TrustVote: Privacy-Preserving Node Ranking in Vehicular Networks

Muhammad Ajmal Azad, Samiran Bag, Simon Parkinson, and Feng Hao *Senior Member, IEEE*

Abstract—The Internet of Vehicles (IoV) is the network of connected vehicles and transport infrastructure units (Roadside Units (RSU)), which utilizes emerging wireless systems (4G, 5G, LTE) for the communication and sharing of information. The network of connected vehicles enables users to disseminate critical information about events happening on the road (for example accidents, traffic congestions, and hazards). The exchange of information between vehicles and roadside units could improve the driving experience and road safety, as well as help drivers to identify the hazardous and safe routes in a timely manner. The sharing of critical information between vehicles is advantageous to the driver; however, at the same time, malicious actors could mislead drivers by spreading fraudulent and fake messages. Fraudulent messages can have a negative impact on the infrastructure, and more significantly, have potential to cause threats to life. It is therefore essential that vehicles can evaluate the credibility of those who send messages (vehicles or roadside units) before taking any action. In this paper, we present TrustVote, a collaborative crowdsourcing-based vehicle reputation system that enables vehicles to evaluate the credibility of other vehicles in a connected vehicular network. The TrustVote system allows participating vehicles to hide their rating/feedback scores and the list of interacted vehicles under a homomorphic cryptographic layer, which can only be unfolded as an aggregate. The proposed approach also considers the trust weight of a vehicle providing the rating scores while computing the aggregate reputation of the vehicles. A prototype of TrustVote is developed and its performance is evaluated in terms of the computational and communication overheads.

Index Terms—Privacy-preservation, Vehicular networks, Reputation system, Secure Multi-party Computation, Private Crowdsourcing

1 Introduction

It is predicted that there will be more than 250 million connected vehicles by the year 2020 [1], [2]. The emergence of new telecommunication technologies (e.g., 4G, 5G networks) enables vehicles to interact with each other and with the transport infrastructure (i.e., roadside units (RSU)) for value-added services. Vehicles in a connected vehicular network could collaborate to provide time-critical information about what is happening on the road, for example, traffic congestion, roadside accidents and road hazardous. Such information is utilized by vehicles and the infrastructure units alike to make decisions. For example, a vehicle may adjust its navigation route due to receiving information regarding congestion, and a traffic management infrastructure might adjust traffic control timings to optimize throughput.

Figure 1 represents the typical system model of a connected vehicular network, which consists of two main participants: vehicles each with a wireless communication link, and the RSUs. The communication between vehicles is termed as Vehicle-to-Vehicle (V2V) communication and the communication between vehicles and RSU is termed as Vehicle-to-Infrastructure (V2I) communication [3]. The RSU is served as the anchor point between vehicles to extend the footprint of the vehicular network. Vehicle-to-vehicle (V2V) and Vehicle-to-Infrastructure communication can take place using any of the communication technologies such as the short-range communication mechanisms (Dedicated Short Range Communication technology (DSRC)) and mobile technologies (3G, 4G, 5G, WiMAX, LTE etc) [4]–[6]. The vehicles share information (such as traffic information, road situations, road congestion etc.) to each other which could be utilized in both safety and navigation systems. Vehicle-to-infrastructure (V2I) is similar in technical implementation; however, the communication is with infrastructure components such as Reduced Speed Zone warning signs, which can automatically communicate key information to any approaching vehicle or other RSUs. The RSUs can be connected to each other through the wired network or a high speed dedicated wireless link. The technique presented in this paper is infrastructure independent, meaning that the implementation is not directly related to the physical network connection. More specifically, it is envisaged that the implementation would reside within the Session layer of the Open Systems Interconnection model (OSI model).

The communication between vehicles and RSUs has provided an opportunity for drivers and vehicles to know about events happening on the road. Drivers can make decisions based on reports they received from other vehicles and RSUs, and more significantly, autonomous vehicles will use received and sensed information to make independent decisions [7]. A malicious adversary could disseminate fake and unwanted messages to other vehicles and RSUs in the network. The distribution of non-trustworthy and malicious messages over a large network (in terms of vehicles and RSUs) could be a serious threat to the safety of drivers. Following on with the example of congestion and its impact on navigation, disseminating spurious messages could cause vehicles to take a remedial action by re-routing, but in doing so could cause the road infrastructure to gridlock.

• *Muhammad Ajmal Azad, Samiran Bag and Feng Hao are with the Department of Computer Science at the University of Warwick, United Kingdom. Simon Parkinson is with the School of Computing and Engineering at the University of Huddersfield, United Kingdom. E-mail:{muhammad.azad,samiran.bag,feng.hao}@warwick.ac.uk, s.parkinson@hud.ac.uk*

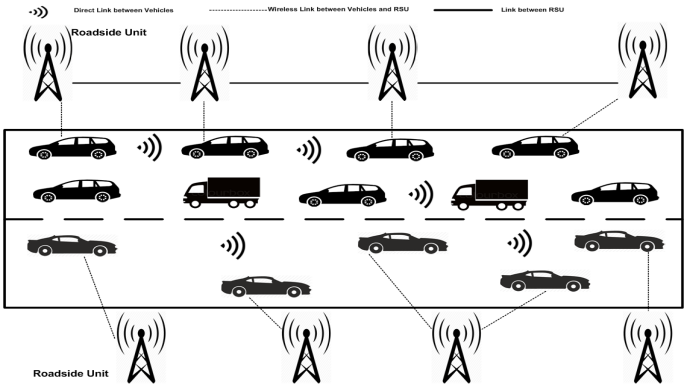


Fig. 1: System Model of Autonomous Vehicular Network.

It is therefore important that vehicles should evaluate the trustworthiness of the message sender before making any maneuver and dissemination of messages to others.

Evaluating the trustworthiness of vehicles in the connected vehicular network is challenging because of the distributed and ad-hoc nature of the vehicular network. The trustworthiness of vehicles can be computed in two ways [8]–[13]: 1) relying on a trusted third party system for handling data reported by the vehicles, 2) implementing a distributed system where vehicles can exchange data directly with each other. In a centralized system, the centralized trusted third party serves as the anchor point between vehicles in the network. However, the centralized system has the following limitations: 1) it can be a single point of failure, 2) it can be a central target of an attack, 3) it can pose threat to the privacy of vehicles as it holds sensitive data submitted by the vehicles [14], [15]. In a distributed system, no central authority exists for evaluating the trustworthiness of nodes, but it brings forth the challenges of privacy-preservation and scalability.

Vehicles can have different levels of trustworthiness in vehicular networks. More specifically, if the vehicle is previously known to be sending correct information or trustworthy information then this vehicle is considered to be more trustworthy than any newer vehicle or some other vehicles in the network. The trust weights of vehicles should be taken into consideration while computing the aggregated reputation as this gives more credibility to messages from the trusted vehicles than other or non-trusted vehicles. The challenges in the design of a weighted reputation system for the connected vehicular network are three-fold: 1) the system should ensure privacy of vehicles providing their feedback about the behavior of other vehicles, 2) it should protect the interaction network of vehicles and positions of vehicles reporting the feedback, and 3) it should perform all computation in a decentralized manner without consuming high computation and communication resources.

In this paper, we present TrustVote, a collaborative ranking system that enables vehicles to evaluate the trustworthiness of nodes in the network without relying on any trusted third party system. The TrustVote system enables vehicles to utilize feedback from a group of vehicles in a privacy-preserving manner. Vehicles are placed in a crowdsourcing group, which provide feedback scores about the behavior of the vehicle by encrypting them and only the aggregate will be revealed to other participants. The TrustVote system is obtained via the integration of a cryptographic mechanism and the randomization technique used in a boardroom voting [16]. The design approach has an important feature of not using the third party for the management of cryptographic parameters in the encryption and decryption phases. In addition, the design of TrustVote has the feature of weighted aggregation, allowing RSUs to assign different weights to vehicles or response providers in the crowd group. The building blocks of the TrustVote system consist of two main components: 1) crowd vehicles providing ratings about the credibility of other vehicles they have had interaction, and 2) a public tally system that collects rating values from the selected crowd group. The proposed approach also has an inherent mechanism for restraining feedback providers from providing out-of-range rating through the use of non-interactive zero-knowledge proofs of correctness. We have prototyped the system and evaluated its performance in terms of computation and bandwidth overheads. Furthermore, we present security proofs to analyze the privacy and security properties of the TrustVote system.

1.1 Contributions

In summary, this paper makes the following contributions:

- We present design of a novel decentralized crowdsourcing-based system for computing the rating of nodes in a vehicular network. The designed system does not require any trusted system for handling cryptographic operations.The system has an inherent mechanism of assigning weights to nodes in a crowd group. The system can also be utilized in any scenario where privacy-preserving weighted voting is required.
- We present security proofs to analyze both security and privacy properties of the system.
- We implement a prototype to evaluate the performance of the system for the computation and communication overheads.

1.2 Outline

The paper is structured as follows. In Section 2 we define the problem. Section 3 presents discussions on the related works. Section 4 presents preliminaries and our threat model. Section 5 presents the system architecture and operations of the TrustVote system followed by a comprehensive discussion on the privacy and security properties in a Section 6. Section 7 presents complexity analysis. Section 8 evaluates the performances of TrustVote system. Finally, Section 9 presents a conclusion and motivates future research.

2 Problem Definition

Assume there are U vehicles that communicate with other vehicles and RSU. The vehicles in the network broadcast their sensed information about what is going on the road to other vehicles. The vehicles also rate other vehicles based on the authenticity of the received messages. Let S be the rating value that a vehicle i $(i \in 1 \ldots, n)$ has assigned to the vehicle j $(j \in 1 \ldots, n)$, and $S \in 0, 1$. We model the

vehicle rating network as a weighted graph network G = (U;S;W). Here, U is a set of vehicles rating each other, S is the rating value assigned by the vehicle U_i to the vehicle U_j, and W is the trust weight of the vehicle rating others. The link between U_i and U_j exists, only if U_i has rated U_j at least once. In this paper, we investigate how to compute the weighted aggregated statistics of a particular vehicle while keeping trust weights and rating values of vehicles private to them.

In the context of rating in the vehicular network, we define the problem as follows. Let there be a network of U vehicles identified as $U_1, U_2, \ldots, U_n$. Each U_i rates the vehicle U_j based on the information it receives from the vehicle with a score $s_i \in \{0, 1\}, \forall i \in [1, n]$. There is a tally system placed at the RSU that provides an anchor point to data provided by the participating vehicles. The RSU uses this data to compute the aggregate ratings of a particular vehicle. Let $W = \{w_1, w_2, \ldots, w_n\}$ be the trust weights of the rater assigned by the RSU based on their past behavior. Each $w_i \in [a]$, where a is a small integer. The tally system provides an opportunity to compute the weighted average reputation $\tilde{S} = \frac{S}{\sum_{i=1}^{n} w_i} = \frac{\sum_{i=1}^{n} w_i s_i}{\sum_{i=1}^{n} w_i}$ of the target vehicle without revealing the values of responses provided by the participating vehicles.

3 State of the Art

Several approaches have been proposed for the management of reputation and trust in the vehicular network. These systems can operate in the centralized and decentralized system settings [17]–[19]. Ray et al. [20] proposed a data-centric framework for evaluating the trustworthiness of nodes in the vehicular ad-hoc network (VANET). The system weighs each individual response according to well-defined rules and takes into account the trust weight of the response provider. However, the privacy of the response provider or nodes in the network has not been considered in this design. A centralized system can be used to get the feedback responses from the vehicles in the network. This centralized system provides an anchor point that announces which vehicles in the network are trustworthy and eliminates the non-trustworthy vehicles from the aggregation process. BiBmeyer et al. [17] proposed a Misbehavior Evaluation Authority (MEA) that operates in the back-end infrastructure. It collects reports from nodes that are directly involved in witnessing the road hazard. For the detection of ghost vehicles, the MEA uses trust information from the participant's reports. Costantino et al. [21] proposed a centralized reputation system that identifies insider attackers by considering contextual information derived from sensors spread along the entire urban network. The roadside unit acts as the centralized system for the data from the vehicles in the network. Li et al. [18] utilized the centralized authentic infrastructure for the collection of trust votes from the legitimate nodes in the VANET. Their system excludes votes from malicious nodes in the final aggregation. Marmol et al. [22] proposed a reputation system that computes the reputation of a node by considering information from three different sources i.e. direct previous experiences with the target node, recommendations from other surrounding nodes and the recommendation provided by a central authority. The RSU holds the direct experience of nodes and distinguishes the malicious or selfish nodes in VANETs with high efficiency and accuracy. Oluoch [23] proposed a system where the vehicle asks other vehicles to provide their feedback about the target vehicle. The system then aggregates all the feedback with the help of RSU. Dewan et al. [24] compute reputation of a node by analyzing the number of packets forwarded by the node in the network. The nodes become highly reputed if they correctly route packets for other nodes. Huang et al. [25] proposed DREAMS, a distributed reputation management system where vehicles outsource their reputation computation to the edge nodes in the network. The centralized system has some limitation: 1) it can be a single point of failure and a single point of attack to get private information about participants, 2) the nodes or users in the network have to trust the centralized system for the protection of their private information.

Decentralized reputation systems, on the other hand, do not require the centralized trusted system for the collection and aggregation of the rating scores [26]. Fischer et al. [27] proposed a reputation system for the Vehicular AdHoc Network that considers the direct and indirect observation of a node towards its neighbors. It does not consider a trusted third party system for the management of data and responses. Huang et al. [28] proposed a weighted voting system for computing the reputation of the nodes in the network. The node's weight is computed on the basis of the nodes distance from the event. Yang et al. [29] proposed a blockchain based decentralized trust management system in the vehicular networks. In this system, vehicles can assign a positive vote if they trust the neighboring vehicles and upload ratings to the Roadside Units (RSU). The RSU then calculate the trust score for the involved vehicles and pack these data into a trusted blockchain which is maintained by all RSUs. The system is decentralized but it has not provided any mechanism for ensuring vote privacy of vehicles. Further, the system does not consider the trustworthy weight of participating vehicles.

Recent research efforts have developed techniques for preserving privacy in the vehicle communication systems. Asuquo et al. [30] discussed the challenges associated with maintaining privacy in location-based systems. The work discusses many potential solutions for maintaining location privacy, which mainly centers around the use of cryptography for maintaining the anonymity of the user. The challenge with the approaches presented in the paper is that they are centralized techniques, whereby a single authority is responsible for implementing the anonymity techniques. This central authority could be the key distribution system within a cryptographic system, which is responsible for restricting the system to valid users. The problem addressed in this paper requires a decentralized approach due to the scale of the challenge.

In current systems, such as those proposed in [30], trust is maintained centrally and is vulnerable to attack. Hussain et al. [31] provide a solution to maintain privacy in witness services in a vehicle infrastructure. Their system implements the ElGamal encryption algorithm with elliptic curve cryptography to protect data and uses pseudonyms mechanism for the identity exchange in order to maintain anonymity.

However, pseudonyms do not provide absolute privacy protection and identities can be linked using background information. Further, the system is dependent on the centralized system for the public and private keys management. Similarly, Zhu et al. [32] present a technique for maintaining privacy in a vehicle social network using the cryptography method. Although the work preserves identities but does not address the need to improve the trust of vehicles within a network. The system requires the trusted authority for the key management.

The existing reputation systems have not given much attention to the privacy and data integrity of participants providing feedback about certain events. The participation in the aggregation process could leak the location privacy and the interaction network of vehicles on the road. It is important that a reputation system should ensure the privacy of nodes in the network without using any centralized system. Furthermore, the system should also consider the weights of nodes while aggregating the individual feedback without revealing the trustworthy weight of the node. The reputation system proposed in this paper is not dependent on the underlying architecture of the vehicular network. The salient features of the TrustVote system are: 1) It allows nodes to compute the reputation of a node without relying on any trusted third-party system, 2) it utilizes weights assigned to nodes based on their behaviour, 3) the system ensures privacy of the nodes with the use of an efficient cryptographic system, 4) the computation is performed in the decentralized setting.

4 Background and Threat Model

In this section, we present definitions, the threat model and background on the homomorphic cryptographic system.

4.1 Definitions

In this section, we present definitions of several important concepts.

Definition 1. **Trust:** Trust represents ones own direct confidence or experience with others. It plays an important role in evaluating the behaviour of entities in the community. In our context of a vehicular network, for any two vehicles v_i and v_j, we use t_{ij} to represent the trust that the vehicle v_i has in the vehicle v_j. We define t_{ij} as the binary value i.e. either the vehicle v_i trust vehicle v_j or not. The trust of vehicle v_j in vehicle v_i is denoted as v_{ji}. In this paper, our focus is on how trust values from the vehicles are used to compute the aggregate reputation of the vehicle without revealing the values of trust scores.

Definition 2. **Reputation:** Reputation can be termed as an aggregate measure of trustworthiness based on the direct trust scores provided by the participants in the community. Let N vehicles have provided their trust scores for the vehicle j i.e. $t_ij = t_{1j}, t_{2j}, t_{3j} \dots t_{Nj}$, then the reputation of vehicle j can be computed as the average of the direct scores provided by the participating vehicles as $R_j = \frac{\sum t_{ij}}{N}$

Definition 3. **Privacy In Vehicular Network** Vehicles in the vehicular network communicate with other vehicles and the fixed infrastructure to have a collaborative value-added computation about conditions on the road. To improve the services, the infrastructure units also like to know the number of vehicles in a particular geographical region without knowing the identities of the vehicles. The exchange of information either vehicle to vehicle or vehicle to infrastructure could bring some benefits towards driver safety but the exchanged information could be used by anyone or adversary to track the private information of vehicles. The trust scores of the vehicles on the other vehicles could provide information about the connectivity network of vehicles and possibly reveal the geographic positions of the vehicles. The emphasis of this paper is to camouflage the trust and the connectivity network of vehicles such that information remains unlikable and untraceable.

Definition 4. **Privacy-preserving Reputation System:** Reputation systems have been particularly designed to collect and analyze personal feedbacks or trust values provided by the participants of the reputation system. The privacy-preserving reputation system can be defined as the reputation system that aggregates the feedback of participants without posing any threat to their private data. There are two major challenges in the design of privacy-preserving reputation systems for the vehicular networks: 1) the value of the trust scores remains hidden and unlikable, i.e. only weighted aggregated value of the trust score is revealed to the protocol participants, 2) the values could not be used to infer the relationship network of vehicles. Additionally, the system should exclude the out-of-range prescribed feedback values. Sometimes vehicles may behave maliciously in reporting their feedback scores. They usually assign exceedingly high feedback scores to maliciously increase the reputation of some vehicles. This misbehaviour may affect the accuracy of the aggregated reputation. In this paper, we consider a vehicle as the malicious if it assigns out-of-range feedback score to others.

4.2 Threat model and security goals

Our system ensures privacy protection of participants under the following conditions and assumptions.

Protecting Responses Unless in full collusion, namely, all $n-1$ vehicles colluding again the remaining one, the adversary will not be able to infer the individual response of a particular vehicle taking part in providing feedback.

Malicious Model The vehicles can be malicious. In a malicious model, the participant tries not only to cheat the system by providing out-of-range prescribe value but also to infer the response score. We assume that Tally System is honest but curious in a sense that it does not alter the provided input data. We assume users are honest in the sense that if they agree to provide feedback then they will not refrain from it by not submitting the cryptograms.

4.3 Notations and Cryptographic Approach

A homomorphic encryption system allows parties to compute mathematical functions over the encrypted data. A homomorphic encryption is a public-key cryptosystem that

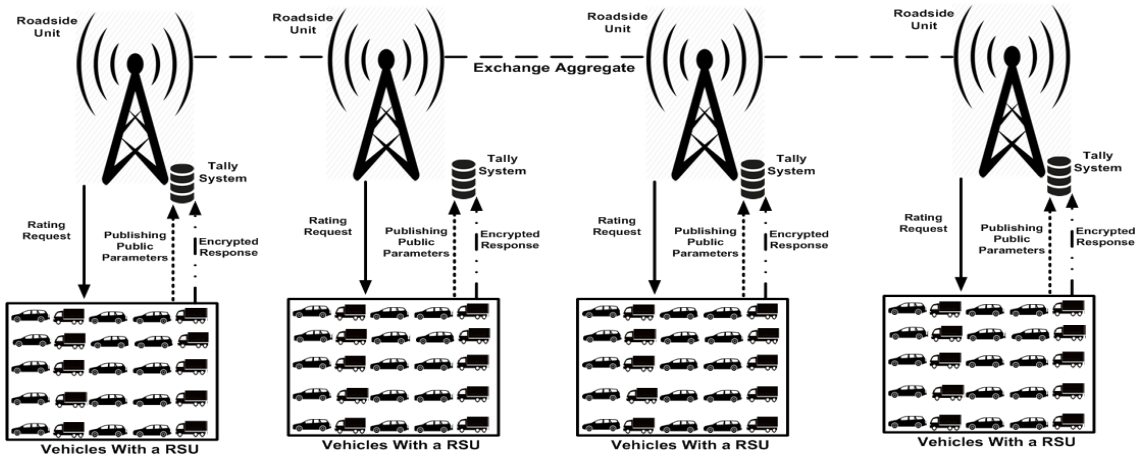


Fig. 2: Building Block of TrustVote System.

Symbols	Description
RSU	Roadside Unit
G	cyclic Group of p elements in which DDH problem is hard
$U_1, U_2, \ldots, U_u$	Set of vehicles registered in a roadside unit
(x_{1i}, x_{2i})	Secret key of U_i
$(g^{y_{1i}}, g^{y_{2i}})$	restructured key of U_i
$(g^{x_{1i}}, g^{x_{2i}})$	public key of U_i
s_i	Secret score of vehicle U_i
w_i	weight assigned to the score of $U_i, 1 \leq w_i \leq a$
α_i	random element generated by each U_i for generating the encrypted feedback
a	max. weight assigned to the scores of vehicles
θ_{1i}, θ_{2i}	First auxiliary variables of U_i
δ_{1i}, δ_{2i}	Second auxiliary variables of U_i
NIZK	non-interactive zero knowledge
$[n]$	the set $\{1, 2, \ldots, n\}$

TABLE 1: Symbols and abbreviations used in the TrustVote System.

allows computation to be performed over the ciphertext, which matches the computation performed over the corresponding plaintext. i.e., $Enc(a) * Enc(b) = Enc(a \oplus b)$. The system has application in a number of domains, for example electronic voting [16], [33], statistical data analysis [34], [35], secure reputation aggregation [36], [37]. In these settings, if we have some n number of parties, say $p_1, p_2 ... p_n$, each with a private input $x_1, x_2, \ldots, x_n$, respectively. The parties would like to compute a collaborative mathematical function over their inputs, say $f(x_1, x_2, \ldots, x_n)$ without revealing their individual inputs to anyone else. The homomorphic cryptosystem consists of three major algorithms: Key generation – responsible for generating the public and private keys, Encryption – responsible for generating the ciphertext corresponding to the plaintext, and the Decryption – responsible for deciphering the result from the encrypted output. In this paper, we consider the additively homomorphic encryption system as we are only aggregating the feedback values from the vehicles.

The cryptographic primitives of TrustVote use the feedback randomization technique proposed for the decentralized and verifiable electronic voting [16], [38], but we adopt the technique for our scenarios by adding weights to each vote. Let $U = \{1, 2, \ldots, n\}$ be the set of vehicles registered with the certain RSU and they provide information about the events to other vehicles in the network. The vehicle rates others on a scale of 0 and 1. The value 1 represents that the vehicle is trusted and 0 represents that it is not trusted. Let there be a DSA/ECDSA-like multiplicative cyclic group where p and q are large primes that satisfy $q \mid p - 1$. Let there be a subgroup $\mathbb{G}_q$ of order q of the group $\mathbb{Z}_p^*$, and g is a generator of $\mathbb{G}_q$. In order to provide the feedback for the vehicles, the user first generates a random value (private key) $sk \in \mathbb{Z}_q$, and generates the public key pk. The public key and a proof of knowledge of the private key are posted at the tally system. The public key is computed as follows:

$$pk_i = g^{sk_i}$$

When all the registered vehicles have generated and published their public keys on the tally server, the encryption key (Y_i) for the vehicle can be computed as:

$$Y_i = \prod_{j \in N, j < i} pk_j \Big/ \prod_{j \in N, j > i} pk_j$$

The computation of Y_i as above ensures that

$$\prod_{i \in N} Y_i^{sk_i} = 1. \quad (1)$$

This property is crucial to the design of the proposed system. Anyone in the system is able to compute Y_i based on the published pk_i keys. This randomization technique was originally intended for encrypting votes in decentralized e-voting [16]. However, we shall be using it in a different context as well as making the aggregation performed over *weighted* votes to suit our new context.

4.4 Assumptions

We assume a DSA/ECDSA-like multiplicative G cyclic group, of prime order q. We also assume that the following cryptographic assumptions hold in G.

Assumption 1. **DDH assumption:** Given g, g^a, g^b and a challenge $\Omega \in \{g^{ab}, R\}$, where $R \xleftarrow{\$} G$, it is computationally hard to find whether $\Omega = g^{ab}$ or $\Omega = R$.

Assumption 2. Given g, g^a, g^b and a challenge $\Omega \in \{g^{ab}, g^{ab}g^a\}$, it is computationally hard to find whether $\Omega = g^{ab}$ or $\Omega = g^{ab}g^a$.

Assumption 3. Given $g, g^a, g^b \in G, t \in \mathbb{Z}_q$ and a challenge $\Omega \in \{g^{ab}, g^{ab}g^t\}$, it is computationally hard to find whether $\Omega = g^{ab}$ or $\Omega = g^{ab}g^t$.

Lemma 1. Assumption 1 implies Assumption 2.

Proof 1. According to the DDH assumption, $(g, g^a, g^b, g^{ab}) \overset{c}{\approx} (g, g^a, g^b, R)$, where $R \xleftarrow{\$} G$. Now, $(g, g^a, g^b, R) \overset{c}{\approx} (g, g^a, g^b, R * g^a) \overset{c}{\approx} (g, g^a, g^b, g^{ab}g^a)$. Hence, the DDH assumption implies Assumption 2.

Lemma 2. Assumption 1 implies Assumption 3.

Proof 2. According to the DDH assumption, $(g, g^a, g^b, g^{ab}) \overset{c}{\approx} (g, g^a, g^b, R)$, where $R \xleftarrow{\$} G$. Now, $(g, g^a, g^b, R) \overset{c}{\approx} (g, g^a, g^b, R * g^t) \overset{c}{\approx} (g, g^a, g^b, g^{ab}g^t)$. Hence, the DDH assumption implies Assumption 3.

Proposition 1. Let, $x_1, x_2, \ldots, x_n \in \mathbb{Z}_q^n$. Then $\sum_{i=1}^{n}(\sum_{j=1}^{i-1} x_j - \sum_{j=i+1}^{n} x_j)x_i = 0$.

proof:

$$\begin{gathered}\sum_{i=1}^{n}(\sum_{j=1}^{i-1} x_j - \sum_{j=i+1}^{n} x_j)x_i \\ = \sum_{i=1}^{n}\sum_{j=1}^{i-1} x_i x_j - \sum_{i=1}^{n}\sum_{j=i+1}^{n} x_i x_j \\ = \sum_{i=1}^{n}\sum_{j<i} x_i x_j - \sum_{i=1}^{n}\sum_{j>i} x_i x_j \\ = \sum_{j=1}^{n}\sum_{j>i} x_i x_j - \sum_{i=1}^{n}\sum_{j>i} x_i x_j.\end{gathered}$$

as, $\sum_{j=1}^{n}\sum_{j>i} x_i x_j = \sum_{j=1}^{n}(x_1 + x_2 + \ldots + x_{j-1})x_j$

$$\begin{gathered}= x_1\sum_{j=2}^{n} x_j + x_2\sum_{j=3}^{n} x_j + \cdots + x_{i-1}\sum_{j=i+1}^{n} x_j + \cdots + \\ x_{n-1}x_n \\ = \sum_{i=1}^{n}\sum_{j>i} x_i x_j\end{gathered}$$

Hence, $\sum_{i=1}^{n}(\sum_{j=1}^{i-1} x_j - \sum_{j=i+1}^{n} x_j)x_i = 0$.

5 TrustVote: System Architecture and Protocol Operations

In this section, we present the system architecture and protocol operations of TrustVote system.

5.1 System Architecture

Figure 2 represents the block diagram of the TrustVote system. The system consists of three major parties: a vehicle that receives messages from other vehicles and rates them on the basis of trustworthiness of the received messages, the RSU that provides links between vehicles, and the tally system which holds cryptographic parameters, the encrypted feedback provided by the participating vehicles and associated zero-knowledge proofs to prove the well-formedness of the encrypted feedback scores. The vehicles communicate with each other and to RSU through the onboard wireless channel. The RSU communicates with each other through a dedicated wireless or wired channel. The tally system is the dynamic database that is placed at the RSU. To use the functionalities of the TrustVote system, the RSU asks vehicles to provide their feedback about the validity and authenticity of the messages they have received from other vehicles. The vehicle responds to the RSU request by presenting the encrypted feedback to the public tally system, which is available to all other parties in the system. Along with the encrypted feedback, the vehicle also transmits its identity to the RSU. The RSUs are interconnected and exchange aggregate scores of vehicles to each other. Further, the RSU maintains the weighted aggregated reputation of the vehicles. Every time when the vehicle receives a message from another vehicle, it first asks the serving RSU for the updated trustworthiness of the vehicle before making any decision on the message.

The TrustVote system is infrastructure-independent, meaning that the implementation is not directly related to the physical network connections. Practically, the TrustVote system is implemented using the standard existing communication practices used in the deployment of RSU and vehicular networks. The TrustVote system can also utilize the current infrastructure and computing devices used within the vehicular network without changing or placing any new system. As such, the TrustVote system is designed to be flexible in terms of the number of vehicles connected to the particular RSU and is capable of being highly distributed. Particularly, the deployment of RSUs in the vehicular network is expensive and difficult to manage. Due to this, we recommend the trade-off between full connectivity through RSUs and their deployment cost. Note that our system does not require any special rule for the placement of RSUs and can be integrated with the current deployment practices of RSUs [39]. To facilitate the communication for the large footprint of RSU and high mobility speed, technologies like Cellular networks (3G, 4G, WiMAX, and LTE) can be used for the communication between RSU and vehicles.

5.2 Protocol Operations of TrustVote System

We now explain the core functionality of the TrustVote system. The system enables Roadside Unit (RSU) to ask a set of trusted and non-trusted vehicles to provide their views about other vehicles in the network. The participating vehicle then responds with the encrypted response to the tally system. The feedback process consists of two steps. 1) The vehicle generates public and private keys, keeps private key secret and shares the public key to the tally system along with a proof of knowledge of the private key. 2) The vehicle computes a restructured key and encrypts its responses with the restructured key and the private key. The vehicles can also get the final aggregate score of any vehicle from the RSU.

Let G be a multiplicative group of p elements, where p is a prime number. The Decisional Diffie Hellman problem is intractable in the group G. Let g be a random generator of G. The protocol consists of the following steps.

5.2.1 Protocol Setup

The roadside unit RSU selects two integers $\omega_1, \omega_2 \in \mathbb{Z}_q$ and computes two variable $\sigma_1 = g^{\omega_1}$ and $\sigma_2 = g^{\omega_2}$. The RSU

posts these variable on the tally system. The RSU also posts non-interactive zero-knowledge proofs of $PW[\omega_1 : \sigma_1]$ and $PW[\omega_2 : \sigma_2]$ on the tally system. The $PW[\omega_i : \sigma_i]$ provides a proof that the RSU knows the value of a ω_i, such that $\sigma_i = g^{\omega_i}$, for all $i \in \{1, 2\}$. The RSU then broadcast its query to the selected vehicles in the network.

The vehicle $U_i; i \in [1, n]$ then selects two random integers $a_{1i}, b_{1i} \in \mathbb{Z}_q$ and computes $\theta_{1i} = g^{a_{1i}}$ and $\delta_{1i} = g^{b_{1i}}$. The vehicle U_i posts θ_{1i} and δ_{1i} on the tally server along with respective NIZK proofs $PW[\theta_{1i} : g]$ and $PW[\delta_{1i} : g]$ of knowledge of θ_{1i} and δ_{1i}.

The RSU computes $\theta_{2i} = (g^{w_i}/(\theta_{1i})^{\omega_1})^{1/\omega_2}$ and $\delta_{2i} = 1/(\delta_{1i}^{\omega_1})^{1/\omega_2}$ for all $i \in [1, n]$. The RSU also computes two non-interactive zero knowledge proofs $PW[\theta_{1i}, \theta_{2i}, \sigma_1, \sigma_2, g]$ and $PW[\delta_{1i}, \delta_{2i}, \sigma_1, \sigma_2, g]$ for all $i \in \{1, 2, \ldots, n\}$. The first NIZK proof proves the fact that $\theta_{1i}^{\omega_1}\theta_{2i}^{\omega_2} \in \{1, g, g^2, \ldots, g^a\}$ without revealing the values of ω_1 or ω_2. The second NIZK proof proves the fact that $\theta_{1i}^{\omega_1}\theta_{2i}^{\omega_2} = 1$. Details of the construction of these two proofs have been elaborated in the Appendix section. The RSU posts θ_{2i}, δ_{2i}, $PW[\theta_{1i}, \theta_{2i}, \sigma_1, \sigma_2, g]$ and $PW[\delta_{1i}, \delta_{2i}, \sigma_1, \sigma_2, g]$ on the tally system, for all $i \in \{1, 2, \ldots, n\}$.

5.2.2 Sharing Cryptographic Parameters

This phase consists of two steps. 1) Vehicles generate cryptographic keys, and 2) creates cryptogram of response. The vehicle $U_i, i \in [1, n]$, selects random values $x_{1i}, x_{2i} \in \mathbb{Z}_q$ as the private key, and publishes public key $pk_i = (X_{1i}, X_{2i}) = (g^{x_{1i}}, g^{x_{2i}})$ on the tally system. The x_{1i}, x_{2i} remains secret to the vehicle. The vehicle U_i generates separate proofs of knowledge for the knowledge of x_{1i} and x_{2i}. These are denoted as $PW_i[x_{ji} : X_{ji}]$ for $j = 1, 2$. Note that the construction of these two NIZK proofs is same as $PW[\omega_1 : \sigma_1]$. The vehicle then computes the restructured key (Y_{1i}, Y_{2i}), where $Y_{ji} = g^{y_{ji}} = g^{\sum_{k=1}^{i-1} x^{jk} - \sum_{k=i+1}^{n} x^{jk}} = \frac{\prod_{k=1}^{i-1} g^{x_{jk}}}{\prod_{k=i+1}^{n} g^{x_{jk}}}, \forall j = 1, 2.$

5.2.3 Reporting Encrypted Responses

The vehicle $U_i, i \in [1, n]$ selects a random $\alpha_i \in \mathbb{Z}_q$ and computes the cryptogram of feedback $\mathbf{c}_i = (B_{1i}, B_{2i}, A_i)$ as follows:

$$B_{1i} = Y_{1i}^{x_{1i}}(\theta_{1i})^{s_i}(\delta_{1i})^{\alpha_i} \tag{2}$$

$$A_i = g^{\alpha_i} \tag{3}$$

$$B_{2i} = Y_{2i}^{x_{2i}}(\theta_{2i})^{s_i}(\delta_{2i})^{\alpha_i} \tag{4}$$

where, $s_i \in \{0, 1\}$ is the secret score of U_i. The vehicle $U_i, i \in [1, n]$ then constructs a NIZK proof

$$PW_i[B_{1i}, B_{2i} : X_{1i}, X_{2i}, Y_{1i}, Y_{2i}, \theta_{1i}, \theta_{2i}, \delta_{1i}, \delta_{2i}, A_i]$$

The cryptogram of feedback and NIZK proof is then posted on the tally system.

5.2.4 Well-formedness of Responses

Each participating vehicle $U_i, i \in [1, n]$ constructs a NIZK proof

$$PW_i[B_{1i}, B_{2i} : X_{1i}, X_{2i}, Y_{1i}, Y_{2i}, \theta_{1i}, \theta_{2i}, \delta_{1i}, \delta_{2i}, A_i]$$

This NIZK proof proves the well-formedness of B_{ji} for $j = 1, 2$ given $X_{ji} = g^{x_{ji}}, Y_{ji} = g^{y_{ji}}, \theta_{ji}, \delta_{ji}, A_i = g^{\alpha_i}$ and $s_i \in \{0, 1\}$. The proof convinces everyone that the encrypted feedback represents either a zero score or a one score. In other words, it proves that the following statement σ is correct: $\sigma \equiv ((B_{1i} = Y_{1i}^{x_{1i}}\delta_{1i}^{\alpha_i}) \wedge (B_{2i} = Y_{2i}^{x_{2i}}\delta_{2i}^{\alpha_i})) \vee ((B_{1i} = Y_{1i}^{x_{1i}}\theta_{1i}\delta_{1i}^{\alpha_i}) \wedge (B_{2i} = Y_{2i}^{x_{2i}}\theta_{2i}\delta_{2i}^{\alpha_i}))$. Here, the secret inputs of the vehicle U_i are x_{1i}, x_{2i}, α_i, and the publicly known variables are $B_{1i}, B_{2i}, Y_{1i}, Y_{2i}, \theta_{1i}, \theta_{2i}, \delta_{1i}, \delta_{2i}, A_i = g^{\alpha_i}$. The NIZK proof is constructed using standard Σ protocol and then it is made non-interactive by means of the Fiat-Shamir heuristic. A detailed construction of this NIZK proof can be found in the Appendix.

5.2.5 Tallying Final Score

In the final step, the RSU accesses data from the tally system and checks its well-formedness proofs. Given the well-formed cryptograms $C = (C_1, C_2)$ for $j = 1, 2$ from the tally server, the RSU executes following steps while computing the final score.

$$C_j = \prod_{i=1}^{n} B_{ji} \tag{5}$$

$$= \prod_{i=1}^{n} Y_{ji}^{x_{ji}}\theta_{ji}^{s_i}\delta_{ji}^{\alpha_i} \tag{6}$$

$$= \prod_{i=1}^{n} g^{x_{ji}y_{ji}}\theta_{ji}^{s_i}\delta_{ji}^{\alpha_i} \tag{7}$$

$$= g^{\sum_{i=1}^{n} x_{ji}y_{ji}} \prod_{i=1}^{n} \theta_{ji}^{s_i}\delta_{ji}^{\alpha_i} \tag{8}$$

From proposition 1, we can see that $\sum_{i=1}^{n} x_{ji}y_{ji} = 0$. Thus, $C_j = \prod_{i=1}^{n} \theta_{ji}^{s_i}\delta_{ji}^{\alpha_i}$ for $j = 1, 2$. Note that this step can be executed by anyone with read access to the tally system. The RSU then computes $L = C_1^{\omega_1}C_2^{\omega_2}$ and posts it on the tally server along with the NIZK proof $PW_L[L : C_1, C_2, \sigma_1, \sigma_2]$. This NIZK proof proves the fact that the L posted on the tally server is indeed equal to $C_1^{\omega_1}C_2^{\omega_2}$ given C_1, C_2, σ_1 and σ_2. Note that $L = \prod_{i=1}^{n}(\theta_{1i}^{s_i}\delta_{1i}^{\alpha_i})^{\omega_1}(\theta_{2i}^{s_i}\delta_{2i}^{\alpha_i})^{\omega_2} = \prod_{i=1}^{n}(\theta_{1i}^{\omega_1}\theta_{2i}^{\omega_2})^{s_i}(\delta_{1i}^{\omega_1}\delta_{2i}^{\omega_2})^{\alpha_i}$. Now, $\theta_{1i}^{\omega_1}\theta_{2i}^{\omega_2} = g^{w_i}$ and $\delta_{1i}^{\omega_1}\delta_{2i}^{\omega_2} = 1$ for all $i \in [1, n]$. Hence, $L = \prod_{i=1}^{n} g^{w_i s_i} = g^{\sum_{i=1} w_i s_i} = g^S$. The service provider then performs a brute force search on L to find final reputation score S. Brute force search will be feasible only if S is not too large. The maximum value that S can take is $a * n$.

Once the weighted sum S has been computed, the RSU can calculate the average weighted reputation as follows:

$$\tilde{S} = S/\sum_{i=1}^{n} w_i \tag{9}$$

Once the average reputation of the vehicle is computed, the RSU lists it as the malicious actor if the aggregate reputation

of the vehicle is less than some threshold. The simplest approach to consider the vehicle as malicious is to apply the fixed or automated threshold. Alternatively, the RSU can also use machine learning methods for this purpose. The vehicles can also query the RSU for the final aggregated score of any vehicle it receives messages before taking any action.

6 Security and Privacy Analysis

In this section, we discuss the security properties of our scheme. We assume that all participants in the aggregation process, including participating vehicles and roadside units may be malicious. The NIZK proofs ensure that they strictly adhere to the protocol specification. They are forced to follow the protocol operations honestly but they might try to infer information from the shared data. We also assume that the participating vehicle can be malicious in providing their feedback scores, however, we ensure they provide input within the correct range via the use of NIZK proofs. The RSU and the tally system only see the cryptograms of vehicles without having access to the private key of vehicles.

Further, we consider an adversary who, in collusion with a subset of vehicles, attempts to learn the score of one or more honest vehicle. We show that the adversary will not learn anything which the weighted sum $S = \sum_{i=1}^{n} w_i s_i$ does not allow her to know. The security of this scheme depends upon the intractability of the Decisional Diffie Hellman problem. Hence, as long as the DDH problem is hard in the group G, the scheme is secure.

6.1 Correctness of Computation

Here, we show that the scheme described in Section 5 is correct. The feedback each vehicle $U_i : i \in [n]$ sends to the tally system (i.e., a public bulletin board) is of the form $(B_{1i}, B_{2i}, g^{\alpha_i})$, where $B_{ji} = Y_{ji}^{x_{ji}} \theta_{ji}^{s_i} \delta_{ji}^{\alpha_i} = g^{x_{ji} y_{ji}} \theta_{ji}^{s_i} \delta^{\alpha_i} : j = 1, 2$. The RSU downloads all such feedbacks from the tally system and computes $C = (C_1, C_2)$. $C_j = \prod_{i=1}^{n} B_{ji} = g^{x_{ji} y_{ji}} \theta_{ji}^{s_i} \delta_{ji}^{\alpha_i} : j \in \{1, 2\}$. From Proposition 1, we can see that $\sum_{i=1}^{n} x_{ji} y_{ji} = 0$ for $j = 1, 2$. This implies that $C_j = \prod_{i=1}^{n} \theta_{ji}^{s_i} \delta_{ji}^{\alpha_i}$. We know that $\theta_{ji} = g^{a_{ji}}$ and $\delta_{ji} = g^{b_{ji}}$ for $j = 1, 2$. Hence, $C_j = \prod_{i=1}^{n} g^{a_{ji} s_i} g^{b_{ji} \alpha_i}$. Now, $L = C_1^{\omega_1} C_2^{\omega_2} = \prod_{i=1}^{n} g^{(\omega_1 a_{1i} + \omega_2 a_{2i}) s_i} * \prod_{i=1}^{n} g^{(\omega_1 b_{1i} + \omega_2 b_{2i}) \alpha_i}$. a_{2i} and b_{2i}s are made to satisfy $\omega_1 a_{1i} + \omega_2 a_{2i} = w_i$ and $\omega_1 b_{1i} + \omega_2 b_{2i} = 0$ hold. Hence, $L = \prod_{i=1}^{n} g^{w_i s_i} = g^{\sum_{i=1}^{n} w_i s_i} = g^S$. Therefore, the scheme is correct.

6.2 Privacy of Trust Weights

The secret inputs of the RSU are the set of weights $\{w_1, w_2, \ldots, w_n\}$. Each w_i is used to compute $g^{a_{2i}} = (g^{w_i} / g^{a_{1i} \omega_1})^{1/\omega_2}$. We need to show that $g^{a_{2i}}$ will not reveal the value of w_i. $g^{a_{2i} \omega_2} = (g^{w_i} / g^{a_{1i} \omega_1})$. Let us assume that there is an adversary $\mathcal{A}$ that can distinguish between the two cases $w_i = w$ and $w_i = w'$, where $w' > w$. We show that $\mathcal{A}$ can be used against Assumption 3. Let the inputs to the DDH adversary be $t, g^{\omega_2}, g^{a_{2i}}$ and a challenge $\Omega \in \{g^{a_{2i} \omega_2}, g^{a_{2i} \omega_2} g^t\}$. Here $t = w' - w$. $\mathcal{A}$ selects random $a_{1i} \in \mathbb{Z}_q$ and computes $g^{\omega_1} = (g^{w'} / \Omega)^{1/a_{1i}}$. Thus, if $\Omega = g^{a_{2i} \omega_2}$, then $g^{a_{1i} \omega_1 + a_{2i} \omega_2} = g^{w'}$ holds. Alternatively if $\Omega = g^{a_{2i} \omega_2} g^{w' - w}$, then $g^{a_{1i} \omega_1 + a_{2i} \omega_2} = g^{w}$ holds. Hence, if $\mathcal{A}$ can distinguish between these two cases, it will amount to distinguishing between the two possible values of Ω viz. $g^{a_{2i} \omega_2}$ and $g^{a_{2i} \omega_2} g^{w' - w}$. Thus, we can say that the weights assigned by the RSU to all the vehicles will remain secret.

6.3 Privacy of Vehicles

The participating vehicles simply provide their feedback in the encrypted form using the encryption key derived from the public keys of all vehicles. The vehicles send encrypted responses to the tally server. The sent responses cannot be decrypted individually and no vehicle would learn the values of other vehicles even if some vehicles collude with each other. The malicious vehicle could report out-of-range feedback values in order to modify the final aggregated values, however, the use of non-interactive zero-knowledge proof of knowledge limits them to provide their inputs within the prescribed range.

In Lemma 7, we assume there are some k distinct honest vehicles identified as $P_{h_1}, P_{h_2}, \ldots, P_{h_k}$, $h_t \in [1, n]$ for all $t = 1, 2, \ldots, k$. The weights assigned to P_{h_t} are w_{h_t} for all $h_t \in [1, n]$ for all $t = 1, 2, \ldots, k$. We show that for any two sets of private inputs of the honest vehicles namely $\mathcal{S}' = \{s'_{h_t} : 1 \leq t \leq k\}$ and $\mathcal{S}'' = \{s''_{h_t} : 1 \leq t \leq k\}$ such that $s'_{h_t}, s''_{h_t} \in \{0, 1\}, \forall t \in [1, k]$ and $\sum_{i=1}^{k} w_{h_i} s'_{h_i} = \sum_{i=1}^{k} w_{h_i} s''_{h_i}$, if the k honest vehicles choose either $\mathcal{S}'$ or $\mathcal{S}''$ as the set of score, the RSU will not be able to find whether $\mathcal{S}'$ or $\mathcal{S}''$ was chosen by the honest vehicles. That is, as long as the partial weighted sum of the honest vehicles is the same, the RSU cannot distinguish between two tally systems corresponding to two different executions of this protocol where the honest vehicles have used different inputs. In other words, the protocol only allows the adversary to learn nothing more than the partial weighted sum of the secret scores of the honest vehicles. As stated above, the adversary can trivially compute this partial weighted sum if she colludes with all other participants (RSU and vehicles), excepting these k honest vehicles. Thus, we can conclude that the protocol allows the colluding adversary to know nothing more than what she could trivially learn using the inputs of the colluding vehicles. We used Assumption 4 in order to prove Lemma 7. In Lemma 3 we show that Assumption 4 follows from Assumption 1. In order to prove Lemma 3, we used Lemma 6. In order to prove Lemma 6, we needed Assumption 5 and Lemma 5. In Lemma 4, we show that Assumption 5 follows directly from Assumption 1.

Note that, if only a subset of vehicles colludes, they will learn nothing as the weights assigned to their inputs are only known to the RSU. and without them, they cannot even compute the partial weighted sum of their own inputs. Thus, even if S is made public, the partial weighted sum of the non-colluding vehicles cannot be computed by them.

6.3.1 Assumptions & Lemmas

Assumption 4. Given $g, \mathbb{X}_1 = \{g^{x_{11}}, g^{x_{12}}, \ldots, g^{x_{1(k-1)}}\}, \mathbb{X}_2 = \{g^{x_{21}}, g^{x_{22}}, \ldots, g^{x_{2(k-1)}}\}$, $\mathbb{Y}_1 = \{g^{y_{11}}, g^{y_{12}}, \ldots, g^{y_{1(k-1)}}\}, \mathbb{Y}_2 = \{g^{y_{21}}, g^{y_{22}}, \ldots, g^{y_{2(k-1)}}\}$,$\mathbb{A}_1 = \{g^{a_{11}}, g^{a_{12}}, \ldots, g^{a_{1k}}\}$, $\mathbb{A}_2 = \{g^{a_{21}}, g^{a_{22}}, \ldots, g^{a_{2k}}\}$,$\mathbb{B}_1 = \{g^{b_{11}}, g^{b_{12}}, \ldots, g^{b_{1k}}\}$,

$\mathbb{B}_2 = \{g^{b_{21}}, g^{b_{22}}, \ldots, g^{b_{2k}}\}$,$\omega_1, \omega_2$, such that $a_{1i}\omega_1 + a_{2i}\omega_2 = w_i, \forall i \in [1, k-1]$ and $b_{1i}\omega_1 + b_{2i}\omega_2 = 0, \forall i \in [1, k-1]$, it is hard to distinguish between $\Omega_1 = (U_1, U_2)$ and $\Omega_2 = (V_1, V_2)$, where $U_1 = (l_{11}, l_{12}, \ldots, l_{1k})$,$U_2 = (l_{21}, l_{22}, \ldots, l_{2k})$, $V_1 = (l'_{11}, l'_{12}, \ldots, l'_{1k})$,$V_2 = (l'_{21}, l'_{22}, \ldots, l'_{2k})$, $l_{ji} = g^{x_{ji}y_{ji}} g^{a_{ji}s_i} g^{b_{ji}\alpha_i}, \forall j = 1, 2; i \in [1, k-1]$, $l'_{ji} = g^{x_{ji}y_{ji}} g^{a_{ji}s'_i} g^{b_{ji}\alpha_i}, \forall j = 1, 2; i \in [1, k-1]$, $l_{jk} = \frac{g^{a_{jk}s_k} g^{b_{jk}\alpha_k}}{\prod_{i=1}^{k-1} g^{x_{ji}y_{ji}}}$, $l'_{jk} = \frac{g^{a_{jk}s'_k} g^{b_{jk}\alpha_k}}{\prod_{i=1}^{k-1} g^{x_{ji}y_{ji}}}, \forall j = 1, 2$ and $\sum_{i=1}^{k} w_i s_i = \sum_{i=1}^{k} w_i s'_i$.

Lemma 3. Assumption 1 implies assumption 4.

Proof 3. $l_{1i} = g^{x_{1i}y_{1i}} g^{a_{1i}s_i} g^{b_{1i}\alpha_i}$ for all $i \in [1, k-1]$ and $l_{1k} = \frac{g^{a_{1k}s_k} g^{b_{1k}\alpha_k}}{\prod_{i=1}^{k-1} g^{x_{1i}y_{1i}}}$
$l'_{1i} = g^{x_{1i}y_{1i}} g^{a_{1i}s'_i} g^{b_{1i}\alpha_i}, \forall i \in [1, k-1]$ and $l'_{1k} = \frac{g^{a_{1k}s'_k} g^{b_{1k}\alpha_k}}{\prod_{i=1}^{k-1} g^{x_{1i}y_{1i}}}$.
Again, $l_{1i} = g^{x_{1i}y_{1i}} g^{a_{1i}s_i} g^{b_{1i}\alpha_i}$ for all $i \in [1, k-1]$ and $l_{1k} = \frac{g^{a_{1k}s_k} g^{b_{1k}\alpha_k}}{\prod_{i=1}^{k-1} g^{x_{1i}y_{1i}}}$
$l'_{1i} = g^{x_{1i}y_{1i}} g^{a_{1i}s'_i} g^{b_{1i}\alpha_i}, \forall i \in [1, k-1]$ and $l'_{1k} = \frac{g^{a_{1k}s'_k} g^{b_{1k}\alpha_k}}{\prod_{i=1}^{k-1} g^{x_{1i}y_{1i}}}$
Similarly, $l_{2i} = g^{x_{2i}y_{2i}} g^{a_{2i}s_i} g^{b_{2i}\alpha_i}$ for all $i \in [1, k-1]$ and $l_{2k} = \frac{g^{a_{2k}s_k} g^{b_{2k}\alpha_k}}{\prod_{i=1}^{k-1} g^{x_{2i}y_{2i}}}$
$l'_{2i} = g^{x_{2i}y_{2i}} g^{a_{2i}s'_i} g^{b_{2i}\alpha_i}, \forall i \in [1, k-1]$ and $l'_{2k} = \frac{g^{a_{2k}s'_k} g^{b_{2k}\alpha_k}}{\prod_{i=1}^{k-1} g^{x_{2i}y_{2i}}}$.
Again, $l_{2i} = g^{x_{2i}y_{2i}} g^{a_{2i}s_i} g^{b_{2i}\alpha_i}$ for all $i \in [1, k-1]$ and $l_{2k} = \frac{g^{a_{2k}s_k} g^{b_{2k}\alpha_k}}{\prod_{i=1}^{k-1} g^{x_{2i}y_{2i}}}$
$l'_{2i} = g^{x_{2i}y_{2i}} g^{a_{2i}s'_i} g^{b_{2i}\alpha_i}, \forall i \in [1, k-1]$ and $l'_{2k} = \frac{g^{a_{2k}s'_k} g^{b_{12}\alpha_k}}{\prod_{i=1}^{k-1} g^{x_{2i}y_{2i}}}$

$\prod_{i=1}^{k} l_{ji} = g^{\sum_{i=1}^{k} a_{ji}s_i} g^{\sum_{i=1}^{k} b_{ji}\alpha_i}$ for all $j \in \{1, 2\}$. Similarly, $\prod_{i=1}^{k} l'_{ji} = g^{\sum_{i=1}^{k} a_{ji}s'_i} g^{\sum_{i=1}^{k} b_{ji}\alpha_i}$ for all $j \in \{1, 2\}$. Let us choose a random $r \in [1, k]$. Now, let $K_j = \prod_{i=1}^{k} l_{ji}$ and $K'_j = \prod_{i=1}^{k} l_{ji}$ for $j \in \{1, 2\}$. So, $K_j = \prod_{i=1}^{k-1} l_{ji} * g^{b_{jk}\alpha_k}$ and $K'_j = \prod_{i=1}^{k-1} l'_{ji} * g^{b_{jk}\alpha_k}$ for all $j \in \{1, 2\}$.
We know that
$K_1^{\omega_1} K_2^{\omega_2} = (g^{\sum_{i=1}^{k} a_{1i}s_i} g^{\sum_{i=1}^{k} b_{1i}\alpha_i})^{\omega_1} (g^{\sum_{i=1}^{k} a_{2i}s_i} g^{\sum_{i=1}^{k} b_{2i}\alpha_i})^{\omega_2} = g^{\sum_{i=1}^{k} w_i s_i}$ and
$K'^{\omega_1}_1 K'^{\omega_2}_2 = (g^{\sum_{i=1}^{k} a_{1i}s'_i} g^{\sum_{i=1}^{k} b_{1i}\alpha_i})^{\omega_1} (g^{\sum_{i=1}^{k} a_{2i}s'_i} g^{\sum_{i=1}^{k} b_{2i}\alpha_i})^{\omega_2} = g^{\sum_{i=1}^{k} w_i s'_i}$. Since, $\sum_{i=1}^{k} w_i s_i = \sum_{i=1}^{k} w_i s'_i$,
$K_1^{\omega_1} K_2^{\omega_2} = (g^{\sum_{i=1}^{k} a_{1i}s_i} g^{\sum_{i=1}^{k} b_{1i}\alpha_i})^{\omega_1} (g^{\sum_{i=1}^{k} a_{2i}s_i} g^{\sum_{i=1}^{k} b_{2i}\alpha_i})^{\omega_2} = (g^{\sum_{i=1}^{k} a_{1i}s'_i} g^{\sum_{i=1}^{k} b_{1i}\alpha_i})^{\omega_1} (g^{\sum_{i=1}^{k} a_{2i}s'_i} g^{\sum_{i=1}^{k} b_{2i}\alpha_i})^{\omega_2} = K'^{\omega_1}_1 K'^{\omega_2}_2$. Hence from Lemma 6, we may say that if $K_1 \overset{c}{\approx} K'_1$, then $(K_1, K_2) \overset{c}{\approx} (K'_1, K'_2)$ Hence, $(U_1, U_2) \overset{c}{\approx} (V_1, V_2)$ or $\Omega_1 \overset{c}{\approx} \Omega_2$.

Assumption 5. Let, $x = (g^{a_1}, g^{a_2}, \ldots, g^{a_{k-1}})$ and $y = (g^{b_1}, g^{b_2}, \ldots, g^{b_{k-1}})$. Let, $X = (g^{a_1 b_1}, g^{a_2 b_2}, \ldots, g^{a_{k-1} b_{k-1}})$. Also let, $\mathcal{X} = \mathcal{X}_1 * \mathcal{X}_2 * \ldots * \mathcal{X}_k$ and $\mathcal{Y} = \mathcal{Y}_1 * \mathcal{Y}_2 * \ldots * \mathcal{Y}_k$. If $\mathcal{X} \overset{c}{\approx} \mathcal{Y}$, then $\left(g^{a_1 b_1} * \mathcal{X}_1, g^{a_2 b_2} * \mathcal{X}_2, \ldots, g^{a_{k-1} b_{k-1}} * \mathcal{X}_{k-1}, \frac{\mathcal{X}_k}{\prod_{i=1}^{k-1} g^{a_i b_i}}\right) \overset{c}{\approx} \left(g^{a_1 b_1} * \mathcal{Y}_1, g^{a_2 b_2} * \mathcal{Y}_2, \ldots, g^{a_{k-1} b_{k-1}} * \mathcal{Y}_{k-1}, \frac{\mathcal{Y}_k}{\prod_{i=1}^{k-1} g^{a_i b_i}}\right)$.

Lemma 4. Assumption 1 implies assumption 5.

Proof 4. Let $A_i = g^{a_i b_i}$ for $i \in [1, k-1]$. $(A_1 * \mathcal{X}_1, A_2 * \mathcal{X}_2, \ldots, A_{k-1} * \mathcal{X}_{k-1}, \frac{\mathcal{X}_k}{\prod_{i=1}^{k-1} A_i}) = (A_1 * \mathcal{X}_1, A_2 * \mathcal{X}_2, \ldots, A_{k-1} * \mathcal{X}_{k-1}, \frac{\mathcal{X}}{\prod_{i=1}^{k-1} A_i * \mathcal{X}_i})$. Since, according to assumption 1; $A_i = g^{a_i b_i} \overset{c}{\approx} R, \forall i \in [1, k-1]$, $A_i * \mathcal{X}_i \overset{c}{\approx} A_i * \mathcal{Y}_i, \forall i \in [1, k-1]$. $\left(A_1 * \mathcal{X}_1, A_2 * \mathcal{X}_2, \ldots, A_{k-1} * \mathcal{X}_{k-1}, \frac{\mathcal{X}}{\prod_{i=1}^{k-1} A_i * \mathcal{X}_i}\right) \overset{c}{\approx} \left(A_1 * \mathcal{Y}_1, A_2 * \mathcal{Y}_2, \ldots, A_{k-1} * \mathcal{Y}_{k-1}, \frac{\mathcal{X}}{\prod_{i=1}^{k-1} A_i * \mathcal{Y}_i}\right)$. Now, we claim that:
$\left(A_1 * \mathcal{Y}_1, A_2 * \mathcal{Y}_2, \ldots, A_{k-1} * \mathcal{Y}_{k-1}, \frac{\mathcal{X}}{\prod_{i=1}^{k-1} A_i * \mathcal{Y}_i}\right) \overset{c}{\approx} \left(A_1 * \mathcal{Y}_1, A_2 * \mathcal{Y}_2, \ldots, A_{k-1} * \mathcal{Y}_{k-1}, \frac{\mathcal{Y}}{\prod_{i=1}^{k-1} A_i * \mathcal{Y}_i}\right)$,
otherwise anyone can distinguish between $\mathcal{X}$ and $\mathcal{Y}$ by choosing random A_is and random $\mathcal{X}_i$'s and thus computing a challenge $\left(A_1 * \mathcal{Y}_1, A_2 * \mathcal{Y}_2, \ldots, A_{k-1} * \mathcal{Y}_{k-1}, \frac{\mathcal{Q}}{\prod_{i=1}^{k-1} A_i * \mathcal{Y}_i}\right)$, where $Q \in \{\mathcal{X}, \mathcal{Y}\}$. If the challenge $\left(A_1 * \mathcal{Y}_1, A_2 * \mathcal{Y}_2, \ldots, A_{k-1} * \mathcal{Y}_{k-1}, \frac{\mathcal{Q}}{\prod_{i=1}^{k-1} A_i * \mathcal{Y}_i}\right)$ is correctly identified then so will be Q. Hence, $\left(A_1 * \mathcal{Y}_1, A_2 * \mathcal{Y}_2, \ldots, A_{k-1} * \mathcal{Y}_{k-1}, \frac{\mathcal{X}}{\prod_{i=1}^{k-1} A_i * \mathcal{Y}_i}\right) \overset{c}{\approx} \left(A_1 * \mathcal{Y}_1, A_2 * \mathcal{Y}_2, \ldots, A_{k-1} * \mathcal{Y}_{k-1}, \frac{\mathcal{Y}}{\prod_{i=1}^{k-1} A_i * \mathcal{Y}_i}\right)$. Thus, the lemma holds.

Lemma 5. Given $X_1 = \mathcal{X}_1 \mathcal{Z}$, $X_2 = \mathcal{X}_2 (\mathcal{Z})^{-\omega_1/\omega_2}$ and $Y_1 = \mathcal{Y}_1 \mathcal{Z}$, $Y_2 = \mathcal{Y}_2 (\mathcal{Z})^{-\omega_1/\omega_2}$. Also assume that $X_1^{\omega_1} X_2^{\omega_2} = Y_1^{\omega_1} Y_2^{\omega_2}$. If $\mathcal{Z} \overset{c}{\approx} R$, then $(X_1, X_2) \overset{c}{\approx} (Y_1, Y_2)$.

Proof 5. $X_2 = \mathcal{X}_2 (\frac{X_1}{\mathcal{X}_1})^{-\omega_1/\omega_2} = (\mathcal{X}_1^{\omega_1} \mathcal{X}_2^{\omega_2})^{1/\omega_2} X_1^{-\omega_1/\omega_2}$.
Similarly, $Y_2 = \mathcal{Y}_2 (\frac{Y_1}{\mathcal{Y}_1})^{-\omega_1/\omega_2} = (\mathcal{Y}_1^{\omega_1} \mathcal{Y}_2^{\omega_2})^{1/\omega_2} Y_1^{-\omega_1/\omega_2}$.
Now, $(X_1, X_2) \overset{c}{\approx} (X_1, (\mathcal{X}_1^{\omega_1} \mathcal{X}_2^{\omega_2})^{1/\omega_2} X_1^{-\omega_1/\omega_2}) = (X_1, (\mathcal{X}_1^{\omega_1} \mathcal{X}_2^{\omega_2})^{1/\omega_2} X_1^{-\omega_1/\omega_2}) = (\mathcal{X}_1 * \mathcal{Z}, (\mathcal{X}_1^{\omega_1} \mathcal{X}_2^{\omega_2})^{1/\omega_2} (\mathcal{X}_1 * \mathcal{Z})^{-\omega_1/\omega_2}) \overset{c}{\approx} (\mathcal{X}_1 * R, (\mathcal{X}_1^{\omega_1} \mathcal{X}_2^{\omega_2})^{1/\omega_2} (\mathcal{X}_1 * R)^{-\omega_1/\omega_2}) \overset{c}{\approx} (R, (\mathcal{X}_1^{\omega_1} \mathcal{X}_2^{\omega_2})^{1/\omega_2} R^{-\omega_1/\omega_2})$. Similarly, $(Y_1, Y_2) \overset{c}{\approx} (R, (\mathcal{Y}_1^{\omega_1} \mathcal{Y}_2^{\omega_2})^{1/\omega_2} R^{-\omega_1/\omega_2}) = (R, (\mathcal{X}_1^{\omega_1} \mathcal{X}_2^{\omega_2})^{1/\omega_2} R^{-\omega_1/\omega_2})$. Thus, $(X_1, X_2) \overset{c}{\approx} (Y_1, Y_2)$.

Lemma 6. Let us assume,

$$X_1 = (g^{x_{11}}, g^{x_{12}}, \ldots, g^{x_{1(k-1)}})$$

$$X_2 = (g^{x_{21}}, g^{x_{22}}, \ldots, g^{x_{2(k-1)}})$$

$$Y_1 = (g^{y_{11}}, g^{y_{12}}, \ldots, g^{y_{1(k-1)}})$$

$$Y_2 = (g^{y_{21}}, g^{y_{22}}, \ldots, g^{y_{2(k-1)}})$$

$$A = (g^{x_{11}y_{11}}, g^{x_{12}y_{12}}, \ldots, g^{x_{1(k-1)}y_{1(k-1)}})$$

$$B = (g^{x_{21}y_{21}}, g^{x_{22}y_{22}}, \ldots, g^{x_{2(k-1)}y_{2(k-1)}})$$

$$X = \mathcal{X} * \mathcal{Z} = \mathcal{X}_1 * \mathcal{X}_2 * \ldots * \mathcal{X}_k$$

$$Y = \mathcal{Y} * (\mathcal{Z})^{-\omega_1/\omega_2} = \mathcal{Y}_1 * \mathcal{Y}_2 * \ldots * \mathcal{Y}_k$$

$$X' = \mathcal{X}' * \mathcal{Z} = \mathcal{X}'_1 * \mathcal{X}'_2 * \ldots * \mathcal{X}'_k$$

$$Y' = \mathcal{Y}' * (\mathcal{Z})^{-\omega_1/\omega_2} = \mathcal{Y}'_1 * \mathcal{Y}'_2 * \ldots * \mathcal{Y}'_k$$

$\exists \omega_1, \omega_2 \in \mathbb{Z}_q$ such that $\mathcal{X}^{\omega_1}\mathcal{Y}^{\omega_2} = \mathcal{X}'^{\omega_1}\mathcal{Y}'^{\omega_2}$ holds. If $\mathcal{Z} \overset{c}{\approx} R$, where $R \overset{\$}{\leftarrow} G$, then $(\psi_1, \psi_1') \overset{c}{\approx} (\psi_2, \psi_2')$, where

$$\psi_1 = \left(A_1 * \mathcal{X}_1, A_2 * \mathcal{X}_2, \ldots, A_{k-1} * \mathcal{X}_{k-1}, \frac{\mathcal{X}_k}{\prod_{i=1}^{k-1} A_i}\right)$$

$$\psi_1' = \left(B_1 * \mathcal{Y}_1, B_2 * \mathcal{Y}_2, \ldots, B_{k-1} * \mathcal{Y}_{k-1}, \frac{\mathcal{Y}_k}{\prod_{i=1}^{k-1} B_i}\right)$$

$$\psi_2 = \left(A_1 * \mathcal{X}_1', A_2 * \mathcal{X}_2', \ldots, A_{k-1} * \mathcal{X}_{k-1}', \frac{\mathcal{X}_k'}{\prod_{i=1}^{k-1} A_i}\right)$$

$$\psi_2' = \left(B_1 * \mathcal{Y}_1', B_2 * \mathcal{Y}_2', \ldots, B_{k-1} * \mathcal{Y}_{k-1}', \frac{\mathcal{Y}_k'}{\prod_{i=1}^{k-1} B_i}\right)$$

and $A_i = g^{x_{1i}y_{1i}}, B_i = g^{x_{2i}y_{2i}}, i \in [1, k-1]$.

Proof 6. According to assump 5, $\psi_1 \overset{c}{\approx} \psi_2$ if $X \overset{c}{\approx} X'$ and $\psi_1' \overset{c}{\approx} \psi_2'$ if $Y \overset{c}{\approx} Y'$. Since, $\mathcal{X}^{\omega_1}\mathcal{X}'^{\omega_2} = \mathcal{Y}^{\omega_1}\mathcal{Y}'^{\omega_2}$. Again, according to Lemma 5, since $\mathcal{Z} \overset{c}{\approx} R$, then $(X, X') \overset{c}{\approx} (Y, Y')$. Hence, $(\omega_1, \omega_2, \psi_1, \psi_2) \overset{c}{\approx} (\omega_1, \omega_2, \psi_1', \psi_2')$. So, the lemma holds.

Lemma 7. In our protocol, if there are some k distinct honest vehicles identified as $P_{h_1}, P_{h_2}, \ldots, P_{h_k}$, $h_t \in [1, n]$ for all $t = 1, 2, \ldots, k$. Let us assume that the weights assigned to P_{h_t} be w_{h_t}. Also assume that there exist the following two sets:
$\mathcal{S}' = \{s'_{h_t} : 1 \leq t \leq k\}$ and $\mathcal{S}'' = \{s''_{h_t} : 1 \leq t \leq k\}$ such that $s'_{h_t}, s''_{h_t} \in \{0, 1\}, \forall t \in [1, k]$ and $\sum_{i=1}^{k} w_{h_i} s'_{h_i} = \sum_{i=1}^{k} w_{h_i} s''_{h_i}$. If the k honest vehicles choose either $\mathcal{S}'$ or $\mathcal{S}''$ as the set of score, the RSU will not be able to find whether $\mathcal{S}'$ or $\mathcal{S}''$ was chosen by the honest vehicles.
That is, the protocol initiator will not be able to distinguish between multiple possible sets of distinct values of inputs of a group of honest vehicles, provided the partial weighted sum of all inputs of those honest vehicles remain the same in all the sets.

Proof 7. We assume that $P_{c_1}, P_{c_2}, \ldots, P_{c_{n-k}}$ be the identifier of the $n-k$ corrupt vehicles. Hence, $\{c_j : 1 \leq j \leq n-k\} \cup \{h_j : 1 \leq j \leq k\} = [1, n]$. Since, the $P_{c_1}, P_{c_2}, \ldots, P_{c_{n-k}}$ are corrupted by the RSU, she can set the keys and scores on their behalf. Let us assume that s_{c_j} is the score chosen by the RSU for the vehicle P_{c_j}, where $j \in [1, n-k]$. We assume that the secret key of P_{c_j} is (x_{1c_j}, x_{2c_j}) and the public key will be $(g^{x_{1c_j}}, g^{x_{2c_j}}), \forall j \in [1, n-k]$. Also, assume that the restructured key of P_{c_j} is $(g^{y_{1c_j}}, g^{y_{2c_j}}), \forall j \in [1, n-k]$. Similarly, we assume that the score of P_{h_j} be s_{h_j} and the secret key of P_{h_j} is (x_{1c_j}, x_{2c_j}) and the public key will be $(g^{x_{1c_j}}, g^{x_{2c_j}}), \forall j \in [1, n-k]$. Also, assume that the restructured key of P_{c_j} is $(g^{y_{1c_j}}, g^{y_{2c_j}}), \forall j \in [1, n-k]$. The RSU can compute the ballot of a corrupt vehicle P_{c_j} as (B_{1c_j}, B_{2c_j}), where $j \in [1, n-k]$. Here, $B_{1c_j} = g^{x_{1c_j}y_{1c_j}}(\theta_{1c_j})^{s_{c_j}}(\delta_{1c_j})^{\alpha_{c_j}} = g^{x_{1c_j}y_{1c_j}} g^{a_{1c_j}s_{c_j}} g^{b_{1c_j}\alpha_{c_j}}$, and $B_{2c_j} = g^{x_{2c_j}y_{2c_j}}(\theta_{2c_j})^{s_{c_j}}(\delta_{2c_j})^{\alpha_{c_j}} = g^{x_{2c_j}y_{2c_j}} g^{a_{2c_j}s_{c_j}} g^{b_{2c_j}\alpha_{c_j}}$.
Similarly, the ballots of the k honest vehicles are of the form (B_{1h_j}, B_{2h_j}), where $j \in [1, k]$. Here,

$$B_{1h_j} = g^{x_{1h_j}y_{1h_j}}(\theta_{1h_j})^{s_{h_j}}(\delta_{1h_j})^{\alpha_{h_j}}$$
$$= g^{x_{1h_j}y_{1h_j}} g^{a_{1h_j}s_{h_j}} g^{b_{1h_j}\alpha_{h_j}}$$
$$B_{2h_j} = g^{x_{2h_j}y_{2h_j}}(\theta_{2h_j})^{s_{h_j}}(\delta_{2h_j})^{\alpha_{h_j}}$$
$$= g^{x_{2h_j}y_{2h_j}} g^{a_{2h_j}s_{h_j}} g^{b_{2h_j}\alpha_{h_j}}$$

Obviously, $\sum_{j=1}^{k} x_{1h_j}y_{1h_j} + \sum_{j=1}^{n-k} x_{1c_j}y_{1c_j} = 0$ and $\sum_{j=1}^{k} x_{2h_j}y_{2c_j} + \sum_{j=1}^{n-k} x_{2c_j}y_{2c_j} = 0$.
So, $\sum_{j=1}^{k} x_{1h_j}y_{1h_j} = -\sum_{j=1}^{n-k} x_{1c_j}y_{1c_j}$ and $\sum_{j=1}^{k} x_{2h_j}y_{2h_j} = -\sum_{j=1}^{n-k} x_{2c_j}y_{2c_j}$.
$g^{x_{1h_k}y_{1h_k}} = g^{-\sum_{j=1}^{n-k} x_{1c_j}y_{1c_j}} * g^{-\sum_{j=1}^{k-1} x_{1h_j}y_{1h_j}}$ and $g^{x_{2h_k}y_{2h_k}} = g^{-\sum_{j=1}^{n-k} x_{2c_j}y_{2c_j}} * g^{-\sum_{j=1}^{k-1} x_{2h_j}y_{2h_j}}$.
Hence, we can rewrite B_{1h_k} and B_{2h_k} as,
$B_{1h_k} = \frac{g^{a_{1h_k}s_{h_k}} g^{b_{1h_k}\alpha_{h_k}}}{g^{\sum_{j=1}^{n-k} x_{1c_j}y_{1c_j}} * g^{\sum_{j=1}^{k-1} x_{1h_j}y_{1h_j}}}$ and $B_{2h_k} = \frac{g^{a_{2h_k}s_{h_k}} g^{b_{2h_k}\alpha_{h_k}}}{g^{\sum_{j=1}^{n-k} x_{2c_j}y_{2c_j}} * g^{\sum_{j=1}^{k-1} x_{2h_j}y_{2h_j}}}$.
Since, the values of x_{1h_j} and x_{2h_j} for $j \in [1, n-k]$ are set by the RSU, she can compute $g^{\sum_{j=1}^{n-k} x_{1c_j}y_{1c_j}}$ and $g^{\sum_{j=1}^{n-k} x_{2c_j}y_{2c_j}}$. So, she can compute $l_{1h_k} = B_{1h_k} * g^{\sum_{j=1}^{n-k} x_{1c_j}y_{1c_j}} = \frac{g^{a_{1h_k}s_{h_k}} g^{b_{1h_k}\alpha_{h_k}}}{g^{\sum_{j=1}^{k-1} x_{1h_j}y_{1h_j}}}$ and $l_{2h_k} = B_{2h_k} * g^{\sum_{j=1}^{n-k} x_{2c_j}y_{2c_j}} = \frac{g^{a_{2h_k}s_{h_k}} g^{b_{2h_k}\alpha_{h_k}}}{g^{\sum_{j=1}^{k-1} x_{2h_j}y_{2h_j}}}$.
Let us denote $l_{1h_j} = B_{1h_j}$ and $l_{2h_j} = B_{2h_j}$ for all $j \in [1, k-1]$. Also denote, $l'_{1h_j} = l_{1h_j}$ and $l'_{2h_j} = l_{2h_j}$ if $s_{h_j} = s'_{h_j}$ for $j \in [1, k]$.
Also denote, $l''_{1h_j} = l_{1h_j}$ and $l''_{2h_j} = l_{2h_j}$ if $s_{h_j} = s''_{h_j}$ for $j \in [1, k]$.
Again, denote $U_1 = (l'_{1h_1}, l'_{1h_2}, \ldots, l'_{1h_k})$, $U_2 = (l'_{2h_1}, l'_{2h_2}, \ldots, l'_{2h_k})$ and $V_1 = (l''_{1h_1}, l''_{1h_2}, \ldots, l''_{1h_k})$, $V_2 = (l''_{2h_1}, l''_{2h_2}, \ldots, l''_{2h_k})$. Now, since, $\sum_{i=1}^{k} w_{h_i} s'_{h_i} = \sum_{i=1}^{k} w_{h_i} s''_{h_i}$ from assumption 4, we can say that $(U_1, U_2) \overset{c}{\approx} (V_1, V_2)$. Hence, the result holds.

7 Complexity Analysis

In this section, we discuss the computation and communication complexity of our proposed scheme. Table 2 provides the computational and communication overheads for a vehicle and the RSU. In our settings, RSU initiates the voting and aggregation phases. During the initialization the RSU P_i computes two elements $\theta_{1i} = g^{a_{1i}}$ and $\delta_{1i} = g^{b_{1i}}$. This requires 2 exponentiations. An encrypted feedback is of the form $\langle B_{1i}, B_{2i}, g^{\alpha_i} \rangle$, where $B_{ji} = Y_{ji}^{x_{ji}} \theta_{ji}^{s_i} \delta_{ji}^{\alpha_i}, j = 1, 2$, and $\alpha_i \in \mathbb{Z}_q$. Hence computation of an encrypted response requires 5 exponentiations. Computation of the NIZK proof of well-formedness of the feedback needs 22 exponentiations. The RSU needs to compute σ_1, σ_2 and $\{\theta_{2i}, \delta_{2i} : i \in [1, n]\}$ during initialization stage. This requires doing $2n + 2$ exponentiations. Again, the RSU needs to do $15n + 8$ exponentiations to compute all the NIZK proofs.

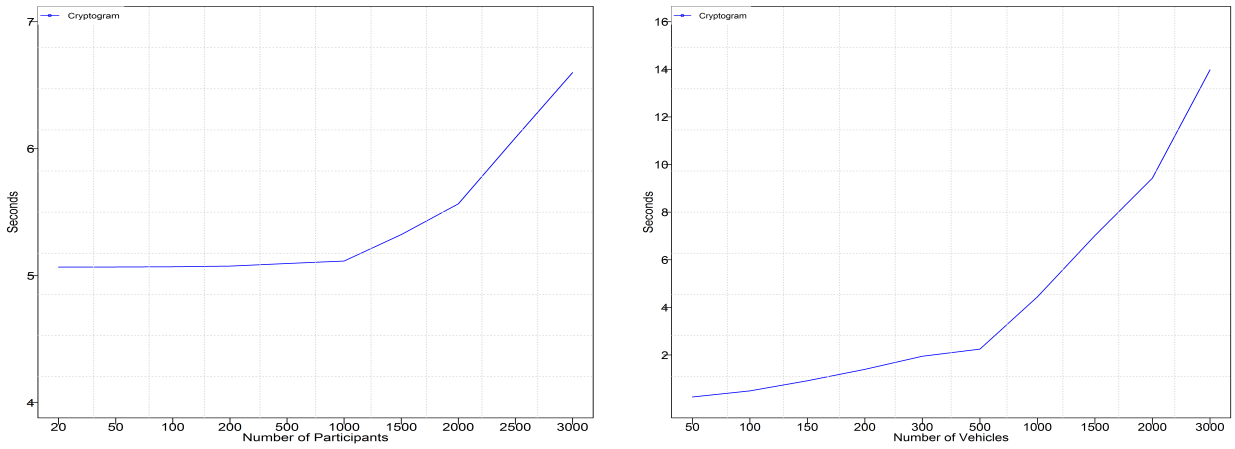
Initially, each participating vehicle needs to post θ_{1i}, δ_{1i}. Hence, during initialization, the communication overhead on a vehicle is 2. The communication overhead to post the encrypted response is 3. Also, each vehicle needs to communicate the NIZK proof of well-formedness of an encrypted feedback which is of size 22. The RSU needs to communicate σ_1, σ_2 and $\{\theta_{2i}, \delta_{2i} : i \in [1, n]\}$ to the tally system during the initialization. Hence, the overhead during

Entity	Computational overhead (number of exponentiations)				Communication overhead		
	Initialization	Feedback	NIZK Proof	Tallying	Initialization	Feedback	NIZK Proof
Vehicle	2	5	22	-	2	3	22
CP	$2n+2$	-	$15n+8$	2	$2n+2$	-	$18n+12$

TABLE 2: Protocol Overhead

Operation	Computation Cost	Communication Cost
Setup	-	$4n+2$
Key	-	$2n$
Feedback	-	$3n$
Verification of all NIZK Proofs	$38n+5an+9$ exponentiations	$32n+5na+11$

TABLE 3: Cost for Public Verification



(a) Vehicle (b) Vehicle

Fig. 3: Vehicle computation time: a) computation time for varying number of participating vehicles, b) computation time for varying number of vehicles for which feedback are provided.

initialization is $2n+2$. Also, the communication overhead of storing all NIZK proofs computed by the RSU is $18n+12$.

The feedback verification is important to handle the malicious vehicles providing out-of-range feedback scores. Table 3 provides the computational and communication overhead on a public verifier to verify all the information (including the final reputation score) provided on the tally system.

8 Performance Analysis

We developed a prototype for the vehicle to report the responses and the analyst for aggregating the responses using Java programs. We evaluated the performance on a system with a CPU 3.6 GHz core i7 and 8 GB memory. At the vehicle side, we evaluated the performance for two aspects: 1) the varying number of vehicles agreed on providing the responses and 2) the number of cryptograms. The performance measurements have been taken for the single core.

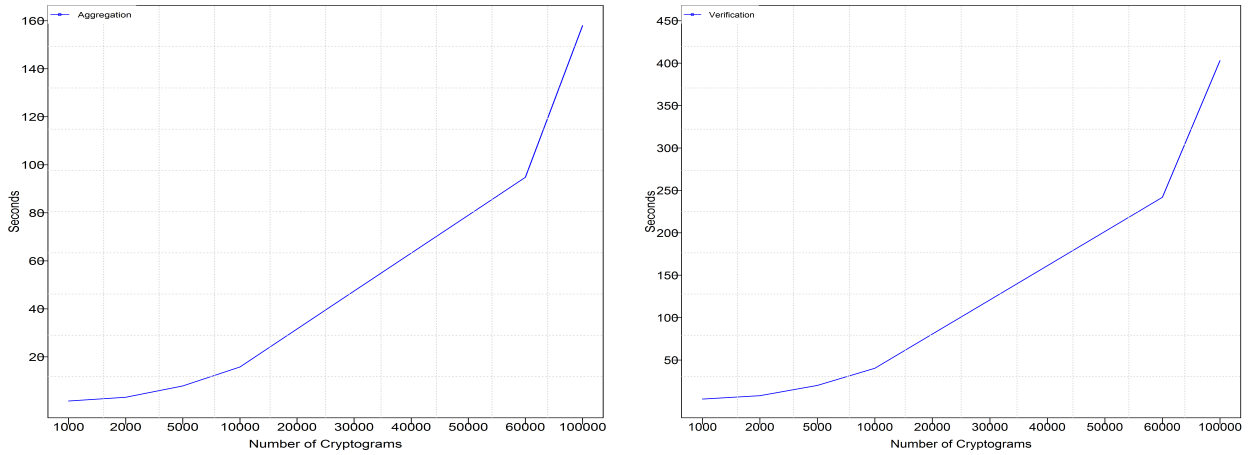
Our evaluation network involves the following scenario. The RSU requests vehicles registered with it to provide their feedback response for other vehicles. The vehicle provides the encrypted responses to the tally system (a public bulletin board) located at the roadside unit. We assume that RSUs are connected to each other and exchange aggregate score of vehicles attached to them with each other.

8.1 Computation Benchmarks

The vehicle providing the feedback response has two major responsibilities: 1) generating the cryptogram of feedback score, and 2) generating the NIZK proof to prove the well-formedness of the encrypted score. Figure 3 shows the vehicle's computation time for two scenarios: a) varying the number of vehicles who agreed to provide the feedback, and b) varying the number of vehicles for which vehicle is required to submit feedback while fixing the number of agreed vehicles to 1000. For the first case, the number of vehicles agreed to provide the feedback is varied from 100 to 10000 while the number of vehicles for which response is provided is fixed at 1000. For the second case, the number of vehicles is varied from 50 to 1000. It can be seen from Figure 3.B that computation time increases linearly with the number of target vehicles. Specifically, for the 3000 vehicles, each vehicle requires around 15 seconds to generate the complete cryptogram (encrypted feedback and NIZK). We can observe from Figure 3.A that varying the number of vehicles would not largely affect the computation time.

Figure 4 presents the time required by the RSU for computing the final tally for each of the vehicles. The RSU would compute the tally from 100K feedback in around 8 seconds. This time does not include the checking of NIZK proof which is the most expensive operation at the aggregator side.

The safety critical messages require an immediate response from the vehicles. In the proposed scheme, the



(a) RSU Aggregation (b) Verification

Fig. 4: Computation Time for the aggregation and verification a) aggregation time for the varying number of cryptograms, b) Verification for the varying number of cryptograms.

vehicle needs to evaluate the trustworthiness of the vehicle from which it received the messages before executing any action. This would incur a delay. Delay in V2I is the time duration when the vehicles receive the message (critical safety message or non-critical messages) from other vehicles and retrieve the reputation score from the RSU. In our scheme, most of reputation computation and aggregation is performed by the RSUs. The same global reputation score is then replicated across all the RSUs. The delay incurred by our method is negligible as it only needs to search the database and purely dependent on the quality of the wireless link between the vehicle and the RSU. The use of high speed dedicated channel would greatly minimize the delay between the RSU and the vehicle. Alternatively, the vehicle also fetches the complete list of vehicles from its connected RSU and decides about the vehicle based on its query to the local database. However, this would increase the load on the vehicle both in terms storage and searching the local database.

# of Vehicles	Vehicle to Tally	# of Cryptograms	Verification
50	41Kb	1000	18.9Mb
100	81Kb	2000	37.8Mb
150	121Kb	5000	94.8Mb
200	162Kb	10000	190.3Mb
300	244Kb	20000	378.8Mb
500	412Kb	30000	568.8Mb
1000	816Kb	40000	757.76Mb
2000	1.7Mb	60000	1.36GB
3000	2.6Mb	100000	1.849GB

TABLE 4: Bandwidth Overhead for the Vehicle and the Tally System. # of cryptogram is the total number of cryptograms at the tally system.

8.2 Communication Benchmark

We now turn to discuss the communication bandwidth required by each vehicle and the storage overhead required by the tally system to handle the responses from 1000 vehicles. The communication overhead for the vehicle and RSU is shown in Table 8.2. A single-vehicle sends around 11 MB data during the whole protocol operations for 10 videos and 9 different questions. Again it can be seen that the most bandwidth intensive operation is the transport of NIZK proofs that alone consume around 9 MB of bandwidth. At the vehicle side, the bandwidth required does not go beyond 10 MB over the period of one day which is also reasonable in the presence of WIFI network or mobile data enabled smartphones. The storage requirement at RSU side is reasonable as RSU already has a large storage space and can handle responses even from hundreds of thousands of users.

One of the major properties of TrustVote along with privacy-preservation is the verification of aggregating score of the object without relying on any trusted setup and a trusted third party system. In the verification process, participating vehicles should be able to verify the score stated by the RSU. Figure 4(b) represents the verification overheads for the experiments mentioned above. The verification process consists of four major operations: system setup, public and private key generation, the verification of feedback scores and NIZK proof. The aggregate time for 100K is around 380 seconds.

Table 8.2 reports the communication bandwidth required for the exchange of data from the vehicle to the tally system. The number of objects varies from 50 to 3000. Table also represents the communication overhead required for the verification of vehicle score and the aggregating of the final score. The number of feedback providers varies from 1000 to 100K. The verification overhead for a large number of cryptograms is high but in a real scenario, it is acceptable as the RSU normally has high-end resources for processing and storing user data.

9 Conclusion

The Internet of Vehicles (IoV) is a network of connected vehicles and roadside units that enable participants to disseminate information for an improved driving experience. The dissemination of false information from malicious nodes in the vehicular network could have negative consequences on

traffic management and more significantly even cause a serious threat to life. There is a strong need to have a system that enables drivers to evaluate the trustworthiness of a message sender before making any decision based upon the received information. In this paper, we have presented a TrustVote system, a decentralized privacy-preserving reputation protocol to evaluate the trustworthiness of vehicles on the Internet of Vehicle network without relying on the trusted system. Our system utilizes a homomorphic cryptographic system for hiding the feedback of vehicles and reveals nothing more than just the aggregate reputation of vehicles in the end. The system prevents malicious vehicles from providing out-of-range feedback values and ensures correct computation even in the presence of malicious vehicles. The salient features of a TrustVote system include: 1) a decentralized system without requiring trust third parties, 2) providing weighted aggregation, 3) preventing malicious participants from providing illegal feedback, and 4) low communication and computation overheads. We evaluate the protocol through a prototype implementation. The evaluation results show that the system has reasonable communication and communication overheads while having important features of privacy-preservation and decentralization.

ACKNOWLEDGEMENT

This work is supported by the ERC Starting Grant, No. 306994. We thank the anonymous reviewers for their insightful comments and suggestions.

REFERENCES

[1] J. Worner, "Focused delivery of key market enablers in 2017/18," 2017. [Online]. Available: https://www.gsma.com/iot/newscat/mautomative-newscat/

[2] (2017) Gartner says by 2020, a quarter billion connected vehicles will enable new in-vehicle services and automated driving capabilities. [Online]. Available: http://www.gartner.com/newsroom/id/2970017.

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

[4] L. Cheng, B. E. Henty, D. D. Stancil, F. Bai, and P. Mudalige, "Mobile vehicle-to-vehicle narrow-band channel measurement and characterization of the 5.9 ghz dedicated short range communication (dsrc) frequency band," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 8, pp. 1501–1516, Oct 2007.

[5] K. Dar, M. Bakhouya, J. Gaber, M. Wack, and P. Lorenz, "Wireless communication technologies for its applications [topics in automotive networking]," *IEEE Communications Magazine*, vol. 48, no. 5, pp. 156–162, May 2010.

[6] P. Papadimitratos, A. D. L. Fortelle, K. Evenssen, R. Brignolo, and S. Cosenza, "Vehicular communication systems: Enabling technologies, applications, and future outlook on intelligent transportation," *IEEE Communications Magazine*, vol. 47, no. 11, pp. 84–95, November 2009.

[7] S. Parkinson, P. Ward, K. Wilson, and J. Miller, "Cyber threats facing autonomous and connected vehicles: Future challenges," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 11, pp. 2898–2915, 2017.

[8] M. A. Azad, S. Bag, and F. Hao, "M2m-rep: Reputation of machines in the internet of things," in *Proceedings of the 12th International Conference on Availability, Reliability and Security*, 2017, pp. 28:1–28:7.

[9] M. A. Azad, S. Bag, F. Hao, and K. Salah, "M2m-rep: Reputation system for machines in the internet of things," *Computers & Security*, vol. 79, pp. 1 – 16, 2018.

[10] Z. Lu, G. Qu, and Z. Liu, "A survey on recent advances in vehicular network security, trust, and privacy," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–17, 2018.

[11] S. Wang and N. Yao, "A rsu-aided distributed trust framework for pseudonym-enabled privacy preservation in vanets," *Wireless Networks*, Feb 2018.

[12] "Etsi ts 102 941 v1.1.1- intelligent transport systems (its); security; trust and privacy management", technical report, 2012."

[13] "Etsi ts 103 097 v1.2.1- intelligent transport systems (its); security; security header and certificate formats", technical report, 2016."

[14] C. Lai, K. Zhang, N. Cheng, H. Li, and X. Shen, "Sirc: A secure incentive scheme for reliable cooperative downloading in highway vanets," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 6, pp. 1559–1574, June 2017.

[15] R. Silva and R. Iqbal, "Ethical implications of social internet of vehicles systems," *IEEE Internet of Things Journal*, pp. 1–1, 2018.

[16] F. Hao, P. Y. A. Ryan, and P. Zielinski, "Anonymous voting by two-round public discussion," *IET Information Security*, vol. 4, no. 2, pp. 62–67, 2010.

[17] N. Bissmeyer, J. Njeukam, J. Petit, and K. M. Bayarou, "Central misbehavior evaluation for vanets based on mobility data plausibility," in *Proceedings of the Ninth ACM International Workshop on Vehicular Inter-networking, Systems, and Applications*, ser. VANET '12, 2012, pp. 73–82.

[18] X. Li, J. Liu, X. Li, and W. Sun, "Rgte: A reputation-based global trust establishment in vanets," in *2013 5th International Conference on Intelligent Networking and Collaborative Systems*, Sept 2013, pp. 210–214.

[19] O. Hasan, L. Brunie, E. Bertino, and N. Shang, "A decentralized privacy preserving reputation protocol for the malicious adversarial model," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 6, pp. 949–962, 2013.

[20] M. Raya, P. Papadimitratos, V. D. Gligor, and J. P. Hubaux, "On data-centric trust establishment in ephemeral ad hoc networks," in *IEEE INFOCOM 2008 - The 27th Conference on Computer Communications*, April 2008.

[21] G. Costantino, F. Martinelli, I. Matteucci, A. Bertolino, A. Calabro, and E. Marchetti, "CARS: Context Aware Reputation Systems to Evaluate Vehicles' Behaviour," in *2018 26th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP)*, Mar. 2018, pp. 446–453.

[22] F. G. Màrmol and G. M. Pèrez, "Trip, a trust and reputation infrastructure-based proposal for vehicular ad hoc networks," *Journal of Network and Computer Applications*, vol. 35, no. 3, pp. 934 – 941, 2012, special Issue on Trusted Computing and Communications.

[23] J. Oluoch, "A distributed reputation scheme for situation awareness in vehicular ad hoc networks (vanets)," in *2016 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)*, March 2016, pp. 63–67.

[24] P. Dewan, P. Dasgupta, and A. Bhattacharya, "On using reputations in ad hoc networks to counter malicious nodes," in *Proceedings of Tenth International Conference on Parallel and Distributed Systems,2004.*, July 2004, pp. 665–672.

[25] X. Huang, R. Yu, J. Kang, and Y. Zhang, "Distributed reputation management for secure and efficient vehicular edge computing and networks," *IEEE Access*, vol. 5, pp. 25 408–25 420, 2017.

[26] M. Azad, S. Bag, S. Tabassum, and F. Hao, "privy: Privacy preserving collaboration across multiple service providers to combat telecoms spam," *IEEE Transactions on Emerging Topics in Computing*, pp. 1–1, 2017.

[27] L. Fischer, F. Dtzer, and P. Magiera, "Vars: A vehicle ad-hoc network reputation system," in *Sixth IEEE International Symposium on a World of Wireless Mobile and Multimedia Networks(WOWMOM)*, vol. 01, 06 2005, pp. 454–456.

[28] Z. Huang, S. Ruj, M. A. Cavenaghi, M. Stojmenovic, and A. Nayak, "A social network approach to trust management in vanets," *Peer-to-Peer Networking and Applications*, vol. 7, no. 3, pp. 229–242, 2014.

[29] Z. Yang, K. Yang, L. Lei, K. Zheng, and V. C. M. Leung, "Blockchain-based decentralized trust management in vehicular networks," *IEEE Internet of Things Journal*, pp. 1–1, 2018.

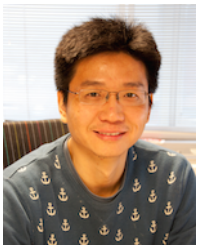
[30] P. Asuquo, H. Cruickshank, J. Morley, C. P. A. Ogah, A. Lei, W. Hathal, S. Bao, and Z. Sun, "Security and privacy in location-based services for vehicular and mobile communications: An overview, challenges and countermeasures," *IEEE Internet of Things Journal*, 2018.

[31] R. Hussain, D. Kim, J. Son, J. Lee, C. A. Kerrache, A. Benslimane, and H. Oh, "Secure and privacy-aware incentives-based witness service in social internet of vehicles clouds," *IEEE Internet of Things Journal*, vol. 5, no. 4, pp. 2441–2448, 2018.
[32] L. Zhu, C. Zhang, C. Xu, X. Du, R. Xu, K. Sharif, and M. Guizani, "Prif: A privacy-preserving interest-based forwarding scheme for social internet of vehicles," *IEEE Internet of Things Journal*, 2018.
[33] P. Y. A. Ryan, D. Bismark, J. Heather, S. Schneider, and Z. Xia, "Prêt à voter: A voter-verifiable voting system," *Trans. Info. For. Sec.*, vol. 4, no. 4, pp. 662–673, Dec. 2009.
[34] E. Shi, R. Chow, T. h. Hubert Chan, D. Song, and E. Rieffel, "Privacy-preserving aggregation of time-series data," in *In NDSS*, 2011.
[35] H. Corrigan-Gibbs and D. Boneh, "Prio: Private, robust, and scalable computation of aggregate statistics," in *Proceedings of the 14th USENIX Conference on Networked Systems Design and Implementation*, ser. NSDI'17, 2017, pp. 259–282.
[36] M. A. Azad and S. Bag, "Decentralized privacy-aware collaborative filtering of smart spammers in a telecommunication network," in *Proceedings of the Symposium on Applied Computing*, ser. SAC '17, New York, NY, USA, 2017, pp. 1711–1717.
[37] M. A. Azad, S. Bag, S. Tabassum, and F. Hao, "privy: Privacy preserving collaboration across multiple service providers to combat telecoms spam," *IEEE Transactions on Emerging Topics in Computing*, vol. PP, no. 99, pp. 1–1, 2017.
[38] F. Hao, M. N. Kreeger, B. Randell, D. Clarke, S. F. Shahandashti, and P. H.-J. Lee, "Every vote counts: Ensuring integrity in large-scale electronic voting," in *2014 Electronic Voting Technology Workshop/Workshop on Trustworthy Elections (EVT/WOTE 14)*, San Diego, CA, 2014.
[39] B. Aslam, F. Amjad, and C. C. Zou, "Optimal roadside units placement in urban areas for vehicular networks," in *2012 IEEE Symposium on Computers and Communications (ISCC)*, July 2012, pp. 000 423–000 429.

Muhammad Ajmal Azad received the Ph.D. (2016) degree in Electrical and Computer Engineering from the University of Porto, Portugal and MS (2008) in Electronics Engineering from the International Islamic University Pakistan. He is currently a research fellow (an equivalence of lecturer in the UK) in the department of computer science at The University of Warwick. Previously, he spent more two years at Newcastle University as the research associate. He also worked in a leading telecommunication company for four years. His research interests include privacy-aware collaboration, reputation aggregation, privacy protection, privacy-aware outsourcing of network logs and spam detection in telecommunication networks.

Samiran Bag received the M.Tech degree in computer science from the Indian Statistical Institute, and the Ph.D. degree from the Indian Statistical Institute. He was a Post-Doctoral Researcher with Kyushu University, Japan and a Research associate at the School of Computing Science, Newcastle University, UK. He is currently a Research Fellow with the Department of Computing Science, University of Warwick, UK. His primary research areas are electronic voting, cryptocurrency, and secure multiparty computation.

Simon Parkinson is a Reader in Cybersecurity within the School of Computing and Engineering at the University of Huddersfield. He has an honours degree in Secure and Forensic computing and a Ph.D. in the cross-discipline use of domain-independent artificial intelligence planning to autonomously produce measurement plans for machine tool calibration. This resulted in the ability to produce measurement plans to reduce both machine tool downtime and the uncertainty of measurement. His research interest is in developing intelligent systems for manufacturing and cybersecurity. Simon is currently researching in the area of improving cybersecurity awareness and audits through developing novel, intelligent software tools. As part of his research, Simon performs research into the vulnerabilities of systems and to develop mitigation technologies.



Feng Hao is a Professor of Security Engineering in the Department of Computer Science, University of Warwick. He graduated with a Ph.D. in 2007 from the Computer Laboratory, University of Cambridge. His research interests include applied cryptography, security engineering, and efficient computing. Since 2013, he has been serving as an associate editor for the IEEE Security and Privacy magazine. He is supported by ERC Starting Grant (No. 306994) and ERC Proof of Concept Grant (No. 677124).