, 2011*, 2011, pp. 3190–3193.
- [49] R. Kim, H. Lim, and B. Krishnamachari, "Prefetching-based data dissemination in vehicular cloud systems," *IEEE Trans. Vehicular Technology*, vol. 65, no. 1, pp. 292–306, 2016.

-
- [50] N. Kumar and J. Lee, "Peer-to-peer cooperative caching for data dissemination in urban vehicular communications," *IEEE Systems Journal*, vol. 8, no. 4, pp. 1136–1144, 2014.
- [51] K. Shafiee and V. C. M. Leung, "A reliable robust fully ad hoc data dissemination mechanism for vehicular networks," in *The Second International Conference on Future Generation Communication and Networking, FGCN 2008, Volume 2, Workshops, Hainan Island, China, December 13-15, 2008*, 2008, pp. 369–373.
- [52] J. Zhao and G. Cao, "VADD: vehicle-assisted data delivery in vehicular *Ad Hoc* networks," *IEEE Trans. Vehicular Technology*, vol. 57, no. 3, pp. 1910–1922, 2008.
- [53] J. Bronsted and L. M. Kristensen, "Specification and performance evaluation of two zone dissemination protocols for vehicular ad-hoc networks," in *Proceedings of the 39th annual Symposium on Simulation*, 2006, pp. 68–79.
- [54] I. Achour, T. Bejaoui, A. Busson, and S. Tabbane, "SEAD: A simple and efficient adaptive data dissemination protocol in vehicular ad-hoc networks," *Wireless Networks*, vol. 22, no. 5, pp. 1673–1683, 2016.
- [55] S. M. Bilal, C. J. Bernardos, and C. Guerrero, "Position-based routing in vehicular networks: A survey," *J. Network and Computer Applications*, vol. 36, no. 2, pp. 685–697, 2013.
- [56] J. Gong, C. Xu, and J. Holle, "Predictive directional greedy routing in vehicular ad hoc networks," in *27th International Conference on Distributed Computing Systems Workshops (ICDCS 2007 Workshops), June 25-29, 2007, Toronto, Ontario, Canada*, 2007, p. 2.
- [57] K. Cho and M. Ryu, "A survey of greedy routing protocols for vehicular ad hoc networks," *Smart CR*, vol. 2, no. 2, pp. 125–137, 2012.
- [58] X. Lin, X. Sun, X. Wang, C. Zhang, P. Ho, and X. Shen, "TSVC: timed efficient and secure vehicular communications with privacy preserving," *IEEE Trans. Wireless Communications*, vol. 7, no. 12-1, pp. 4987–4998, 2008.

- [59] W. Shen, L. Liu, X. Cao, Y. Hao, and Y. Cheng, “Cooperative message authentication in vehicular cyber-physical systems,” *IEEE Trans. Emerging Topics Comput.*, vol. 1, no. 1, pp. 84–97, 2013.
- [60] R. S. Bali and N. Kumar, “Secure clustering for efficient data dissemination in vehicular cyber-physical systems,” *Future Generation Comp. Syst.*, vol. 56, pp. 476–492, 2016.
- [61] C. Castelluccia, A. C. Chan, E. Mykletun, and G. Tsudik, “Efficient and provably secure aggregation of encrypted data in wireless sensor networks,” *TOSN*, vol. 5, no. 3, pp. 20:1–20:36, 2009.
- [62] E. Shi, T. H. Chan, E. G. Rieffel, R. Chow, and D. Song, “Privacy-preserving aggregation of time-series data,” in *Proceedings of the Network and Distributed System Security Symposium, NDSS 2011, San Diego, California, USA, 6th February - 9th February 2011*, 2011.
- [63] R. Lu, X. Liang, X. Li, X. Lin, and X. Shen, “EPPA: an efficient and privacy-preserving aggregation scheme for secure smart grid communications,” *IEEE Trans. Parallel Distrib. Syst.*, vol. 23, no. 9, pp. 1621–1631, 2012.
- [64] R. Lu, K. Heung, A. H. Lashkari, and A. A. Ghorbani, “A lightweight privacy-preserving data aggregation scheme for fog computing-enhanced iot,” *IEEE Access*, vol. 5, pp. 3302–3312, 2017.
- [65] L. Chen, R. Lu, Z. Cao, K. Alharbi, and X. Lin, “Muda: Multifunctional data aggregation in privacy-preserving smart grid communications,” *Peer-to-Peer Networking and Applications*, vol. 8, no. 5, pp. 777–792, 2015.
- [66] S. Özdemir and Y. Xiao, “Secure data aggregation in wireless sensor networks: A comprehensive overview,” *Computer Networks*, vol. 53, no. 12, pp. 2022–2037, 2009.
- [67] H. Chan, A. Perrig, B. Przydatek, and D. X. Song, “SIA: secure information aggregation in sensor networks,” *Journal of Computer Security*, vol. 15, no. 1, pp. 69–102, 2007.

-
- [68] S. Özdemir, “Concealed data aggregation in heterogeneous sensor networks using privacy homomorphism,” in *Proceedings of the IEEE International Conference on Pervasive Services, ICPS 2007, 15-20 July, 2007, Istanbul, Turkey, 2007*, pp. 165–168.
- [69] C. Li, R. Lu, H. Li, L. Chen, and J. Chen, “PDA: a privacy-preserving dual-functional aggregation scheme for smart grid communications,” *Security and Communication Networks*, vol. 8, no. 15, pp. 2494–2506, 2015.
- [70] H. Bao and R. Lu, “A lightweight data aggregation scheme achieving privacy preservation and data integrity with differential privacy and fault tolerance,” *Peer-to-Peer Networking and Applications*, vol. 10, no. 1, pp. 106–121, 2017.
- [71] R. Lu, K. Alharbi, X. Lin, and C. Huang, “A novel privacy-preserving set aggregation scheme for smart grid communications,” in *2015 IEEE Global Communications Conference, GLOBECOM 2015, San Diego, CA, USA, December 6-10, 2015*, 2015, pp. 1–6.
- [72] H. Bao and R. Lu, “A new differentially private data aggregation with fault tolerance for smart grid communications,” *IEEE Internet of Things Journal*, vol. 2, no. 3, pp. 248–258, 2015.
- [73] G. Zhuo, Q. Jia, L. Guo, M. Li, and P. Li, “Privacy-preserving verifiable data aggregation and analysis for cloud-assisted mobile crowdsourcing,” in *35th Annual IEEE International Conference on Computer Communications, INFOCOM 2016, San Francisco, CA, USA, April 10-14, 2016*, 2016, pp. 1–9.
- [74] X. Liu, Y. Zhang, B. Wang, and H. Wang, “An anonymous data aggregation scheme for smart grid systems,” *Security and Communication Networks*, vol. 7, no. 3, pp. 602–610, 2014.
- [75] H. Zhu, L. Gao, and H. Li, “Secure and privacy-preserving body sensor data collection and query scheme,” *Sensors*, vol. 16, no. 2, p. 179, 2016.

- [76] L. Chen, R. Lu, and Z. Cao, "PDAFT: A privacy-preserving data aggregation scheme with fault tolerance for smart grid communications," *Peer-to-Peer Networking and Applications*, vol. 8, no. 6, pp. 1122–1132, 2015.
- [77] D. Boneh, E. Goh, and K. Nissim, "Evaluating 2-dnf formulas on ciphertexts," in *Theory of Cryptography, Second Theory of Cryptography Conference, TCC 2005, Cambridge, MA, USA, February 10-12, 2005, Proceedings*, 2005, pp. 325–341.
- [78] J. Vaidya and C. Clifton, "Privacy preserving association rule mining in vertically partitioned data," in *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, July 23-26, 2002, Edmonton, Alberta, Canada, 2002*, pp. 639–644.
- [79] B. Goethals, S. Laur, H. Lipmaa, and T. Mielikäinen, "On private scalar product computation for privacy-preserving data mining," in *Information Security and Cryptology - ICISC 2004, 7th International Conference, Seoul, Korea, December 2-3, 2004, Revised Selected Papers*, 2004, pp. 104–120.
- [80] A. Amirbekyan and V. Estivill-Castro, "A new efficient privacy-preserving scalar product protocol," in *Data Mining and Analytics 2007, Proceedings of the Sixth Australasian Data Mining Conference (AusDM 2007), Gold Coast, Queensland, Australia, December 3-4, 2007, Proceedings*, 2007, pp. 209–214.
- [81] M. Li, N. Cao, S. Yu, and W. Lou, "Findu: Privacy-preserving personal profile matching in mobile social networks," in *INFOCOM 2011. 30th IEEE International Conference on Computer Communications, Joint Conference of the IEEE Computer and Communications Societies, 10-15 April 2011, Shanghai, China, 2011*, pp. 2435–2443.
- [82] R. Lu, X. Lin, X. Liang, and X. Shen, "A secure handshake scheme with symptoms-matching for mhealthcare social network," *MONET*, vol. 16, no. 6, pp. 683–694, 2011.

-
- [83] R. Zhang, J. Zhang, Y. Zhang, J. Sun, and G. Yan, "Privacy-preserving profile matching for proximity-based mobile social networking," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9-Supplement, pp. 656–668, 2013.
- [84] R. Lu, X. Lin, and X. S. Shen, "SPOC: A secure and privacy-preserving opportunistic computing framework for mobile-healthcare emergency," *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 3, pp. 614–624, 2013.
- [85] R. Chergui, "Privacy preserving distributed profile matching in mobile social network," *CoRR*, vol. abs/1502.06993, 2015.
- [86] G. Zhuo, Q. Jia, L. Guo, M. Li, and Y. Fang, "Privacy-preserving verifiable proximity test for location-based services," in *2015 IEEE Global Communications Conference, GLOBECOM 2015, San Diego, CA, USA, December 6-10, 2015*, 2015, pp. 1–6.
- [87] J. Yuan and S. Yu, "Privacy preserving back-propagation neural network learning made practical with cloud computing," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 1, pp. 212–221, 2014.
- [88] I. Leontiadis, M. Önen, R. Molva, M. J. Chorley, and G. B. Colombo, "Privacy preserving similarity detection for data analysis," in *2013 International Conference on Cloud and Green Computing, Karlsruhe, Germany, September 30 - October 2, 2013*, 2013, pp. 547–552.
- [89] R. Lu, H. Zhu, X. Liu, J. K. Liu, and J. Shao, "Toward efficient and privacy-preserving computing in big data era," *IEEE Network*, vol. 28, no. 4, pp. 46–50, 2014.
- [90] Z. Gheid and Y. Challal, "An efficient and privacy-preserving similarity evaluation for big data analytics," in *8th IEEE/ACM International Conference on Utility and Cloud Computing, UCC 2015, Limassol, Cyprus, December 7-10, 2015*, 2015, pp. 281–289.
- [91] K. Shilton, "Four billion little brothers?: privacy, mobile phones, and ubiquitous data collection," *Commun. ACM*, vol. 52, no. 11, pp. 48–53, 2009.

- [92] R. Lu, X. Lin, T. H. Luan, X. Liang, and X. S. Shen, "Pseudonym changing at social spots: An effective strategy for location privacy in vanets," *IEEE Trans. Vehicular Technology*, vol. 61, no. 1, pp. 86–96, 2012.
- [93] —, "Pseudonym changing at social spots: An effective strategy for location privacy in vanets," *IEEE Trans. Vehicular Technology*, vol. 61, no. 1, pp. 86–96, 2012.
- [94] 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.
- [95] D. Huang, S. Misra, M. Verma, and G. Xue, "PACP: an efficient pseudonymous authentication-based conditional privacy protocol for vanets," *IEEE Trans. Intelligent Transportation Systems*, vol. 12, no. 3, pp. 736–746, 2011.
- [96] S. Jiang, X. Zhu, and L. Wang, "An efficient anonymous batch authentication scheme based on HMAC for vanets," *IEEE Trans. Intelligent Transportation Systems*, vol. 17, no. 8, pp. 2193–2204, 2016.
- [97] J. Freudiger, M. Raya, M. Félegyházi, P. Papadimitratos, and J.-P. Hubaux, "Mix-zones for location privacy in vehicular networks," in *ACM Workshop on Wireless Networking for Intelligent Transportation Systems (WiN-ITS)*, no. LCA-CONF-2007-016, 2007.
- [98] G. Ghinita, P. Kalnis, A. Khoshgozaran, C. Shahabi, and K. Tan, "Private queries in location based services: anonymizers are not necessary," in *Proceedings of the ACM SIGMOD International Conference on Management of Data, SIGMOD 2008, Vancouver, BC, Canada, June 10-12, 2008*, 2008, pp. 121–132.
- [99] L. Li, R. Lu, and C. Huang, "EPLQ: efficient privacy-preserving location-based query over outsourced encrypted data," *IEEE Internet of Things Journal*, vol. 3, no. 2, pp. 206–218, 2016.
- [100] T. Xu and Y. Cai, "Exploring historical location data for anonymity preservation in location-based services," in *INFOCOM 2008. 27th IEEE International Conference on*

- Computer Communications, Joint Conference of the IEEE Computer and Communications Societies, 13-18 April 2008, Phoenix, AZ, USA, 2008, pp. 547–555.*
- [101] W. Ni, M. Gu, and X. Chen, “Location privacy-preserving k nearest neighbor query under user’s preference,” *Knowl.-Based Syst.*, vol. 103, pp. 19–27, 2016.
- [102] K. Vu, R. Zheng, and J. Gao, “Efficient algorithms for k-anonymous location privacy in participatory sensing,” in *Proceedings of the IEEE INFOCOM 2012, Orlando, FL, USA, March 25-30, 2012, 2012*, pp. 2399–2407.
- [103] K. L. Huang, S. S. Kanhere, and W. Hu, “Preserving privacy in participatory sensing systems,” *Computer Communications*, vol. 33, no. 11, pp. 1266–1280, 2010.
- [104] A. Machanavajjhala, J. Gehrke, D. Kifer, and M. Venkatasubramanian, “l-diversity: Privacy beyond k-anonymity,” in *Proceedings of the 22nd International Conference on Data Engineering, ICDE 2006, 3-8 April 2006, Atlanta, GA, USA, 2006*, p. 24.
- [105] N. Li, T. Li, and S. Venkatasubramanian, “t-closeness: Privacy beyond k-anonymity and l-diversity,” in *Proceedings of the 23rd International Conference on Data Engineering, ICDE 2007, The Marmara Hotel, Istanbul, Turkey, April 15-20, 2007, 2007*, pp. 106–115.
- [106] B. Niu, Q. Li, X. Zhu, G. Cao, and H. Li, “Achieving k-anonymity in privacy-aware location-based services,” in *2014 IEEE Conference on Computer Communications, INFOCOM 2014, Toronto, Canada, April 27 - May 2, 2014, 2014*, pp. 754–762.
- [107] S. Tian, Y. Cai, and Q. Zheng, “A hybrid approach for privacy-preserving processing of knn queries in mobile database systems,” in *22nd ACM International Conference on Information and Knowledge Management, CIKM’13, San Francisco, CA, USA, October 27 - November 1, 2013, 2013*, pp. 1161–1164.
- [108] B. Niu, Q. Li, X. Zhu, and H. Li, “A fine-grained spatial cloaking scheme for privacy-aware users in location-based services,” in *23rd International Conference on Computer*

- Communication and Networks, ICCCN 2014, Shanghai, China, August 4-7, 2014*, 2014, pp. 1–8.
- [109] M. Ghaffari, N. Ghadiri, M. H. Manshaei, and M. S. Lahijani, “P4QS: A peer to peer privacy preserving query service for location-based mobile applications,” *CoRR*, vol. abs/1606.02373, 2016.
- [110] H. Zhu, R. Lu, C. Huang, L. Chen, and H. Li, “An efficient privacy-preserving location-based services query scheme in outsourced cloud,” *IEEE Trans. Vehicular Technology*, vol. 65, no. 9, pp. 7729–7739, 2016.
- [111] I. Bilogrevic, M. Jadliwala, K. Kalkan, J. Hubaux, and I. Aad, “Privacy in mobile computing for location-sharing-based services,” in *Privacy Enhancing Technologies - 11th International Symposium, PETS 2011, Waterloo, ON, Canada, July 27-29, 2011. Proceedings*, 2011, pp. 77–96.
- [112] K. P. N. Puttaswamy, S. Wang, T. Steinbauer, D. Agrawal, A. El Abbadi, C. Kruegel, and B. Y. Zhao, “Preserving location privacy in geosocial applications,” *IEEE Trans. Mob. Comput.*, vol. 13, no. 1, pp. 159–173, 2014.
- [113] G. F. Marias, C. Delakouridis, L. Kazatzopoulos, and P. Georgiadis, “Location privacy through secret sharing techniques,” in *2005 International Conference on a World of Wireless, Mobile and Multimedia Networks (WOWMOM 2005), 13-16 June 2005, Taormina, Italy, Proceedings*, 2005, pp. 614–620.
- [114] J. Shao, R. Lu, and X. Lin, “FINE: A fine-grained privacy-preserving location-based service framework for mobile devices,” in *2014 IEEE Conference on Computer Communications, INFOCOM 2014, Toronto, Canada, April 27 - May 2, 2014*, 2014, pp. 244–252.
- [115] R. Paulet, M. G. Kaosar, X. Yi, and E. Bertino, “Privacy-preserving and content-protecting location based queries,” *IEEE Trans. Knowl. Data Eng.*, vol. 26, no. 5, pp. 1200–1210, 2014.

-
- [116] H. Zhu, F. Liu, and H. Li, "Efficient and privacy-preserving polygons spatial query framework for location-based services," *IEEE Internet of Things Journal*, vol. 4, no. 2, pp. 536–545, 2017.
- [117] M. Naor and B. Pinkas, "Oblivious transfer with adaptive queries," in *Advances in Cryptology - CRYPTO '99, 19th Annual International Cryptology Conference, Santa Barbara, California, USA, August 15-19, 1999, Proceedings, 1999*, pp. 573–590.
- [118] "3GPP system architecture evolution (SAE); security architecture (Release 11)," *3GPP TS 33.401, tech. rep.*, Sep. 2011.
- [119] "Technical specification group radio access network; mobile relay for evolved universal terrestrial radio access (e-utra)," *3GPP TR 36.836, tech. rep.*, Nov. 2012.
- [120] L. Chen, Y. Huang, F. Xie, Y. Gao, L. Chu, H. He, Y. Li, F. Liang, and Y. Yuan, "Mobile relay in lte-advanced systems," *IEEE Communications Magazine*, vol. 51, no. 11, pp. 144–151, 2013.
- [121] J. Kokkonen, J. Ylitalo, P. Luoto, S. Scott, J. Leinonen, and M. Latva-aho, "Performance evaluation of vehicular LTE mobile relay nodes," in *24th IEEE Annual International Symposium on Personal, Indoor, and Mobile Radio Communications, PIMRC 2013, London, United Kingdom, September 8-11, 2013*, 2013, pp. 1972–1976.
- [122] O. Altrad and S. Muhaidat, "Intra-frequency handover algorithm design in LTE networks using doppler frequency estimation," in *Workshops Proceedings of the Global Communications Conference, GLOBECOM 2012, 3-7 December 2012, Anaheim, California, USA, 2012*, pp. 1172–1177.
- [123] Y. Sui, J. Vihriälä, A. Papadogiannis, M. Sternad, W. Yang, and T. Svensson, "Moving cells: a promising solution to boost performance for vehicular users," *IEEE Communications Magazine*, vol. 51, no. 6, 2013.

- [124] M. Pan, T. Lin, and W. Chen, "An enhanced handover scheme for mobile relays in LTE-A high-speed rail networks," *IEEE T. Vehicular Technology*, vol. 64, no. 2, pp. 743–756, 2015.
- [125] W. Lee and D. Cho, "Enhanced group handover scheme in multiaccess networks," *IEEE Trans. Vehicular Technology*, vol. 60, no. 5, pp. 2389–2395, 2011.
- [126] W. Lee and D. H. Cho, "Group handover scheme using adjusted delay for multi-access networks," in *Proceedings of IEEE International Conference on Communications, ICC 2010, Cape Town, South Africa, 23-27 May 2010*, 2010, pp. 1–5.
- [127] S. Chae, T. Nguyen, and Y. M. Jang, "A novel handover scheme in moving vehicular femtocell networks," in *2013 Fifth International Conference on Ubiquitous and Future Networks (ICUFN), Da Nang, 2013*, 2013, pp. 144–148.
- [128] R. Balakrishnan, X. Yang, M. Venkatachalam, and I. F. Akyildiz, "Mobile relay and group mobility for 4g wimax networks," in *2011 IEEE Wireless Communications and Networking Conference, WCNC 2011, Proceedings, Cancun, Mexico, 28-31 March, 2011*, 2011, pp. 1224–1229.
- [129] J. Zhao, T. Q. S. Quek, and Z. Lei, "Coordinated multipoint transmission with limited backhaul data transfer," *IEEE Trans. Wireless Communications*, vol. 12, no. 6, pp. 2762–2775, 2013.
- [130] A. Tölli, H. Pennanen, and P. Komulainen, "Decentralized minimum power multi-cell beamforming with limited backhaul signaling," *IEEE Trans. Wireless Communications*, vol. 10, no. 2, pp. 570–580, 2011.
- [131] P. Marsch and G. Fettweis, "Uplink comp under a constrained backhaul and imperfect channel knowledge," *IEEE Trans. Wireless Communications*, vol. 10, no. 6, pp. 1730–1742, 2011.

-
- [132] Y. Xia, X. Fang, W. Luo, M. Liu, S. Li, and Y. Zhao, "Coordinated of multi-point and bi-casting joint soft handover scheme for high-speed rail," *IET Communications*, vol. 8, no. 14, pp. 2509–2515, 2014.
- [133] C.-C. Lin, K. Sandrasegaran, and Z. Xu, "Performance testing of comp handover algorithms in lte-advanced," in *Workshop on Advances in Real-time Information Networks*, 2013.
- [134] S. Abdelhamid, H. S. Hassanein, and G. Takahara, "Vehicle as a mobile sensor," in *The 9th International Conference on Future Networks and Communications*, 2014, pp. 286–295.
- [135] U. Lee, E. Magistretti, M. Gerla, P. Bellavista, and A. Corradi, "Dissemination and harvesting of urban data using vehicular sensing platforms," *IEEE Trans. Vehicular Technology*, vol. 58, no. 2, pp. 882–901, 2009.
- [136] S. Devarakonda, P. Sevusu, H. Liu, R. Liu, L. Iftode, and B. Nath, "Real-time air quality monitoring through mobile sensing in metropolitan areas," in *Proceedings of the 2nd ACM SIGKDD International Workshop on Urban Computing, UrbComp@KDD 2013, Chicago, Illinois, USA, August 11, 2013*, 2013, pp. 15:1–15:8.
- [137] J. Eriksson, L. Girod, B. Hull, R. Newton, S. Madden, and H. Balakrishnan, "The pothole patrol: using a mobile sensor network for road surface monitoring," in *Proceedings of the 6th International Conference on Mobile Systems, Applications, and Services (MobiSys 2008), Breckenridge, CO, USA, June 17-20, 2008*, 2008, pp. 29–39.
- [138] G. Yang, S. He, Z. Shi, and J. Chen, "Promoting cooperation by the social incentive mechanism in mobile crowdsensing," *IEEE Communications Magazine*, vol. 55, no. 3, pp. 86–92, 2017.
- [139] R. Lu, X. Lin, T. H. Luan, X. Liang, and X. S. Shen, "Pseudonym changing at social spots: An effective strategy for location privacy in vanets," *IEEE Trans. Vehicular Technology*, vol. 61, no. 1, pp. 86–96, 2012.

- [140] M. Wen, R. Lu, X. Liang, J. Lei, and X. S. Shen, *Querying over Encrypted Data in Smart Grids*, ser. Springer Briefs in Computer Science. Springer, 2014.
- [141] U. Lee and M. Gerla, “A survey of urban vehicular sensing platforms,” *Computer Networks*, vol. 54, no. 4, pp. 527–544, 2010.
- [142] Y. Elmehdwi, B. K. Samanthula, and W. Jiang, “Secure k-nearest neighbor query over encrypted data in outsourced environments,” *CoRR*, vol. abs/1307.4824, 2013.
- [143] P. Paillier, “Public-key cryptosystems based on composite degree residuosity classes,” in *Advances in Cryptology - EUROCRYPT '99, International Conference on the Theory and Application of Cryptographic Techniques, Prague, Czech Republic, May 2-6, 1999, Proceeding*, 1999, pp. 223–238.
- [144] “<https://environment.google/projects/airview/>.”
- [145] “802.11 p-2010-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,” *URL <http://standards.ieee.org/findstds/standard/802.11p-2010.html>*, 2010.
- [146] M. S. Al-kahtani, “Survey on security attacks in vehicular ad hoc networks (vanets),” in *6th International Conference on Signal Processing and Communication Systems, ICSPCS 2012, Gold Coast, Australia, December 12-14, 2012*, 2012, pp. 1–9.
- [147] R. Schlegel, C. Chow, Q. Huang, and D. S. Wong, “User-defined privacy grid system for continuous location-based services,” *IEEE Trans. Mob. Comput.*, vol. 14, no. 10, pp. 2158–2172, 2015.
- [148] “Beijing fangshan industrial park,” *URL <https://www.researchandmarkets.com/reports/1946272/beijing-fangshan-industrial-park.pdf>*, 2011.
- [149] K. Liu, “Paillier’s homomorphic cryptosystem (java implementation),” *<https://www.csee.umbc.edu/~kunliu1/research/Paillier.html>*.

-
- [150] M. Gerla, E. Lee, G. Pau, and U. Lee, "Internet of vehicles: From intelligent grid to autonomous cars and vehicular clouds," in *IEEE World Forum on Internet of Things, WF-IoT 2014, Seoul, South Korea, March 6-8, 2014*, 2014, pp. 241–246.
 - [151] R. N. Rajaram, E. Ohn-Bar, and M. M. Trivedi, "Refinenet: Refining object detectors for autonomous driving," *IEEE Trans. Intelligent Vehicles*, vol. 1, no. 4, pp. 358–368, 2016.
 - [152] S. E. Li, S. Xu, K. Deng, and Q. Zhang, "Cruising control of hybridized powertrain for minimized fuel consumption," in *Automotive Air Conditioning*, 2016, pp. 267–289.
 - [153] K. M. Alam, M. K. Saini, and A. El-Saddik, "Toward social internet of vehicles: Concept, architecture, and applications," *IEEE Access*, vol. 3, pp. 343–357, 2015.
 - [154] C. Xu, R. Lu, H. Wang, L. Zhu, and C. Huang, "PAVS: A new privacy-preserving data aggregation scheme for vehicle sensing systems," *Sensors*, vol. 17, no. 3, p. 500, 2017.
 - [155] Sensors, "Url <http://jenniskens.livedsl.nl/technical/tips/files/bosch>
 - [156] "Anti-lock braking system," URL https://en.wikipedia.org/wiki/Anti-lock_braking_system.
 - [157] J. Yin, T. A. ElBatt, G. Yeung, B. Ryu, S. Habermas, H. Krishnan, and T. Talty, "Performance evaluation of safety applications over DSRC vehicular ad hoc networks," in *Proceedings of the First International Workshop on Vehicular Ad Hoc Networks, 2004, Philadelphia, PA, USA, October 1, 2004*, 2004, pp. 1–9.
 - [158] E. Ayday, J. L. Raisaro, P. J. McLaren, J. Fellay, and J. Hubaux, "Privacy-preserving computation of disease risk by using genomic, clinical, and environmental data," in *2013 USENIX Workshop on Health Information Technologies, HealthTech '13, Washington, D.C., August 12, 2013*, 2013.
 - [159] C. Ding, D. Pei, and A. Salomaa, *Chinese remainder theorem: applications in computing, coding, cryptography*. World Scientific, 1996.
 - [160] T. L. Saaty, *Elements of queueing theory: with applications*, 1961.

- [161] E. Bresson, D. Catalano, and D. Pointcheval, “A simple public-key cryptosystem with a double trapdoor decryption mechanism and its applications,” in *Advances in Cryptology - ASIACRYPT 2003, 9th International Conference on the Theory and Application of Cryptology and Information Security, Taipei, Taiwan, November 30 - December 4, 2003, Proceedings*, 2003, pp. 37–54.
- [162] W. Chen, R. K. Guha, T. J. Kwon, J. Lee, and Y. Hsu, “A survey and challenges in routing and data dissemination in vehicular *ad hoc* networks,” *Wireless Communications and Mobile Computing*, vol. 11, no. 7, pp. 787–795, 2011.
- [163] M. Chiang and T. Zhang, “Fog and iot: An overview of research opportunities,” *IEEE Internet of Things Journal*, vol. 3, no. 6, pp. 854–864, 2016.
- [164] X. Hou, Y. Li, M. Chen, D. Wu, D. Jin, and S. Chen, “Vehicular fog computing: A viewpoint of vehicles as the infrastructures,” *IEEE Trans. Vehicular Technology*, vol. 65, no. 6, pp. 3860–3873, 2016.
- [165] F. Al-Turjman, “Cognitive caching for the future sensors in fog networking,” *Pervasive and Mobile Computing*, vol. 42, pp. 317–334, 2017.
- [166] K. Liu and V. C. Lee, “Rsu-based real-time data access in dynamic vehicular networks,” in *Intelligent Transportation Systems (ITSC), 2010 13th International IEEE Conference on*, 2010, pp. 1051–1056.
- [167] F. Ye, S. Roy, and H. Wang, “Efficient data dissemination in vehicular ad hoc networks,” *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 4, pp. 769–779, 2012.
- [168] K. Rabieh, M. M. E. A. Mahmoud, and M. F. Younis, “Privacy-preserving route reporting schemes for traffic management systems,” *IEEE Trans. Vehicular Technology*, vol. 66, no. 3, pp. 2703–2713, 2017.
- [169] R. Li, H. Asaeda, J. Li, and X. Fu, “A verifiable and flexible data sharing mechanism for information-centric iot,” in *IEEE International Conference on Communications, ICC 2017, Paris, France, May 21-25, 2017*, 2017, pp. 1–7.

- [170] J. Liu, J. Han, L. Wu, R. Sun, and X. Du, "VDAS: verifiable data aggregation scheme for internet of things," in *IEEE International Conference on Communications, ICC 2017, Paris, France, May 21-25, 2017*, 2017, pp. 1–6.
- [171] J. Bohli and A. Pashalidis, "Relations among privacy notions," *ACM Trans. Inf. Syst. Secur.*, vol. 14, no. 1, pp. 4:1–4:24, 2011.
- [172] O. Cetinkaya and A. Doganaksoy, "Pseudo-voter identity (PVID) scheme for e-voting protocols," in *Availability, Reliability and Security, 2007. ARES 2007. The Second International Conference on*, 2007, pp. 1190–1196.
- [173] S. Cui, P. Duan, C. W. Chan, and X. Cheng, "An efficient identity-based signature scheme and its applications," *I. J. Network Security*, vol. 5, no. 1, pp. 89–98, 2007.
- [174] A. D. Caro and V. Iovino, "jpbcc: Java pairing based cryptography," in *Proceedings of the 16th IEEE Symposium on Computers and Communications, ISCC 2011, Kerkyra, Corfu, Greece, June 28 - July 1, 2011*, 2011, pp. 850–855.
- [175] A. Le, J. Loo, A. Lasebae, M. Aiash, and Y. Luo, "6lowpan: a study on qos security threats and countermeasures using intrusion detection system approach," *Int. J. Communication Systems*, vol. 25, no. 9, pp. 1189–1212, 2012.
- [176] Z. Zhou, M. Dong, K. Ota, G. Wang, and L. T. Yang, "Energy-efficient resource allocation for D2D communications underlaying cloud-ran-based LTE-A networks," *IEEE Internet of Things Journal*, vol. 3, no. 3, pp. 428–438, 2016.
- [177] J. Wu, M. Dong, K. Ota, L. Liang, and Z. Zhou, "Securing distributed storage for social internet of things using regenerating code and blom key agreement," *Peer-to-Peer Networking and Applications*, vol. 8, no. 6, pp. 1133–1142, 2015.
- [178] L. Costantino, N. Buonaccorsi, C. Cicconetti, and R. Mambrini, "Performance analysis of an LTE gateway for the iot," in *2012 IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks, WoWMoM 2012, San Francisco, CA, USA, June 25-28, 2012*, 2012, pp. 1–6.

- [179] R. W. H. Jr., M. L. Honig, S. Nagata, S. Parkvall, and A. C. K. Soong, "Lte-advanced pro: part 1 [guest editorial]," *IEEE Communications Magazine*, vol. 54, no. 5, pp. 74–75, 2016.
- [180] Z. Zhou, M. Dong, K. Ota, and Z. Chang, "Energy-efficient context-aware matching for resource allocation in ultra-dense small cells," *IEEE Access*, vol. 3, pp. 1849–1860, 2015.
- [181] M. Peng, Y. Li, Z. Zhao, and C. Wang, "System architecture and key technologies for 5g heterogeneous cloud radio access networks," *IEEE Network*, vol. 29, no. 2, pp. 6–14, 2015.
- [182] L. Atzori, A. Iera, and G. Morabito, "The internet of things: A survey," *Computer Networks*, vol. 54, no. 15, pp. 2787–2805, 2010.
- [183] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of things (iot): A vision, architectural elements, and future directions," *Future Generation Comp. Syst.*, vol. 29, no. 7, pp. 1645–1660, 2013.
- [184] N. Lin, X. Huang, and X. Ma, "Analysis of the uplink capacity in the high-speed train wireless communication with full-duplex mobile relay," in *IEEE 83rd Vehicular Technology Conference, VTC Spring 2016, Nanjing, China, May 15-18, 2016*, 2016, pp. 1–5.
- [185] F. Y. Lin, C. Hsiao, K. Chu, and Y. Liu, "Minimum-cost qos-constrained deployment and routing policies for wireless relay networks," *J. Applied Mathematics*, vol. 2013, pp. 1–19, 2013.
- [186] C. Lai, H. Li, R. Lu, R. Jiang, and X. Shen, "SEGR: A secure and efficient group roaming scheme for machine to machine communications between 3gpp and wimax networks," in *IEEE International Conference on Communications, ICC 2014, Sydney, Australia, June 10-14, 2014*, 2014, pp. 1011–1016.

-
- [187] G. M. Kjøien, "Mutual entity authentication for LTE," in *Proceedings of the 7th International Wireless Communications and Mobile Computing Conference, IWCMC 2011, Istanbul, Turkey, 4-8 July, 2011*, 2011, pp. 689–694.
- [188] D. Forsberg, "LTE key management analysis with session keys context," *Computer Communications*, vol. 33, no. 16, pp. 1907–1915, 2010.
- [189] M. L. Das, "Two-factor user authentication in wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 8, no. 3, pp. 1086–1090, 2009.
- [190] M. Rezvani, A. Ignjatovic, E. Bertino, and S. Jha, "Secure data aggregation technique for wireless sensor networks in the presence of collusion attacks," *IEEE Trans. Dependable Sec. Comput.*, vol. 12, no. 1, pp. 98–110, 2015.
- [191] H. Lin, J. Shao, C. Zhang, and Y. Fang, "CAM: cloud-assisted privacy preserving mobile health monitoring," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 6, pp. 985–997, 2013.
- [192] N. Savage, "Proofs probable," *Commun. ACM*, vol. 56, no. 6, pp. 22–24, 2013.
- [193] C. Lai, H. Li, R. Lu, and X. S. Shen, "SE-AKA: A secure and efficient group authentication and key agreement protocol for LTE networks," *Computer Networks*, vol. 57, no. 17, pp. 3492–3510, 2013.
- [194] C. Yang, S. Han, X. Hou, and A. F. Molisch, "How do we design comp to achieve its promised potential?" *IEEE Wireless Commun.*, vol. 20, no. 1, pp. 67–74, 2013.
- [195] N. Lu, N. Cheng, N. Zhang, X. Shen, and J. W. Mark, "Connected vehicles: Solutions and challenges," *IEEE Internet of Things Journal*, vol. 1, no. 4, pp. 289–299, 2014.
- [196] Q. Wu, Y. Mu, W. Susilo, B. Qin, and J. Domingo-Ferrer, "Asymmetric group key agreement," in *Advances in Cryptology - EUROCRYPT 2009, 28th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Cologne, Germany, April 26-30, 2009. Proceedings*, 2009, pp. 153–170.

- [197] R. Lu, X. Lin, H. Zhu, P. Ho, and X. Shen, “ECPP: efficient conditional privacy preservation protocol for secure vehicular communications,” in *INFOCOM 2008. 27th IEEE International Conference on Computer Communications, Joint Conference of the IEEE Computer and Communications Societies, 13-18 April 2008, Phoenix, AZ, USA, 2008*, pp. 1229–1237.