Secure and Privacy-aware Traffic Information as a Service in VANET-based Clouds
Rasheed Hussain*, Zeinab Rezaeifar*, Yong-Hwan Lee**, and Heekuck Oh*

*Dept. of Computer Science and Engineering, Hanyang University, South Korea
**Dept. of Smart Mobile, Far East University, South Korea

Email: {rasheed, hko}@hanyang.ac.kr

Corresponding Author:
Professor, Heekuck Oh, PhD
Department of Computer Science and Engineering,
Hanyang University, ERICA Campus,
Sa 3-dong, Sangnok-gu,
Ansan, Gyeonggi, 426-791, South Korea.
Phone: 0082-31-400-5197
Fax: 0082-31-400-1874
Email: hko@hanyang.ac.kr
Abstract

The multitude of applications and services in Vehicular Ad hoc NETwork (VANET) and cloud computing technology that are dreamed of in the near future, are visible today. VANET applications are set to explode in the next couple of years as a result of the advancements in the wireless communication technologies and automobile industry. Nevertheless, it is speculated that future high-end vehicles will potentially under-utilize their on-board storage, computation, and communication resources. This phenomenon set the ground for the evolution of traditional VANET to a rather more applications-rich paradigm referred to as VANET-based clouds. In this paper, we aim at a framework of VANET-based clouds namely VANET using Clouds (VuC) and propose a novel secure and privacy-aware service referred to as Traffic Information as a Service (TIaaS) atop the cloud computing services stack. TIaaS provides vehicles (more precisely subscribers) with fine-grained traffic information from the cloud as a result of subscribers’ cooperation with the cloud in a privacy-preserving way. Legitimate VANET users share their frequent whereabouts information referred to as Mobility Vectors (MV) with the cloud infrastructure through gateways (static Road Side Units-RSUs and mobile vehicles with 3/4G internet). The gateways forward coarse-grain traffic information (MVs) from vehicles to the cloud whereas after processing, cloud modules construct and re-forward the fine-grained traffic information along with location-based warnings to the subscribers based on their physical locations and moving directions. The communication among vehicles, gateways, and the cloud infrastructure is carried out in a privacy-preserving way. More precisely vehicles share their MVs with cloud infrastructure anonymously. The MVs are hard to link to the sender, until and unless necessary, otherwise. Similarly every vehicle receives fine-grained traffic information in a privacy-preserving manner. The proposed TIaaS keeps the adversaries at bay from abusing users’ privacy and/or constructing profiles against targeted users. Moreover for location confidentiality and privacy, we also propose a novel location-based encryption technique that keeps the insider and outsider adversaries at bay from manipulating the contents of the message. Furthermore, the proposed TIaaS preserves conditional privacy and with the help of an efficient revocation mechanism, revocation authorities can revoke any node in case of a dispute. The proposed TIaaS also introduces the thin-client concept for vehicles where most of the time-consuming processing is offloaded to the cloud and the processing resources of the vehicles can be used elsewhere, for instance for the critical safety related applications. More precisely the cloud processes the big traffic data (BTD) and produces timely, decisive, and meaningful results.

Keywords: VANET clouds, security, privacy, traffic information dissemination, cloud computing
1. Introduction

Vehicular Ad Hoc NETwork (VANET) is widely regarded as a promising concept for enabling safe, reliable, and infotainment-rich driving experience. However governments, service providers, and automobile manufacturers are still reluctant to deploy VANET because it brings challenges such as cost, safety and security issues, need for infrastructure, and user preferences, to name a few. It is worth noting that in the past, automotive industry focused on the in-car technology to improve navigation and entertainment systems, but recently research community put their research emphasis on VANET because of the improved processing, computation, and computation capabilities of the high-end vehicles. These capabilities enabled vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communication. These communication paradigms are capable of providing consumers with safety and infotainment applications rather than providing them with only specific applications such as a queried destination (for instance a restaurant or a friend’s home in other city) [1].

Documented proof from US Department of Transportation (DoT) in 2008 state that worker productivity of worth US$ 75 billion was lost. Moreover in accordance with worker productivity, around 8.4 billion gallons of fuel is wasted [2]. Besides, it is a common understanding that most of the highway turbulence situations and congestions occur due to rush hours, but it has been pointed out that almost half of these congestions and road blockage occurred due to road incidents [3]. To counter these problems, infrastructure such as traffic cameras, road sensors, and other traffic management systems have been deployed in metropolitan cities on a large scale, however; from the cost, management, maintenance perspective, other robust, reliable, and cost-effective solutions are essential. In other words, the current advancements in technologies clearly mandate the need for intelligent transportation system (ITS) to enable safe, reliable and infotainment-rich driving experience. ITS is realized through VANET. World leading automobile manufacturers, government agencies, standardization agencies, academia, and consortia have already tapped their research resources to work out the problems in deploying VANET technology in the masses of vehicles (for instance Networks-on-Wheels, Car-2-Car Communication Consortium, the Vehicle Safety Communication Consortium, and Honda’s Advanced Safety Vehicle Program) [1].

On the other hand, the bandwidth allocated for the exclusive use of dedicated short range communication (DSRC) standard is in excess and exceeds the baseline requirements of VANET applications. Therefore in order to utilize the bandwidth in an efficient way, there are more possibilities rather than just focusing on simple VANET applications [4-5]. To be more precise, US Federal Communication Commission (FCC) has allocated a rich 75 MHz of spectrum (5.850~5.925 GHz band) for VANET standard, i.e.
DSRC or WAVE 802.11p\(^1\). As aforementioned, today’s high-end vehicles are capable of storage, communication and computation capabilities and most of the times these resources are wasted, for instance when the cars are parked. This phenomenon gives motivation to other technologies to merge with VANET in order to fully utilize vehicular resources. To this end, cloud computing is the ideal technology to utilize vehicular resources even if they are parked. In other words, it can be possible that the vehicle rents out its resources for some revenue. The reason for merging cloud computing with VANET is straightforward, it fits to the resource utilization scenario, and the cloud infrastructure is already in place. Moreover the virtualization concept of cloud computing has made it a buzz word and with the virtue of cloud computing, consumers can have unlimited resources if they can pay for it\(^6\).

Olariu et al. for the first time tossed the term VANET clouds that is the combination of the signature VANET and cloud computing\(^7\). In VANET clouds, vehicular nodes rent out their resources to the cloud and/or use resources form cloud as well. The driving force of merging VANET with cloud computing is the motive of cloud computing which almost guarantees ‘Anything as a Service’. Moving a step ahead, Hussain et al.\(^8,41\) put forth VANET clouds taxonomy and proposed potential architectural frameworks for VANET-based clouds namely Vehicular Clouds (VC), Vehicles using Clouds (VuC), and Hybrid Vehicular Clouds (HVC).

In this paper we aim at VuC framework and propose another secure and privacy-aware service layer atop the cloud computing stack. We name the new service layer as Traffic Information as a Service (TIaaS). This layer provides vehicular nodes on the road with fine-grained traffic information service as a result of cooperation between vehicles and the cloud infrastructure through gateways. Interested vehicles can subscribe to TIaaS and receive fine-grained traffic information and possibly other services based on their current and near-future locations. It is worth noting that this fine-grained traffic information is leveraged to construct long range extended traffic view, which otherwise would require multi-hop communication. Road-Side Units (RSUs) and high-end vehicles with 3/4G internet support act as Gateway Terminal (GT) between the vehicles on the road and the cloud infrastructure. The role of GT is to forward the coarse-grained cooperation from vehicles to the cloud and disseminate fine-grained traffic information from cloud to vehicles based on their locations. Legitimate VANET users share their coarse-grained traffic information in the form of beacon messages, hereafter referred to as Mobility Vectors (MVs), with each other and with cloud infrastructure through GTs. According to DSRC standard, each vehicle broadcasts MVs with a frequency that is in the order of milliseconds. It is worth noting that in order to preserve users’ conditional privacy, we leverage anonymous MVs. More precisely, MVs do not contain any identity

---

\(^1\) Institute of Electrical and Electronics Engineers (IEEE), “Trial-use standard for wireless access in vehicular environments- security services for applications and management messages,” IEEE Std 1609.2 (rev D15), 2012.
information through which profile could be constructed against targeted users and abuse their privacy. However, in case of any misbehavior or a dispute, users are subject to revocation by the revocation authorities (RAs) with the help of a trapdoor in the MVs. This way the privacy of the users is preserved conditionally. Similarly the fine-grained traffic information is also disseminated among the subscribers in a privacy-preserving manner.

MVVs contain frequent mobility information such as current location, speed, acceleration, direction, heading, and vehicle control information such as brake status and steering wheel angle etc. Normally MVVs are broadcasted to one hop neighbors and every vehicle uses these MVVs to construct its local short range traffic view that is limited to the one-hop radius (about 300 m according to DSRC). Fast and reliable information spraying is an essential precondition for a successful implementation of VANET-based cloud applications and services. Assuming sufficient market penetration, we argue that TIaaS offers the possibility of a quasi-centralized and robust traffic information system, where coarse-grained traffic data is collected by cloud infrastructure, evaluated, refined, and re-distributed the fine-grained traffic information among the cars based on their physical locations. However, the primary requirement of such service is the privacy guarantee for the users. High frequency of the MVs and the fine-grained traffic messages pose challenges for the privacy of the users. The high frequency of MVs enables adversaries to abuse privacy and/or construct movement profiles against targeted users. Therefore the users will not use such service if their privacy is at stake. As a result, it is essential to design the service in such a way that the privacy of the users is conditionally preserved. To fill the gap, we leverage anonymous MVs. Users broadcast identityless MVs with no identity information that could lead the adversary to link the message to the sender. Therefore our proposed MVs make it difficult for insiders and outsiders to construct movement profile against targeted users. Similarly fine-grained traffic information is forwarded to the subscribers in a privacy-preserving manner as well. Our propose TIaaS is secure against both insiders and outsiders, and preserves conditional privacy where in case of any dispute the culprits are subject to revocation by the RAs. Moreover our proposed scheme also provides location confidentiality and non-frameability through our novel geolock-based encryption. The significance of cooperation between cloud infrastructure and VANET users for service exchange is not limited to traffic information dissemination, rather it can provide VANET with a number of applications, for instance, vehicles as witnesses on the road, criminal investigation, and route tracing to name a few. In the aforementioned scenarios, MVs are stored in the cloud and authorities (law enforcement agencies and insurance companies) use the stored information as forensic evidences for investigation. The contribution of this paper can be summarized as follows:
a. **Traffic Information Dissemination:** We propose a novel and efficient cloud-based traffic information dissemination service (TIaaS) to the vehicles on the road where the processing is outsourced to the cloud infrastructure as a result of cooperation between vehicles and the cloud infrastructure.

b. **Conditional privacy and anonymity:** We preserve users’ privacy and anonymity in a conditional manner throughout the service provision starting from the cooperation from the vehicle all the way to the service delivery. Our proposed TIaaS also includes an efficient revocation mechanism in case of any dispute.

c. **Location confidentiality and privacy:** To counter both insiders and outsider attackers, we propose a novel location-based encryption mechanism to keep the outsiders attackers from intruding into the system and minimizing the malicious behavior of the insiders who may try to manipulate the contents.

d. **Efficient MV framework:** We tailor the existing beacon framework and use the coarse-grain information in beacon messages as mobility vectors in a secure and privacy preserving manner. It is also to be noted that MVs are the essence of our proposed TIaaS.

The structure of the rest of the paper is organized as follows. Section 2 summarizes the related work regarding VANET, cloud computing, and vehicular clouds. We provide the readers with a comprehensive background of VANET, cloud computing, and vehicular clouds in section 3. Section 4 briefly outlines our proposed TIaaS scheme. In section 5, we quantitatively evaluate and analyze our proposed scheme followed by concluding remarks and future directions in section 6.

### 2. Related Work and Research Objectives

VANET-based cloud is new, infant, and exciting topic and it is yet to be explored. Conventional VANET and cloud computing are given much attention by research community from services, design and architecture, security, and privacy standpoints [9,10,13,14]. Nevertheless, despite the surge in both VANET and cloud computing researches, the complete deployment of these emerging technologies, their success, and adoption in end-users, consumers, and governments will strictly depend on viable security solutions and consumer satisfaction [11,12]. In this section we outline previous research carried out regarding VANET, cloud computing, and VANET-based clouds. Moreover in the last subsection, we also outline the details of our contributions from the novelty, originality, and significance standpoint.
2.1. VANET

This sub-section puts light on research carried out on VANET architecture, design, security, and privacy. Kargl et al. [13] discussed implementation, performance, and research challenges in VANET in detail. Similarly Papadimitriou et al. [14] outlined design and architectural issues in VANET. Before bringing VANET to the masses, researchers from academia and industries, after great brainstorming proposed remedies to security and privacy issues in VANET. VANET inherited most of the security and privacy issues from general networks [31]. Dotzer et al. discussed the privacy requirements in VANET in a great detail [15] and privacy enhancing solutions and data-centric misbehavior solutions can be found in [9,10,16,36]. Moreover, to date, multiple pseudonymous approach is regarded as one of the best solutions for conditional privacy preservation in VANET [38,39,40]. However despite considerable amount of research, the advancements in VANET deployment have been impeded because of lack of proper infrastructure in place, consumer satisfaction, and economical reasons. Nonetheless, governments, academia, and automobile industries are testing their waters for the initial deployment of this tempting technology.

Optimistically the services offered by VANET are not limited to safety warnings, non-safety applications like traffic congestion and routing information, but also include value-added services like high-speed tolling, mobile infotainment, internet-on-the-road, movies-on-demand, and IPTV [17]. Nonetheless ephemeral nature of VANET is still a daunting challenge for researchers and still needs to be addressed.

2.2. Cloud Computing

Currently many commercial cloud computing services are provided to consumers either for free or on the basis of ‘pay as you go’ depending upon the type of service. But the real question is that whether cloud computing will hold the consumer market or it will become a hyped subject that will be thrown to the corner in the future [18]. Current market-players include Google, Amazon, and Microsoft. The security and privacy problems are alike up to some extent in both VANET and cloud computing, however; merging of these two technologies pose some additional challenges as well, for instance high mobility in VANET, fast authentication, and so forth. Storage security is another hot issue in the cloud computing these days. Bessani et al. [19] proposed a scheme to solve the storage security problem in cloud computing. They used encryption, encoding techniques, and replication to form cloud-of-clouds. Cachin et al. [20] conducted a comprehensive survey about data integrity and consistency in clouds. They

---

outlined the recent research trends in cryptography and distributed computing that can be used to solve the aforementioned security problems in cloud computing.

2.3. VANET Clouds

Olariu et al. for the first time proposed VANET clouds [7, 21]. Furthermore they came up with a new notion of VANET called Autonomous Vehicular Clouds (AVC). Bernstein et al. [22] proposed a Platform as a Service (PaaS) model for mobile vehicular domain with possible potential applications. Moreover Hussain et al. proposed vehicles witness as a service where vehicles moving on the road serve as witness to certain designated incidents and save the forensics in the cloud which can be used by the law enforcement agencies and/or insurance agencies during the trial [42]. Though Olariu et al. put light on some applications of VANET-based clouds, for instance data center at car parking or shopping mall, dynamic scheduling of traffic lights, and so forth, but they did not define any potential variations of such emerging technology and did not outline architectural framework. Their work is remarkable starting point of this field but it needs deeper insight. Yan et al. outlined the security challenges of the VANET clouds and proposed security mechanism [23]. However we argue that the long-chased privacy issues have not been addressed by Yan et al. in their scheme. Moreover their proposed security scheme covers only traditional VANET whereas VANET-based clouds inherit security and privacy issues from both VANET and CC. Additionally their scheme allows frameability, abuse of location and user privacy. Recently Hussain et al. [8, 41] proposed three architectural frameworks for cloud-based VANETs namely VC, VuC, and HVC. Furthermore a comprehensive survey about vehicular clouds, is carried out by Wahiduzzaman et al. [26].

Qin et al. [24] proposed a cloud-based routing scheme in VANET named VehiCloud which provides routing services for VANET. Vehicles share their current and future location information in the form of waypoints with cloud infrastructure and cloud provides them with optimal routing information. The limitation of their proposed scheme is that, the future location of the vehicle depends upon the current velocity and the driving behavior of the driver. Moreover their scheme does not take security and privacy concerns into account.

2.4. Objectives and Contributions

After scrutinizing previously proposed schemes in VANET-based clouds, we extend VuC architecture proposed by Hussain et al. and propose a novel cloud-based, secure, and privacy-aware traffic information dissemination service (TIaaS) in VANET-based clouds. This work is the extended version of our previous work in [43]. In the poster version of this work, we put forth the baseline idea of the TIaaS
whereas in this paper, we encompass the full picture of the TIaaS service and rigorously analyze its security, privacy, and other parameters. Moreover we outline the current state of the art regarding VANET-based cloud services and the working networking model. Our proposed scheme is different from the traditional traffic information dissemination schemes in certain aspects, for instance we introduce a thin client concept for the vehicles where vehicles offshore their data to the cloud for processing. Moreover this type of offshoring is also considered to be ‘cooperation’ from the vehicles to the cloud and in return cloud provides the vehicles with fine-grained traffic information based on their interests and/or current/near-future locations. It is also to be noted that the fine-grained information received from the cloud covers more than single hop thereby enabling vehicles to extend their traffic view from short range to long range depending upon their physical locations. Therefore such cloud based approach avoids complex and computationally expensive multi-hop communication in VANET. Another dimension of our work is the conditional privacy preservation in a highly ephemeral and mobile environment. To date, outsider attackers have been the target of the research community to keep them at the bay from intruding into the VANET system, however; insider attackers can be a serious risk as well because they have access to the benign credentials. Such phenomenon of insiders is more challenging than outsiders. Therefore our proposed scheme also leverages a novel location-based encryption mechanism to 1) keep the outsiders from attacking the system from outside and 2) minimizing the possible damage by the insiders through content manipulation.

3. Overview of VANET, Cloud Computing, and VANET-based Clouds

In this section we outline the baseline overview of the traditional VANET, cloud computing, and VANET-based clouds.

3.1. Traditional Vehicular Networks (VANET)

VANET, a specialized breed of Mobile Ad Hoc NETwork (MANET), employs vehicles as mobile nodes moving on the road. The key components of the VANET system are the management entities and the users. These entities are further divided into sub-classes where management entities consist of registration, certification, and revocation authorities, whereas users are the vehicles equipped with VANET hardware and software. From bird’s view, there are two major communication paradigms in VANET mandated by the DSRC standard, vehicle-to-vehicle communication paradigm (V2 or Zero Infrastructure) and vehicle-to-infrastructure (V2I) communication paradigm. Moreover in compliance with DSRC, SAE J2735, two message classes has been defined for the exclusive use of VANET\textsuperscript{3}. These two message sets are cooperative awareness message (CAM) and other safety related messages. Both of the aforementioned

\textsuperscript{3} SAE J2735, Dedicated Short Range Communication (DSRC) Message Set Dictionary, 2009.
message classes have their own significance in the VANET application, however; CAMs are of paramount importance because the data contained in CAMs is used by almost every VANET application, directly or indirectly. Vehicles intelligently decide on the basis of the statistics drawn from these data. CAM contains vehicle’s frequent mobility statistics (in the order of milliseconds). These statistics include current position, velocity, and so forth. Due to the high frequency, this data can give a clear picture of the target node and its mobility statistics.

There are enormous application offered by the VANET technology ranging from safety related applications (for example an accident on the freeway, black ice on the pavement, and traffic jam in rush hours) to infotainment on the move. For instance CAMs are used to construct short range local traffic view and long range extended traffic views. Another such application is Cooperative Adaptive Cruise Control (CACC) which controls different maneuvers during driving. CACC lets the vehicles know about the traffic dynamics of the surroundings (in the front), and decide to maneuver the vehicle through V2V communication [25]. It is worth noting that many industries and governments consortia have put their research resources together to develop value-added applications and services for the consumers. These organizations include Crash Avoidance Metrics Partnership (CAMP) which is the collaboration of vehicle companies like Mercedes Benz, KIA, Hyundai, Ford, GM, Honda, Nissan, Toyota, and Volkswagen, and the famous application is Automatic Collision Notification System (CNS) similar to GM’s Onstar system with Advanced Automatic Collision Notification (CAN). Moreover, US department of transportation (DoT) has already mandated the legislation of V2V communication. This legislation will motivate investors and other market players to invest in the deployment of this technology.

3.2. Cloud Computing Services

Cloud computing is more of a business model than a technology. The possibilities of virtually limitless resources have successfully convinced investors, young companies and entrepreneurs to migrate from traditional technologies to cloud computing because it almost has zero infrastructure and maintenance cost. Cloud computing has made it possible to deliver services to the users in a virtual manner where the service users have no idea where the hardware and the servers reside. Therefore as aforementioned, many businesses are migrating to the cloud computing, however; it is worth noting that the complete migration towards cloud computing will take time. Cloud computing offers a generic service model referred to as ‘anything as a service’ where end users only use the service from the cloud without knowing the origin of

---


that service and so forth. Moreover, cloud computing is also known as ‘utility computing’ which is based on ‘pay-as-you-go’ billing mechanism [6]. This scenario can be compared to our household utilities where we pay for the amount of the utility (electricity and gas) that we use. Cloud offers three major service paradigms namely platform as a service (PaaS) where the developers can develop their application without caring for the toolkits and the operating systems, infrastructure as a service (IaaS) where the users can have unlimited hardware resources remotely and software as a service (SaaS) where users can have access to the softwares remotely without worrying about the licensing and so forth. Besides, cloud offers offshore data storage and processing as well. Therefore world leading companies have already established cloud computing infrastructures and are providing services to the users (both free and paid services). The market leaders in the cloud industry include Amazon S3, Google Drive, and Microsoft SkyDrive, to name a few.

Despite exciting opportunities provided by the cloud computing, there are still grey areas that have not been addressed in the cloud, such as user privacy, data privacy, storage security, and control over data, to name a few. From the users’ standpoint, the main problem with using cloud infrastructure is the control over users’ data. Once the users offshore their data either for processing or for simple storage to the cloud, they logically lose control over their own data. Therefore necessary measures are essential to deal with such problems in the cloud computing paradigm. Data integrity and audit are also challenging once the data is stored in the remote service provider’s servers. Users are more concerned about the data manipulation by the service providers without users’ consent. To date, some work has been done to address these issues, for instance Chow et al. [27] believe that integrity of cloud infrastructure is ensured through the use of traditional trusted computing.

3.3. Communication Structure in VANET-based Clouds

In the VANET-based clouds scenario, there are four communication paradigms namely Gateway to Gateway (G2G), V2V, V2I, and in-car communication. The communication structure is shown in Fig 1. It can be seen that the in-car communication consists of different in-car communicating nodes, and they collaborate with each other. Moreover these sensors give data to the OBU for message construction. The entities involved in this type of communication include radar, Global Positioning System (GPS), sensors, and actuators. These entities create a networked environment in the car and share their data with other modules. Moreover, a small sensor network could be established inside the car. Apart from these sensors, there are a number of other sensors throughout the car that share their data with the processing modules. The second type of communication is inter-car level communication that consists of both V2V and V2I. In other words, OBUs communicate the OBUs of other vehicles as well as with the gateways (RSUs) through DSRC standard. This communication takes place for the cooperative awareness, service and
cooperation exchange, safety alerts and so forth. The data from the in-car network is routed to the OBUs and then OBU takes control of that data. At this level an In-Car cloud can be formed. The mid-level communication (V2V and V2I) serves as bridge between the conventional VANET and the VANET-based cloud because procedures to form cloud between the vehicles, are executed at this level. Other communication paradigm consists of gateways where the communication takes places at gateways level. More precisely, the gateways communicate among themselves and with the service providers. It is worth noting that the term gateway is generic and a proper entity either RSU or a police patrol van could serve as a gateway, however; the nomination depends upon the underlying framework of VANET-based clouds. The top most communication paradigm is responsible for the resource provision from other clouds. Moreover, these gateways are also responsible for communicating with the management entities for keys updates, revocation, credentials, and so forth.

![Diagram of Communication paradigm of VANET clouds](image)

3.4. Conventional Clouds Vs VANET-based Clouds

There are similarities between conventional clouds and the VANET clouds at some levels, however; due to the difference in their network models, we need to address their differences as well. Fig. 2 shows the conventional clouds and VANET-based clouds together [41]. It can be seen from the figure that starting from the infrastructure all the way to the clients in the conventional clouds, there is an entity in the VANET-based clouds that mimics the functionality carried out by its counterpart. For instance, at the infrastructure level, conventional clouds have their own infrastructure, whereas in case of VANET-based clouds, they leverage OBUs and RSUs. Same argument holds for the platform, application, and clients as well where in VANET-based clouds we have WAVE standard, VANET applications, and the Vehicular nodes, respectively. It is also worth noting that Fig. 2 gives a very abstract representation of the two architectures and the particulars of application-based VANET-based clouds can be different based on the underlying framework.
3.5. VANET Clouds Taxonomy

VANET clouds taxonomy proposed by Hussain et al. [8,41] is shown in Fig. 3. Our proposed TlaaS is based on VuC framework where VANET users use cloud services on the move. VANET entities (more precisely OBU) collaborate with cloud modules and provide cloud with meaningful data through gateways. This data can be used in decision making and other applications. The virtualization layer is provided by the GTs. VANET-based clouds can be divided into three frameworks namely VC, VuC, and HVC. In VC, vehicles tap their resources to form a disposable cloud and high performance distributed computing environment to carry out certain computational operations. Second kind of framework is VuC where vehicles use other clouds (public or private) for additional services. The third framework is HVC where vehicles form their own cloud and use external cloud services at the same time. Since vehicles use gateways (GTs) to communicate with authorities and other service providers, we assume that high speed wired backbone channel is present between GTs and the cloud infrastructure. Without loss of generality, VuC can be used to provide VANET with CAA, real-time traffic information, infotainment, and forensic services. However in this paper we only emphasize on traffic information dissemination.
4. Proposed Secure and Privacy-aware Traffic Information as a Service (TIaaS)

4.1. Design Rationale

VANET applications can be mainly divided into two classes, safety applications and non-safety applications. Most of the applications from both classes depend upon the constant mobility data from the neighborhood. In other words, these applications rely on the information from beacon messages. According to DSRC standard, every vehicle generates and broadcast beacon messages in the order of milliseconds, that is why a huge amount of data is generated in the neighborhood. Such data is termed as big traffic data (BTD). Keeping in mind the amount of data to be processed by single vehicular node, and the resources needed, it can be implied that a reasonable amount of resources will be spent on the data processing. On the other hand, the resources of the cloud can be easily utilized for the delay-tolerant applications in VANET where vehicles outsource their coarse-grained data to the cloud and receive fine-grained information. This will serve multiple purposes, for instance, the resources of the vehicles can be used elsewhere for critical safety-related applications, and so forth. From another perspective, in order to construct a long range extended traffic view, the vehicular node must receive beacons from an area equal to the length of the view which can be as large as 2 km. Beacon rebroadcasting on this scale can easily cause broadcast storm problem. Whereas the same long range extended traffic view can be received as a single message from the cloud thereby saving the bandwidth and avoiding congestion on the channel. Nonetheless, the computation and communication delay must be addressed in the service level agreement (SLA) with the cloud.
4.2. Assumptions

Our proposed scheme is based on the following assumptions.

1. The vehicles are equipped with OBU and tamper-resistant hardware (TRH). With current technologies, OBUs are capable of carrying out necessary ordinary and cryptographic computations. Some of the Vehicles are also assumed to support 3G/4G data networks in order to channel the information from vehicles to cloud. This assumption is based on the fact that, currently almost all automobile companies produce cars with 3/4G network facility embedded in their high-end cars.

2. TRH is assumed to be semi-trusted entity because recent researches have shown that side-channel attacks are possible to tamper with TRH.

3. Secret keys and certificates are stored in TRH, and only authentic configuration of TRH is possible. Moreover all messages are assembled inside TRH.

4. There is working and policy-based contract between VANET handling authorities (i.e. government) and the cloud service provider(s). We assume certain functions that are not able to be carried out through technology, are carried out based on the policy in the contract, such as the service level agreement.

4.3. Network Model

Fig. 4 illustrates our proposed network model. We divide our proposed architecture into two infrastructures, VANET and cloud computing, connected through GTs. VANET architecture consists of traditional vehicular nodes, RSUs, and management/revocation authorities. Vehicles moving on the road serve both as producers (they share coarse-grained information with the cloud) and consumers (they subscribe for fine-grained traffic information and other services from the cloud). RSUs serve as GT between vehicles and cloud infrastructure. It is worth noting that in case if there is no access to RSU directly, then vehicles with 4G internet could be used as mobile GT.
Cloud architecture consists of authenticator, Cloud Processing Module (CPM), Cloud Knowledge Base (CKB), and Cloud Decision Module (CDM). Authenticator is responsible for handling subscriptions from vehicles and their authentication. Data-contained, coarse-grained MVs from vehicles are collected by cloud, stored at CKB and forwarded to CPM for processing. After processing coarse-grained MVs from the vehicles, CPM constructs fine-grained traffic information in segment chunks and forwards it to CDM. CDM then categorizes the fine-grained data based on physical locations. Without loss of generality, physical roads are divided into small manageable zones and zones are further divided into segments beforehand. The stored MVs are not only used for traffic information dissemination, but also used for variety of other purposes such as forensics for investigation, insurance claims in case of accidents, and so forth. Table 1 outlines the notations that will be used throughout the rest of the paper.

<table>
<thead>
<tr>
<th>S.No</th>
<th>Notation</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MV</td>
<td>Mobility Vector</td>
</tr>
<tr>
<td>2</td>
<td>MD</td>
<td>Mobility Data</td>
</tr>
<tr>
<td>3</td>
<td>GT</td>
<td>Gateway Terminal</td>
</tr>
<tr>
<td>4</td>
<td>VID</td>
<td>Vehicle ID</td>
</tr>
<tr>
<td>5</td>
<td>ZID</td>
<td>Zone ID</td>
</tr>
<tr>
<td>6</td>
<td>SegID</td>
<td>Segment ID</td>
</tr>
<tr>
<td>7</td>
<td>$AV_{segID}$</td>
<td>Average velocity at segment $segID$</td>
</tr>
<tr>
<td>8</td>
<td>$TD_{segID}$</td>
<td>Traffic Density at segment $segID$</td>
</tr>
<tr>
<td>9</td>
<td>$M_{TI}$</td>
<td>Traffic information message</td>
</tr>
<tr>
<td>10</td>
<td>$K_V$</td>
<td>Individual secret key of vehicle $V$</td>
</tr>
<tr>
<td>11</td>
<td>$K_Z$</td>
<td>Zone level secret key</td>
</tr>
<tr>
<td>12</td>
<td>$K_{geolock}$</td>
<td>Location-based encryption key</td>
</tr>
<tr>
<td>13</td>
<td>$K_{RSU}$</td>
<td>RSU’s key shared with vehicles to construct $K_{geolock}$</td>
</tr>
</tbody>
</table>
4.4. Geolock-based Encryption

Location-based encryption refers to an encryption scheme where the receiver of the encrypted message must be physically present at the location defined by the sender of the message. More precisely, location information is used to generate the key used for encryption and decryption. The purpose of location-based encryption is twofold: to provide location confidentiality against outsiders and to keep insiders from manipulating the content of the message.

Traditionally in VANET, cooperative awareness information is sent to the one hop neighbors in plaintext. In such scenarios, outsiders can generate movement profiles against the target nodes. Outsiders can also manipulate the location information in order to create illusions as a result of launching Sybil attacks. Therefore it is essential to provide location confidentiality against outsiders. However, insiders are still a risk since they can manipulate the location information as well. Scott et al. [37] proposed a location-based encryption technique to provide location confidentiality, however; the data used for constructing the key, is public and dictionary attack is possible. Moreover, the content can also be manipulated with such attacks. In order to diminish the effect of content manipulation, we propose an extended version of Yan et al.’s location-based encryption [28]. Their scheme uses only GPS information in order to construct the geolock key; however, GPS information is publicly available and everybody can have access to such information and they assume that GPS information is private. Therefore, Yan et al.’s scheme does not provide location confidentiality. Our proposed geolock-based encryption is given in Fig. 5.

![Geolock Key Construction](image-url)
Fig. 5 shows the construction process of geolock key denoted by $K_{geolock}$ which is used to encrypt MV. Geolock key construction module takes as input, the effective region size, message lifetime, $K_Z$, $K_{RSU}$, and then multiplexes these values altogether in order to calculate hash value from the multiplexed content. The effective region size is used to define physical region where $K_{geolock}$ is effective. Message lifetime defines the validity period of the message, which helps to prevent the useless lingering around of stale messages in the network. $K_Z$ is distributed among vehicles upon their arrival to the respective zone and vehicles moving from one zone to another are subject to change their $K_Z$ accordingly. $K_{RSU}$ is the RSU-specific key issued to the legitimate vehicle when it comes under the RSU’s transmission range. We assume the gradual deployment of VANET where there might be sections of roads without RSUs. We argue that the deployed RSUs know about their neighborhood, hence an additional key is issued to the vehicles approaching to a region where RSUs are not currently deployed. Our proposed location-based encryption mechanism uses different levels of confidentiality as follows: GPS coordinates are publically available, $K_Z$ is known to the legitimate VANET users in a specific zone, and $K_{RSU}$ is known to the vehicles currently present within the transmission range of RSU. Additionally it must be noted that some of the regions without RSUs might cover longer distance with the same $K_{RSU}$ but in that case, the time factor would limit the degree of content abuse. The receiving vehicles must be physically present in a certain geographic region (specified by the sender), at certain time, holding valid $K_Z$ and $K_{RSU}$ in order to decrypt the message encrypted with $K_{geolock}$. We assume that $K_Z$ and $K_{RSU}$ are subject to change on a regular or dynamic interval for security reasons; they have no contribution to security and privacy, otherwise. If the aforementioned keys are not changed, then a vehicle can move around in different zones and RSUs to get these keys, and later on it can manipulate information with those keys. However, the changing frequency may be different in case of $K_Z$ and $K_{RSU}$. The former covers larger area as compared to the latter. Initially these keys are distributed by the department of motor vehicles (DMV) which is the management authority. Outsiders are kept at the bay from manipulating the contents of the message and abusing users’ privacy, unless the aforementioned keys are compromised, otherwise. However, the regular or dynamic change in these keys will be essential to control insiders as well who possess these keys. We assume that the keys change should be based on the geographical region where different geographical regions exhibit different mobility dynamics, for instance highway and urban areas. More precisely, in the urban areas, the keys will be changed more frequently than in case of highway scenario, based on a statistical time interval required to pass through the zone and through an area under the jurisdiction of certain RSU. In other words, the keys are valid for a time period that is enough for the vehicles to traverse through a zone and through area that is under jurisdiction of certain RSU. In normal circumstance, these keys will be changed frequently on the regular intervals, however; besides regular intervals, there can be triggers for the change in these group keys, for instance the compromise of one or
both of these keys. We believe that many efficient key management and updating schemes exist in the literature. To the best of our knowledge, [32,33,34,35] can be incorporated with our scheme to carry out the group key update procedure without shaking the baseline functionality of our proposed scheme.

4.5. Secure Mobility Vector

As a cornerstone of VANET applications, vehicles rely on situational awareness information from neighbors in order to have smooth, safe, and reliable driving experience. Such awareness is realized through beacons shared among the vehicles. The information contained in beacon helps the neighbors to update their awareness about network topology. The traffic view constructed through beacons enables drivers to take right and timely decisions in case of hazard situations (for instance a traffic jam, an accident on the road, or an icy road segment). It is to be noted that the one hop beacons are used to construct short-range local traffic view. But even in case of one hop, the standard effective transmission range is not covered due to many factors including obstacles. Recent research studies found out that vegetation, tall buildings, and even tall vehicles create non line-of-sight and reduce the effective transmission range [30]. Hence in that case, our proposed cloud-based approach is ideal for not only covering the one hop area, but also enables vehicles to construct long range traffic view through fine-grained traffic information from the cloud, based on their locations and the direction of their movement.

We use location-based encryption (already explained in previous sub-section) for security and privacy reasons and encrypt MV with $K_{geolock}$. The format of MV is given below:

$$MV = \{MD, h(K_Y), h_{K_Y}(MD), h_{K_2}(\delta || MD)\}_{K_{geolock}} \quad (1)$$

$$MD = (h(VID), t_{cur}, loc_{cur}, vel_{cur}, dir)$$

$$\delta = h_{K_Y}(MD)$$

$t_{cur}$ is the current time, and $loc_{cur}$ is current position of the vehicle moving with velocity $vel_{cur}$ towards direction $dir$. 

![Fig. 6. Traffic Information dissemination](image-url)
It is worth noting that mobile GTs are chosen among the vehicular nodes with rich communication resources (3/4G data plan), hence they provide an extra communication channel to clouds in the areas where static GTs are not available. When an encrypted MV reaches a GT, GT first decrypts it with current $K_{geolock}$ and then saves $h(VID)$ and $h(K_V)$ in its database for liability reasons. These two parameters will help RAs in revocation, in case of a dispute. After that, GT forwards MV to the cloud for further processing.

4.6. Fine-grained TI (Traffic Information) Dissemination

VANET is divided into manageable physical domains and each domain has its own potential cloud infrastructure for traffic information dissemination. In order to guarantee the relevant and right information delivery to subscribers, domains are further divided into reasonable zones and each zone has its own zone level common secret key $K_Z$. As discussed in previous sub-section, $K_Z$ is used to construct $K_{geolock}$. Without loss of generality, going down to another level of hierarchy, each zone consists of segments which corresponds to small road segments (for instance in urban areas, a road segment between two traffic lights, or a block). CPM constructs traffic information based on physical locations (segments). Similarly after CPM constructs fine-grained traffic information corresponding to segments, it forwards the fine-grained information to CDM where CDM is responsible to provide the subscribers with the right and relevant information according to their current and near-future locations. The reason for segment level delivery is that the subscribers might only be interested in a certain area which logically means that vehicles might want to know the traffic conditions ahead of them in accordance with their future route. The desired information may vary depending upon the direction of the vehicle. The fine-grained traffic information message denoted by $M_{TI}$ is given below.

$$M_{TI} = (t_{cur}, ZID, SegID, TD_{SegID}, AV_{SegID}, others) \tag{2}$$

‘others’ corresponds to any warning message or alarm (for instance a black ice on the road, fog, or an ambulance approaching). It is worth noting that the pair $ZID$ and $SegID$ corresponds to a physical location and this information is used by CDM to forward the traffic information to concerned GT. When GT receives fine-grained traffic information from CDM, it forwards the information with slight tuning of security parameters to the vehicles. Vehicles moving on the road use V2I infrastructure to communicate with GTs and use Zero-Infrastructure (V2V) for inter-car communication. The flow of the messages is given below.

$$V \rightarrow GT: MV = \{MD, h(K_V), h_{K_V}(MD), h_{K_Z}(\delta||MD)\}_{K_{geolock}}$$
\( GT \rightarrow CKB: MV = \{GT, MD, h(K_T), h_{K_Z}(MD)\}_{K_{GT-c}} \)

\( CDM \rightarrow GT: M_{TI} = \{t_{cur}, ZID, SegID, TD_{SegID}, AV_{SegID}, others\}_{K_{GT-c}} \)

\( GT \rightarrow*: \{GT, M_{TI}, h(GT||M_{TI})\}_{K_{geolock}}, Sig_{Pr_GT}(GT||M_{TI}) \)

\( K_{GT-c} \) is the shared secret key between \( GT \) and the cloud module. It is worth noting that \( GT \) and cloud module both save the values \( h(VID) \) and \( h(K_T) \) for liability reasons. These values serve as anonymous proofs for revocation in case of any dispute. When \( GT \) receives fine-grained traffic information about the segments which are under its jurisdiction, it signs the message before forwarding it to the subscriber vehicles in order to prove its authenticity. The overall protocol is shown in Fig. 6.

5. Quantitative Evaluations and Analysis

In this section we quantitatively evaluate our proposed TIaaS scheme from security, privacy, revocation, and computational overhead standpoint. We also analyze the anonymity of the users and the number of subscribers and producers with the quality of information.

5.1. Security

**Lemma 5.1.1:** Our proposed scheme provides loose message authentication, integrity, confidentiality, timeliness, and non-frameability.

**Proof:** The security requirements in our proposed scheme are message authentication, message integrity, confidentiality, timeliness, privacy protection, anonymity revocability, and non-frameability. We argue that the proposed scheme fulfills the aforementioned requirements. Our proposed scheme provides security for \( MV \) and \( M_{TI} \) since the messages are sent in encrypted form. In case of \( MV \), the message is encrypted with \( K_{geolock} \) and only legitimate vehicles and \( GT \) that hold and/or can construct \( K_{geolock} \) can decrypt the message. The security of the message depends upon security of \( K_Z \) and \( K_{RSU} \) that are used to construct \( K_{geolock} \). \( K_{geolock} \) keeps outsiders from manipulating the messages and also limits the effect of stale messages in the network. In other words, \( K_{geolock} \) guarantees the freshness of the message. For insiders to decrypt the message in a timely manner, they must be physically present in the effective area where current \( K_{geolock} \) can be constructed. Therefore our proposed scheme provides location confidentiality through \( K_{geolock} \). The integrity of the contents is checked by calculating \( h_{K_Z}(MD) \). Due to high frequency of \( MV \) (in the order of milliseconds), we use loose authentication for \( MV \) by using keyed HMAC. Hence our proposed scheme fulfills authentication and message integrity in case of \( MV \). In case of \( M_{TI} \), \( GT \) is authenticated with signature, and the value \( h(GT||M_{TI}) \) guarantees the integrity. \( \blacksquare \)
Lemma 5.1.2: If $K_v$ is not compromised, then it is very hard to impersonate or frame another benign entity.

Proof: In our proposed scheme, it is very hard to impersonate or frame another vehicle until $K_v$ is compromised, otherwise. We include $\delta = h_{K_v}(MD) \cdot h_{K_z}(\delta || MD)$ which makes sure that for each $MV$, the value of $\delta$ will be different and only legitimate node with $K_v$ will be able to calculate the aforementioned value.

Confidentiality and timeliness are provided by $K_{\text{geolock}}$. Only legitimate users having $K_z$ and $K_{RSU}$ will be able to decrypt the message and after the message gets stale (in other words the validity time of the message expires), the message will not be able to be decrypted because time factor is included in $K_{\text{geolock}}$ construction. However if $K_z$ and $K_{RSU}$ are compromised, then outsiders can manipulate messages. Nevertheless they have to be physically present in the effective location, where the message can be decrypted. Moreover if the attackers are not present at the right place, then they need to construct $K_{\text{geolock}}$ by brute force which will be challenging.

The following corollary naturally follows from the above lemma.

Corollary: Outdated messages are not processed by the system.

In our geolock-based encryption we used the time parameter which makes sure that only fresh messages get decrypted. Nevertheless we assume that VANET and cloud infrastructure agree upon the message lifetime beforehand or they use standard lifetime in case of DSRC/WAVE. If the message validity is denoted by $t_{\text{validity}}$, then according to the construction mechanism of $K_{\text{geolock}}$, receiver of a beacon message can only decrypt the message if it is in the effective region of the $K_{\text{geolock}}$ and if $\text{Dist}(x, y)$ denotes the Euclidean distance between $x$ and $y$ then the condition $\text{Dist}(s, r) \leq \text{Effective Region}$ must be satisfied. Moreover $t_{\text{cur}} - t_{\text{transmit}} \leq t_{\text{validity}}$ must be satisfied as well where $t_{\text{cur}}$ is the current time and $t_{\text{transmit}}$ is the time when beacon was transmitted. If the aforementioned conditions are not satisfied, $K_{\text{geolock}}$ will not be constructed and henceforth stale messages will not be processed by the system.

Lemma 5.1.3: The privacy of vehicles in the vicinity is not compromised until and unless $K_{RSU}$ and $K_z$ are compromised. In other words, it will be very hard for the adversary to get VID even if $K_{RSU}$ and $K_z$ are compromised.

Proof: Let suppose that the above statement does not hold, then we have $K'_{RSU}$ and $K'_z$ such that $K'_{\text{geolock}}$ can be constructed by a user and:
\[ MV = \{MD, h(K_v), h_{K_v}(MD), h_{K_z}(\delta||MD)\}^{K_{\text{geolock}}} \] (3)

and

\[ \forall mv_t \in MV_{[t,t_f]}, VID \leftarrow h(VID) \wedge \sigma_t \leftarrow h_{K_v}(MD) \] (4)

The above statement states that when an adversary has compromised \( K_{\text{RSU}} \) and \( K_z \), then the adversary can obtain \( h(VID) \) and the vehicle’s privacy can be abused with calculating the integrity of the message from \( h_{K_v}(MD) \). If the hash functions are collision-resistant, then by dictionary attack the statements in above equations contradict. It will be very hard to calculate \( VID \) from \( h(VID) \) and so will be the profilation and privacy abuse.

**Lemma 5.1.4:** The compromise of only \( K_v \) does not pose any threat to the entire system, rather only affects the owner of the compromised \( K_v \).

**Proof:** Let suppose an adversary \( \mathcal{A} \) compromised vehicle \( V \)’s \( K_v \) and uses it as \( K_v' \). With compromised \( K_v' \), \( \mathcal{A} \) can calculate \( h(K_v') \) and \( MD \) as:

\[ MD' = (h(VID'), t'_\text{cur}, \text{loc}'_\text{cur}, \text{vel}'_\text{cur}, \text{dir}'_\text{cur}) \] (5)

\[ MV = \{MD', h(K_v'), h_{K_v'}(MD'), h_{K_z}(\delta'||MD')\}^{K_{\text{geolock}}} \] (6)

It is worth noting that the adversaries are of two kinds, i.e. Outsiders \( \mathcal{A}_O \) and insiders \( \mathcal{A}_I \). In case of \( \mathcal{A}_O \) even if \( K_v \) is compromised, the adversary \( \mathcal{A}_O \) cannot necessarily abuse the privacy of the user because the beacon is encrypted with \( K_{\text{geolock}} \) provided that \( K_{\text{geolock}} \) is not compromised. In case of \( \mathcal{A}_I \), the adversary is a legitimate VANET user and can construct the current \( K_{\text{geolock}} \). In such scenario, all the messages from the point when \( K_v \) was compromised, i.e. \( \forall mv \in MV_{[t_\text{cur},t_\infty]} \) (all future messages onwards) can be manipulated by \( \mathcal{A}_I \). However such compromise does not pose any threat to other users of the system because \( K_v \) is possessed by a single entity. Nevertheless we argue that in case of misbehavior, the entity is subject to revocation anyways.

**Lemma 5.1.5:** It takes universal brute force attack for \( \mathcal{A}_I \) to compromise \( K_{\text{geolock}} \).

**Proof:** In order to decrypt the message that is encrypted with current \( K_{\text{geolock}} \), the decrypter must be physically present in the area where current \( K_{\text{geolock}} \) is valid and at the right time. Any insider legitimate user who is not currently present in the aforesaid area is considered as \( \mathcal{A}_I \). In order for \( \mathcal{A}_I \) who is not
physically present in the area where current $K_{geolock}$ can be used for decryption, must construct all possible combinations of $K_{geolock}$ for the decryption of the message. When $A_t$ receives message encrypted with current $K_{geolock}$ which is constructed with $K_Z, K_{RSU}, t_{cur}$, and $loc_{cur}$, it will be hard for $A_t$ to figure out which $K_Z$ and $K_{RSU}$ are used to construct $K_{geolock}$ and that is why $A_t$ has to try all combinations of these two keys and the GPS locations in each zone. More precisely $A_t$ has to try all zones and RSUs therein. Additionally even in the single zone and under a single RSU, the GPS coordinates are also important. If there are $n$ zones, $s$ RSUs in each zone and $l$ locations under the jurisdiction of each RSU, then $A_t$ must try the following number of keys to decrypt the message encrypted with current $K_{geolock}$.

Moreover the time factor is an important issue in such brute force because after the expiry of the validity time denoted by $t_{validity}$ which is also an input to the $K_{geolock}$, the key cannot be constructed and becomes stale.

5.2. Conditional Privacy and Revocation

Lemma 5.2.1: Our proposed scheme preserves conditional privacy and efficiently revokes the message/user in case of a dispute.

Proof: We do not include any identity information in the messages that leads to link the message to a particular user. $MV$ and $MT_I$ are exchanged anonymously. However for RAs, it is possible to revoke a message. Since $GT$ saves the values $h(VID)$ and $h(K_V)$ in a table and in case of any dispute $GT$ can provide the accused values to RAs, RAs have the vehicle’s secret key information beforehand. We assume that the secret key is stored in RAs in encrypted form and RAs collude to decrypt the key and revoke the user. By looking up the credential table with hash values, the revocation complexity can only $O(1)$ depending upon the hash function and its implementation. It is worth noting that although $GT$s and cloud save $h(VID)$ and $h(K_V)$, it does not have an adverse effect on the privacy because the hash values do not reveal any valuable information about the legitimate users.

5.3. Computation and Communication Overhead, and Comparison with other Known Schemes

In this sub-section we reason about the computation and communication overhead incurred by our proposed scheme. Moreover we also compare our proposed scheme with the known schemes of its kind.
To the best of our knowledge, the most relevant work to our proposed scheme is carried out by Qin et al. [24] and Bernstein et al. [22]. Ours is the first effort to propose a mechanism where MVs are stored in the cloud and leveraged as a source of traffic information as a service to the vehicles on the road. However Qin et al. [24] and Bernstein [22] did not address potential security and privacy issues. Our proposed scheme is secure and conditional privacy-preserved. Qin et al. leverages Time Space Link Graph (TSLG) to define vertices and edges in VANET. In their proposed scheme, every vehicle can be source as well as destination at the same time. So in worst case, the number of links can be \( \left( \frac{n(n-1)}{4} \right) \) on average and the order of routing optimization is \( O(n^2) \) where \( n \) is the number of vehicles in the network.

Bernstein et al. also do not address security concerns in PaaS architecture in connected cars. The computation overhead incurred by our proposed scheme is given below.

i. A vehicle while sending MV to GT performs \( 2h + 1E \) operations where \( h \) denotes hashing operation and \( E \) denotes encryption.

ii. When a vehicle receives \( M_{TI} \) from GT, it performs \( 1h + 1D + 1E \) where \( D \) denotes decryption and \( V \) is the signature verification. Without loss of generality, we assume ECMV [29] scheme by Wasef et al. for signatures in our scheme.

iii. For GT to send an MV to the cloud and receive the fine-grained traffic information from the cloud, performs \( 1S + 2D + 2E + 2h \). GT upon receiving MV from vehicles decrypts it and re-encrypts it with the shared key between itself and cloud module. The same process repeats in opposite direction when GT receives \( M_{TI} \) from cloud and forwards it to the vehicle. \( S \) denotes the signature operation and GT puts its digital signature on the contents of before sending it to the subscriber vehicles.

The communication overhead incurred by the keys update (\( K_{RSU} \) and \( K_Z \)) is different in case of highway and urban scenarios. In the highway scenarios, the key update on a normal regular basis will be lower than in case of urban scenarios because the mobility of the vehicles is mostly quasi-static and the chance of using the same keys is lower. Whereas in case of urban scenarios, the keys update is dependent on the area covered by a zone and the area under jurisdiction of RSU. Therefore \( K_{RSU} \) will be updated with higher frequency than \( K_Z \) because \( 2Tr_{RSU} \ll 2Tr_Z \), where \( Tr_{RSU} \) is the transmission radius of the RSU and \( Tr_Z \) is the transmission radius of the zone. It is worth noting that the aforementioned keys update is carried out in a normal scenario without any key compromise. In case of any key compromise detected by the authorities and/or reported by the users, it should be updated right away. The overhead incurred by the keys updates is outlined in [32,33,34,35]. The aforementioned schemes are semantically close to our scenario and can be implemented in our proposed scheme.
5.4. Analysis on the Number of Producers and Subscribers

Without loss of generality, we assume homogeneous traffic flow with total traffic density \( d \) and global market penetration rate \( \alpha \) such that the relevant density variable is the partial density of OBU-equipped vehicles, i.e.,
\[
\lambda = d \times \alpha \quad (7)
\]

Since \( \lambda \) represents the number of OBU-equipped vehicles, let \( X \) be a random variable representing the number of legitimate VANET users who share their information with cloud, we take the expectation of \( X \) as follows.
\[
E[X] = d \times \alpha \quad (8)
\]

For the sake of understanding, we consider a single zone \( Z \) to analyze the probability of VANET users who share their whereabouts with cloud modules and subscribers who receive fine-grained traffic information from the cloud. It is worth noting that VANET legitimate users can be producers and consumers at the same time. We argue that the local analysis will have global impact based on our homogeneous traffic assumption. Nevertheless variations are not out of question (for instance rush hour or traffic hot spots such as intersections). Let \( Y \) be a random variable representing the number of consumers, then expectation of \( Y \) is given below.
\[
E[Y] = d \times (1 - \alpha) + \beta, \quad b \leq \lambda \quad (9)
\]

Where \( \beta \) represent vehicles that are producers and consumers at the same time (i.e. VANET users). The quality of traffic information produced by the cloud depends upon the traffic density and global penetration rate. Equation 7 shows that if the global penetration rate increases, the number of expected producers also increases whereas a decrease in penetration rate results in a decreased number of producers, hence the low quality of information and vice versa. But it is worth mentioning that places where VANET is deployed in its initial stages also use the current traditional ways of traffic information (CCTV, roadside sensors etc.).

5.5. Anonymity Analysis

To measure the privacy of a node in a particular zone, we measure the anonymity of the reporter by calculating entropy denoted by \( \mathcal{H} \). To calculate the entropy, we assume that the anonymity set is the set of the users around the site of interest (SoI) denoted by \( V \) and let \( p_{V_i} \) be the probability that the node \( V_i \) be the producer to the cloud where \( \forall V_i \in V \) and the summation of all the probabilities collectively is 1, i.e.
\[
\sum_{i=1}^{\vert V \vert} p_{V_i} = 1.
\]
In our case SoI is the area that is the jurisdiction of RSUs in a particular zone. The entropy $H$ of the target producer $V_i$ in the anonymity set $V$ is given below:

$$H = - \sum_{i=1}^{\mid V \mid} p_{V_i} \times \log_2 p_{V_i}$$  \hspace{1cm} (10)$$

Since our anonymity set is $V$, the possible outcomes can be $\mid V \mid$ assuming the fair distribution and the probability of each outcome will be $\frac{1}{\mid V \mid + (N - \mid V \mid)}$. $N$ is the number of all vehicles in the zone under consideration. If the distribution is normal and the occurrence of the nodes to be producers, i.e. share their beacons with cloud infrastructure is equally likely, then the maximum entropy is also given by the following formula.

$$H_{\text{max}} = - \sum_{i=1}^{\mid V \mid} p_{V_i} \times \log_2 p_{V_i} = \log_2 p_{V_i}$$  \hspace{1cm} (11)$$

It must be noted that in such case, the normal entropy is equal to the maximum entropy, i.e. $H = H_{\text{max}}$. However due to the ephemeral nature of VANET, the situation where $H$ equals to $H_{\text{max}}$, is hard to achieve. The entropy of the anonymity does not only depend upon the anonymity set, but also depends upon the individual probability of the producers. However in general, the more are the elements in anonymity set, the higher is the entropy provided that the nodes are distributed equally likely. On the other hand, in our case, the anonymity of the producer can be variable depending upon its location, market penetration rate of VANET users, and traffic density around SoI.

### 6. Conclusions and Future Directions

Traditional Vehicular Ad Hoc NETwork (VANET) and cloud computing have caught the eyes of the academia and industry as a result of remarkable research results in these fields. Recently, keeping in mind the resources-rich capabilities of high-end vehicles, VANET evolved to VANET-based cloud as a result of merging VANET and cloud computing. In this paper, we take the VANET-based clouds a step further towards its applications and services. We leverage Hussain et al.’s Vehicles using Clouds (VuC) framework and propose a new secure and privacy-aware service namely Traffic Information as a Service (TIaaS) atop the service stack of VANET-based clouds. Highly mobile vehicular nodes share their coarse-grained whereabouts information (Mobility Vectors) with the stationary or mobile Gateway terminals (GT) and GT forwards coarse-grained individual information to the cloud. Cloud infrastructure saves MVs and after processing and constructing fine-grained traffic information, sends it back to vehicles on the road through GTs, based on the physical locality of the vehicles. Our proposed scheme preserves conditional privacy and other security parameters like authentication, integrity, and non-frameability. To guarantee
the smooth service to the subscribers, a working service level agreement must be in place that defines the processing and communication delay to the clouds. In future, we plan to focus more on subscribing process of the vehicles. Moreover to stimulate the active participation of the vehicles in the service, we aim at incentive-based TIaaS where vehicles will be credited with incentives based on their contribution, in a privacy-aware manner.

Acknowledgements

This research was supported by the MSIP (Ministry of Science, ICT and Future Planning), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2015-H8501-15-1007) supervised by the IITP (Institute for Information & communications Technology Promotion).

This research was supported by the MSIP (Ministry of Science, ICT and Future Planning), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2015-H8501-15-1018) supervised by the IITP (Institute for Information & communications Technology Promotion).

This work was also supported in part by the NRF (National Research Foundation of Korea) grant funded by the Korea government MEST (Ministry of Education, Science and Technology) (No. NRF-2012R1A1A2009152).

This work was also supported in part by the NRF (National Research Foundation of Korea) grant funded by the Korea government MEST (Ministry of Education, Science and Technology) (No. NRF-2012R1A2A2A01046986).

References


