This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBDATA.2017.2710347, IEEE Transactions on Big Data

Trustworthiness Evaluation-based Routing Protocol for Incompletely Predictable Vehicular Ad hoc Networks

Jian Shen, *Member, IEEE,* Chen Wang, Aniello Castiglione, *Member, IEEE*
Dengzhi Liu and Christian Esposito, *Member, IEEE*

Abstract—Incompletely predictable vehicular ad hoc networks is a type of networks where vehicles move in a certain range or just in a particular tendency, which is very similar to some circumstances in reality. However, how to route in such type of networks more efficiently according to the node motion characteristics and related historical big data is still an open issue. In this paper, we propose a novel routing protocol named trustworthiness evaluation-based routing protocol (TERP). In our protocol, trustworthiness of each individual is calculated by the cloud depending on the attribute parameters uploaded by the corresponding vehicle. In addition, according to the trustworthiness provided by the cloud, vehicles in the network choose reliable forward nodes and complete the entire route. The analysis shows that our protocol can effectively improve the fairness of the trustworthiness judgement. In the simulation, our protocol has a good performance in terms of the packet delivery ratio, normalized routing overhead and average end-to-end delay.

Index Terms—Incompletely Predicable Networks, VANETs, Trustworthiness Evaluation, Routing Protocol, Self-Configured Networks.

1 Introduction

NOWADAYS, *incompletely predictable ad-hoc networks* (IPNs) which is firstly proposed in [1] has received researchers' attention. As a type of ad hoc networks [2], [3], [4], [5], [6], IPNs has many similarities compared with typical ad hoc networks, which lack of pre-existing infrastructures and are self-configuring and decentralized. The topological structure of nodes in such network might change at any time. As a consequence, the communications among nodes are difficult to be guaranteed. In addition, IPNs has a lot of unique characteristics. One of the most important features of IPNs is that nodes in the network only move in their respective specific ranges, or are subject to certain rules of movement. In a vehicle network, the range of motion and the moving trajectory of each vehicle are substantially predictable, which can help the network to make the route choice for message transmission ahead of time [7], [8]. Vehicle networks (Vehicular Ad hoc Networks, VANETs) are considered similar to Online Social Networks (OSNs). There are many predictable elements in social networks. It is possible to simplify the complexity of routing in such a network as long as the rules of variation are partially known [9], [10], [11], [12]. For instance, taxis in a city have a range for regular activities [13], [14], [15], [16]. Besides, each of their business can basically determine the route arrangement, which will greatly take advantage of our routing scheme.

In addition to the traveling path of a vehicle, the trustworthiness of it is also crucial in several field of today daily life [17], [18]. It is significant that the algorithm for calculating the vehicle's trustworthiness should be able to fully express attributes of the vehicle and the relationship among attributes. The trustworthiness of a vehicle can be evaluated in many aspects [5], [19]. For example, the model of the vehicle, mileage, fuel consumption, vehicle accident records, violations of traffic regulations, the number of times that the node acts as a relay node and the performance of the node are all very important attributes affecting the value of a vehicle trustworthiness.

After having an appropriate algorithm to compute the vehicle trustworthiness, a robust center for big data collection, analysis and trustworthiness distribution is also demanded [20]. As is well known, the most important features of cloud computing are distributed computing and mass storage [21], [22], [23], [24], [25], [26], [27], [28]. Applying the features of cloud computing into practice, researchers have proposed many outstanding schemes and protocols in their own fields [29], [30], [31], [32]. It is a good choice to use the cloud to collect, store and analyze attribute parameters [33] to provide the trustworthiness of any vehicle to the nodes querying for the vehicle's trustworthiness. Specifically, current vehicles are equipped

- *J. Shen is with Jiangsu Engineering Center of Network Monitoring, Jiangsu Technology & Engineering Center of Meteorological Sensor Network, Jiangsu Collaborative Innovation Center on Atmospheric Environment and Equipment Technology, School of Computer and Software, Nanjing University of Information Science & Technology, Nanjing, China, 210044. E-mail: s_shenjian@126.com*
- *C. Wang and D. Liu are with School of Computer and Software, Nanjing University of Information Science & Technology, Nanjing, China, 210044. E-mail: wangchennuist@126.com, liudzdh@qq.com*
- *A. Castiglione and C. Esposito are with Department of Computer Science, University of Salerno, I-84084 Fisciano (SA) Italy. E-mail: castiglione@ieee.org, esposito@unisa.it*

Corresponding author: Christian Esposito, Department of Computer Science, University of Salerno, Via Giovanni Paolo II, I-84084 Fisciano (SA) Italy. (e-mail: esposito@unisa.it).

Manuscript received on January 13, 2017; revised on April 18, 2017.
For information on obtaining reprints of this article, please send e-mail to: preprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TBDATA.2017.1234567

2332-7790 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBDATA.2017.2710347, IEEE Transactions on Big Data

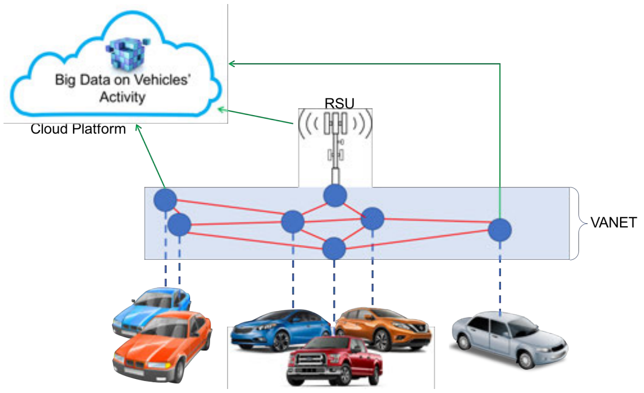


Fig. 1: Integration of cloud with VANETS for collecting and storing big data sets related to vehicles.

with both short-range communication means (such as the ones provided by Bluetooth and Wi-Fi) and long-range ones (such as the cellular networks). The first ones are used to build a vehicular network (VANET) in order to exchange information among neighboring vehicles, while the second ones are used to remote access services for road traffic monitoring, weather forecasting and so on. For the vehicles that do not exhibit the long-range communication means, alongside the road it is possible to have proper units that can provide connectivity support to passing vehicles. Long-rage communications can be used in order to let vehicles to store in the cloud some data related to the driver behavior monitoring and analysis, or the vehicle trajectory and speed, as shown in Fig. 1. Given the enormous number of vehicles and the proliferation over the time of such a kind of data, we should consider that in this kind of applications the cloud have to manage a vast amount of historical information, which is a clear example of Big Data [34]. We advocate that such historical big data is particularly valuable for the proper estimation of trustworthiness of nodes within a VANET and to support an efficient routing scheme in such a context.

The motivation of our work: Although many protocols have been proposed for vehicle networks [35], [36], some open issues, such as network feature utilization and trustworthiness evaluation, etc., have not yet been resolved. From one side, relationships among different timestamps and positions in the same vehicle are not fully used as the important information for routing search. In traditional routing protocols, routing situations at each moment are usually considered separately. However, the continuity of vehicle movement in time is ignored. On the other side, there is no way to completely avoid the current trustworthiness mechanism from providing false or unfair trustworthiness value. In a trusted scheme or mechanism, its own trustworthiness value is mainly given by the node itself, or decided by team members [37], [38].

1.1 Our Contribution

In this paper, we proposed a trustworthiness evaluation-based routing protocol (TERP). The trustworthiness of every vehicle is taken into consideration, which is obtained through the attribute parameters of the vehicle and is provided by the cloud. The cloud is used in order to host the vast amount of historical big data related to vehicle activities, on which the trust evaluation is based. Integrating vehicle control and networks with cloud computing [39] and leveraging big data for the development of smart solutions for sustainable and secure transportation [40], as depicted in Fig. 1, represent the frontier of the future mobility and smart city concepts. Our work represents a valuable innovation in such direction, and its main contribution can be outlined as follows:

- **A balanced node utilization ratio is defined.** The node's contribution into a routing selection process is defined based on the size of the transmission packet. The concept of anti-pheromone is presented to record the contribution of every vehicle. The value of anti-pheromone is sent to the cloud along with other attribute parameters. The cloud takes it into account for the representation of node utilization ratio.
- **A fair trustworthiness evaluation is proposed.** The cloud server acts as a trusted third part to provide fair trustworthiness evaluation. This avoids problems caused by different criteria for the evaluation of vehicle trustworthiness. At the same time, the cloud server can also reduce the calculation and storage overhead of each vehicle terminal.
- **A method for value re-excavation from known information is presented.** This protocol analyzes and summarizes the historical big data of vehicle nodes [41], [42] with a view to mining out valuable information. This information will help the cloud server to conduct a more detailed and comprehensive assessment for vehicle trustworthiness.

The work described in this paper represents an extension of the one previously presented in [43]. The previous approach has been augmented by inserting a trust-based selection of candidate nodes to whom deliver messages, where trust is evaluated based on historical big data hosted in the cloud and by means of a principal component analysis.

1.2 Related Work

Plenty of trustworthiness evaluation methods and routing protocols are presented in the literature [44]. Only some notable works are reviewed here due to space limitation. In ad hoc networks and other network types [3], trustworthiness has always been a hot topic.

Theodorakopoulos *et al.* [45] define the trust evaluation process as a path problem in a direct graph. They claim that their method is robust to current attacks. A graph-based trust evaluation is also contained in [46] where its methodologies and challenges when applied to the case of online social networks are explained.

He *et al.* [37] propose a trust management through the use of simple cryptographic techniques for medical sensor networks. This method can provide some security, but put over-reliance on traditional encryption technology.

Jiang *et al.* [47] present a user-domain-based trusted acquaintance chain discovery algorithm to take advantage of potential relationships in on-line social networks. In fact,

2332-7790 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBDATA.2017.2710347, IEEE Transactions on Big Data

vehicle networks are also a type of social networks, the work mentioned by Jiang *et al.* [47] is worthy of reference.

There have been many trust models that are worthy of reference being presented for other fields in recent years. Although not designed specifically for routing protocols in vehicle networks, these trust models and methods have a certain reference value [48], [49].

Zhou *et al.* [38] determine the new node's trustworthiness as a direct security degree based on the historical security event record. They utilize correlation coefficient to distinguish malicious vehicle nodes.

Esposito *et al.* [50] compute the trust degree of a certain entity by means of the aggregation of linguistic fuzzy sets collected by the neighbors and the quantitative assessment of the entity's trust from direct observation. They use mechanism design in order to force the neighbors to say the truth about the entity's reputation. A similar Multi-dimensional fuzzy trust evaluation is also described in [51] and applied to the mobile social networks.

Mohsenzadeh *et al.* [52] present a fuzzy mathematics based trust evaluation model for cloud computing according to records of interaction between cloud entities. Simulation experiments show that the proposed model can effectively identify malicious entities, and assist the system to correctly make security decisions.

Rajendran *et al.* [53] formulate a hybrid model to calculate the trustworthiness of cloud service providers based on compliance and reputation of providers. In their paper, the user's feedback is used as a basis for trustworthiness evaluation.

Jiang *et al.* [54] propose a solution for evaluating trust In online social networks by considering trust evidence propagated along trusted paths and the challenge of trust decay through propagation. The authors driving idea is to consider the similarity between trust propagation and network flow, then the problem of trust evaluation has been converted into a generalized network flow problem.

Wang *et al.* [55] present a reputation management scheme for pseudonym-enabled VANETs. Two types of reputations are defined, which are named as the service reputation and the feedback reputation. Reputation accumulation algorithms are designed based on the information entropy and the majority rule. They claim that their scheme is robust against the tactical attack and can preserve privacy against the reputation link attack during pseudonym changes.

Pop *et al.* [56] propose a trust management in structured overlay networks. Their management is stated to be established on the basis of the mobile cloud architecture and a honeycomb overlay is considered.

However, most of these methods cannot adapt to characteristics of IPNs, including those in [57], [58], [59], [60], and cannot make full use of features of vehicle networks. So, we put forward a novel protocol named trustworthiness evaluation-based routing protocol for incompletely predictable vehicular ad hoc networks in this paper.

1.3 Organization

The remainder of this paper is organized as follows. Section 2 provides some preliminaries about concepts of IPNs mentioned in the paper. Section 3 proposes a detailed explanation of the TERP protocol. Section 4 presents a trustworthiness evaluation test and simulation results of the novel protocol are shown. Finally, the conclusions are drawn in Section 5.

2 Preliminaries

Delay Tolerant Networks (DTNs) [61], Wireless Mesh Networks (WMNs) [2] and Wireless Sensor Networks (WSNs) [3], [8], [62], [63], [64], [65] widely mention the so-called incompletely predictable networks, whose example is given in Fig. 2. In vehicular networks, cars move in a particular tendency or in a specific range in the city [66], [67], [68]. We put the topology constructed by vehicles into a space-time graph [69], [70], [71]. In social networks, the node mobility can be predicted with a potential accuracy of about 93 percent [9], [11], [72]. Moreover, there is a specific situation where the node positions and the link status are fixed [73], [74]. IPNs are a suitable tool to model VANETs and we have used them as the basis for our protocol.

Definition 1 (The fundamental model of IPNs). We define IPNs as an undirected graph G. G is a two-tuple constructed with a finite nonempty set $V(G)$ and an unordered pair set $E(G)$. In other words, $G = (V(G), E(G))$. In detail, $V(G) = \{v_1, v_2, ..., v_n\}$ is called the *basic position* or *BP* set and each element $v_i (i = 1, 2, ..., n)$ in $V(G)$ represents *BP* of each node. $E(G) = \{e_1, e_2, ..., e_m\}$ is the link set of graph G. Every element e_k in $E(G)$ is an unordered pair of two specific elements v_i and v_j, reported as $e_k = (v_i, v_j)$ or $e_k = v_i v_j = v_j v_i (k = 1, 2, ..., m)$.

Definition 1 gives the definition of points, the relation between points, and the concrete meaning of *BP*s in IPNs. A visual representation of IPNs is that IPN nodes have their own fixed movement range or trajectory, which is the main reason for the semi-predictability of the network.

Definition 2 (Moving range). R_{mov} represents the radius of the movement range of each node. The movement range of a node in which it can travel through is described as a circle with the *BP* as its center and with the radius of R_{mov}. The moving range or active range of node i is defined as a circle taking *BP* v_i as its center and R_{mov} as its radius.

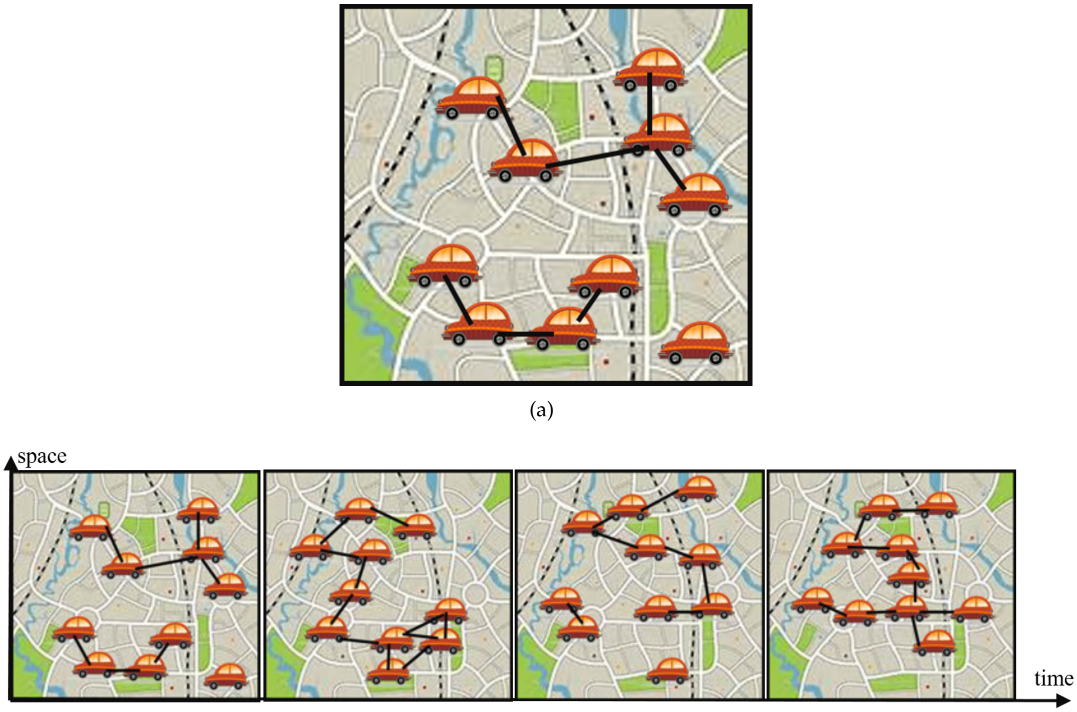
The moving range refers to the maximum radius that nodes can move from the *BP*s according to the model. This range may vary with each node's own situation. However, in this paper we assume that the ranges of mode movement are the same size, and remain unchanged in the process of routing transmission.

Definition 3 (Transmission range). The radius of transmission range is defined as R_{trans}. This definition indicates that nodes, in the circular area with the corresponding node of each *BP* as its center and with the radius of R_{trans}, can receive massages from the node.

As with the moving range, each node in the network has its own broadcast radius, which is called the transmission range [74]. The difference is that the area covered by the transmission range moves as the node moves. Similarly, in

2332-7790 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBDATA.2017.2710347, IEEE Transactions on Big Data



(a)

(b)

Fig. 2: A time-evolving network: (a) a snapshot of the network and (b) time-evolving topologies of the network

TABLE 1: Notations in our scheme

Symbol	Description
v_i	A vehicle whose identity is i
φ	An attribute parameter (AP)
$\mathcal{P}$	AP list
APh	Anti-pheromone
τ_i^t	The trustworthiness of node i at time t
$Pr_{i\to j}^t$	The probability that node i can be successfully heard by j at time t
$Pr_{i\to\forall}^t$	The probability that node i can be successfully heard by any nodes at time t
$\mathcal{C}$	The candidate node list
Pr_d	The default transmission probability value
Δt	The freshness of the trustworthiness value

order to simplify the description of the protocol, we assume that transmission ranges of nodes in an IPN are the same size and unchanged.

The above three definitions basically cover the concepts, characteristics and specific details of IPNs. As shown in Fig. 2, the positions of vehicles at each moment are different, but their movement will only change within a certain area. This will greatly assist the design of routing protocol. In order to better illustrate the designed protocol, some important notations are listed in Table 1.

3 The Proposed Protocol

The detailed routing protocol is presented in this section. An overview of the proposed TERP is given before the description of the specific protocol. The protocol is composed of three phases: the preparatory phase, the routing phase and the delivery phase. Then the process of trustworthiness evaluation is presented, and the cloud server judges the trustworthiness of every individual in the network according to the trustworthiness evaluation process. Finally, detailed routing steps are stated.

3.1 Overview of TERP

The proposed TERP utilizes the cloud sever to evaluate the trustworthiness of every vehicle in the network. The trustworthiness is given by the cloud based on attribute parameters (APs) uploaded by vehicles themselves. Nodes select an appropriate route by choosing an optimal relay node according to the trustworthiness of every candidate node. These candidate nodes are determined in the view of transmission probability. The process of TERP is given in Fig. 3.

The basic flow of TERP is as follows. In preparatory phase, the vehicle will upload all the required attribute parameters to the cloud. The cloud will record the data along with the timestamp accurately. These data will be treated as attribute parameters and be utilized to calculate the users' trustworthiness. In the routing phase, the node first determines the set of candidate nodes by the transmission probabilities of its neighboring nodes. The set of candidate nodes is uploaded to the cloud. Notre that the node that uploads the candidate node list to the cloud is called the query node. After obtaining the list, the cloud queries the entire storage space, and APs of all the nodes involved in the list will be utilized to calculate the trustworthiness of each node. The cloud then feeds back the trustworthiness list to the query node. In the delivery phase, the query node that has obtained the trustworthiness of the candidate nodes sends the message to the node with the optimal trustworthiness. To summarize, TERP evaluates the vehicle according

2332-7790 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBDATA.2017.2710347, IEEE Transactions on Big Data

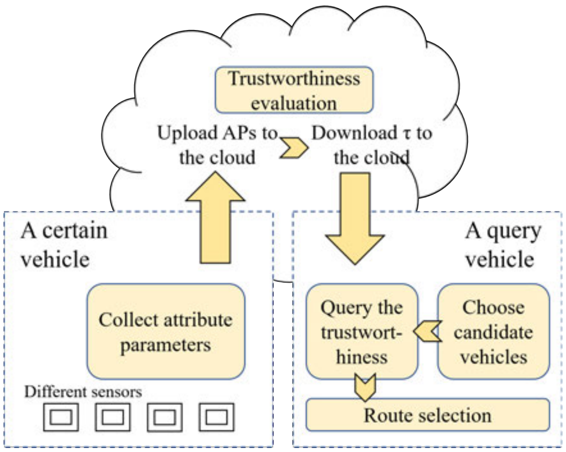


Fig. 3: The process of TERP

to its attributes with the assistance of the cloud, and filters in candidate vehicles according to their trustworthiness to help the network make route decisions.

3.2 The Detailed TERP

The TERP is composed of three phases: the preparatory phase, the routing phase and the delivery phase. The relationship between the three phases is illustrated in Fig. 3. Note that the detailed evaluation method will be stated in the next subsection.

3.2.1 Preparatory Phase

The preparatory phase represents the event happening from time to time in the proposed protocol. In the preparatory phase, vehicles in the network upload attribute parameters of themselves collected by various sensors in vehicles. These parameters include but not limited to the following: the model of the vehicle, mileage, fuel consumption, vehicle accident records, violations of traffic regulations, the number of times that the node acts as a relay node and the performance of the node. For instance, vehicles always have the similar quality if they are the same model. In addition, the mileage and time of travel can reflect the frequency of the vehicle utilization. The uploaded attribute parameters (APs) are defined as follows. Definition 4 and Definition 5 provides the definitions of AP and AP list respectively.

Definition 4 (Attribute Parameter, AP). An AP φ is defined as a parameter which can reflect the performance of the corresponding vehicle.

Definition 5 (AP list). An AP list is a list $\mathcal{P}$ containing all the APs collected by sensors on the vehicle for the evaluation process. Suppose that there are h valid APs of the vehicle. An AP list can be represented by $\mathcal{P} = \{\varphi_1, ..., \varphi_h\}$.

Among all the h APs, there is one important parameter needs to be noted, which is named as anti-pheromone, which is utilized to reflect the node utilization rate. Here gives its specific connotation.

Anti-pheromone (APh). In Ant Colony Optimization (ACO) algorithms, people define the concept of "pheromone" from natural phenomenon [75]. The so-called "pheromone" gradually evaporates as time goes by. The value of pheromone indicates the characterization of a path. In our protocol, to ensure data security and route reliability, the vehicle utilization ratio needs to be controlled to prevent a portion of nodes from controlling the data transmitted on the routes. Under these circumstances, anti-pheromone is designed to create a more utilization efficient network, to achieve utilization-balance through the entire network and guarantee the relative fairness of the information obtained by nodes in the network. Algorithm 1 shows how to calculate and transmit the value of anti-pheromone APh. In this protocol, the increase of anti-pheromone is determined by the ratio r of the data size of each transmission to the total storage space of the node. In order to make the added value meaningful, the ratio r is multiplied by an artificially set coefficient ϵ which is in front of this ratio in the algorithm.

Algorithm 1 Upon receiving RREQ$< S, D >$ or Msg$< S, D >$

```
$r \leftarrow$ package size / storage size
**if** $n_i == D$ **then**
    **return** RREP
**else**
    APh ← APh + $\epsilon \cdot r$
    **return** ACK<Confirmation message, APh>
    //APh: the value of Anti-pheromone of the certain node
**end if**
```

Specifically, when a node receives RREQs or messages from other nodes, the node first calculates the size of the package. The resulting value is compared with the storage size of the node itself to obtain a ratio r. If the node is the destination node, it directly accepts the message or feeds back RREP. Otherwise it will add the original anti-pheromone with a value, which is just a multiple of ratio r. The new anti-pheromone is stored in the acknowledgment packet and sent to the previous node.

In simple terms, the anti-pheromone is used to record the number of times each vehicle node serves as a relay node. This is because the size of most routing packets is similar, although the message packet size will be different. In this protocol, the anti-pheromone will be listed in the AP list along with other parameters and uploaded to the cloud.

Once the AP list $\mathcal{P}$ is completed, $\mathcal{P}$ will be sent to the cloud server. The cloud stores all the AP lists preparing for the trustworthiness evaluation. Here we give the definition of trustworthiness.

Definition 6 (Trustworthiness). The trustworthiness τ_i^t of node i is calculated by the cloud according to the APs that node i uploads to the cloud at time t.

In this paper, according to the vehicle's AP list, the principal component analysis (PCA) algorithm is adopted to obtain the corresponding trustworthiness. To make the description more clear, the specific acquisition method will be provided in detail later.

3.2.2 Routing Phase

In the routing phase, a node wants to transmit its message to a certain destination node. The node chooses a set of

2332-7790 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBDATA.2017.2710347, IEEE Transactions on Big Data

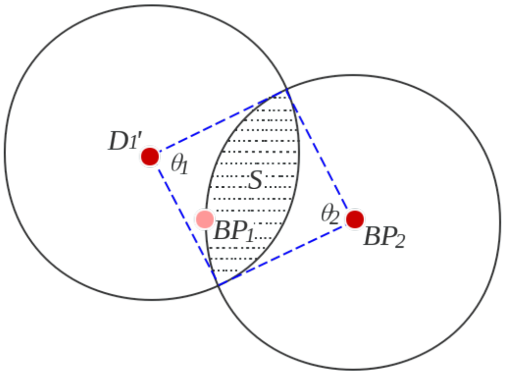


Fig. 4: The calculation of transmission probability between two nodes

candidate nodes which can be relay nodes for sending messages to the destination node.

In the proposed TERP protocol, the transmission probability is utilized to choose candidate nodes for the delivery phase. The concept of transmission probability is defined by Definition 7. When a node is going to transmit a message to a certain destination, the node figures out the proper range which is suitable for delivery and checks out nodes that meet requirements for high transmission probabilities.

Definition 7 (Transmission probability). The transmission probability $Pr^t_{i \to j}$ is the probability that a message can be successfully transmitted from node i to node j at time t. $Pr^t_{i \to \forall}$ represents the probability that node i can be heard by any node around it.

In particular, the transmission probability can be calculated by the following equations. In Fig. 4, BP_1 and BP_2 are the basic positions of node 1 and node 2, which is well defined in Definition 1. At this moment shown in Fig. 4, node 1 leaves its basic position BP_1 and located at the point $D^{'}_1$, whereas node 2 is still at the position of BP_2. The value of transmission probability is evaluated by the positions of target nodes by Eq. (1). If there are more nodes around the source node, the total probability is required to be calculated by Eq. (2).

$$\begin{aligned} Pr^t_{1 \to 2} &= \frac{S}{\pi R^2_{mov}} \\ &= \frac{S}{2\pi}\left[\frac{R^2_{trans}}{R^2_{mov}}(\theta_1 - \sin\theta_1) + (\theta_2 - \sin\theta_2)\right] \end{aligned} \tag{1}$$

$$Pr^t_{1 \to \forall} = 1 - \prod_{i=1}^{c}\left(1 - Pr^t_{1 \to i}\right) \tag{2}$$

The node that wants to deliver the message masters a list based on the range of acceptable transmission probabilities that should be set in advance. This list contains all the nodes that have the chance to be chosen as a relay node. We define this list as a candidate node list by Definition 8.

Definition 8 (Candidate node list). A candidate node list $\mathcal{C}$ contains the ID numbers of all nodes whose transmission probabilities satisfy the transmission requirement.

List $\mathcal{C}$ is obtained by comparing the transmission probability of every node which might be the relay node with the default probability value Pr_d set in advance as a condition. The process for obtaining $\mathcal{C}$ is given by Algorithm 2.

Algorithm 2 Upon the query node's acknowledgment of nodes which can be relay nodes

```
$N \leftarrow$ all nodes who can act as a relay node
$\mathcal{C} \leftarrow$ an empty set
while $N$ do
    $n \leftarrow$ one of the nodes in $N$
    $N \leftarrow N - n$
    if $Pr^t_{n \to \forall} \geq Pr_d$ then
        $\mathcal{C} \leftarrow \mathcal{C} + n$
    else
        return
    end if
end while
return $\mathcal{C}$
```

The specific meaning of Algorithm 2 is that the node chooses the candidate next hop after obtaining list N of all the possible nodes in the reachable range. Node n in set N is taken out and its transmission probability $Pr^t_{n \to \forall}$ is compared with a preset value Pr_d. If node n has a high probability of transmission, it can be selected as one of the candidate nodes in the candidate node list $\mathcal{C}$.

The candidate node list $\mathcal{C}$ should be sent to the cloud by the node who wants to know their trustworthiness and this node is named as the query node in this paper.

3.2.3 Delivery Phase

When the query node wants to select an appropriate node in candidates to be the relay node, he sends the candidate node list to the cloud. The cloud will search the cloud storage to find out the up-to-date trustworthiness τ of the nodes in the list.

The cloud will not directly send the trustworthiness value τ back to the query node. To prevent a too old trustworthiness value, the cloud first compares the timestamps of τ and list $\mathcal{C}$. If the interval between these two timestamps exceeds half of the period since AP uploading, the trustworthiness of the node will be recalculated. The freshness is defined as the interval in this paper in Definition 9. The system generally pre-sets a constant as an acceptable freshness.

Definition 9 (Freshness of trustworthiness). Freshness Δt represents the interval between the timestamps of the most recent trustworthiness record that can be found by the cloud and the current timestamp. The greater the freshness value is, the older the record is.

After receiving the list of candidate nodes from the query node, the cloud feedbacks the trustworthiness of these nodes. This process is described in detail by Algorithm 3.

In Algorithm 3, $\mathcal{C}$ is the candidate node list given by the query node according to their transmission probability. T represents the timestamp when the query node uploads the candidate node list. Note that Δt is defined well in Definition 9. E is utilized to record the up-to-date trustworthiness of the nodes in the list. The cloud server takes a node from the list and finds the corresponding trustworthiness through the node's unique identity information. To avoid

2332-7790 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBDATA.2017.2710347, IEEE Transactions on Big Data

Algorithm 3 Upon the cloud server's acquisition of query node's requirement for trustworthiness of nodes in the candidate node list $\mathcal{C}$

```
$\mathcal{C} \leftarrow$ nodes in the candidate node list
$T \leftarrow$ the current timestamp
$\Delta t \leftarrow$ pre-set acceptable freshness of trustworthiness
$E \leftarrow$ an empty set used to record the trustworthiness
evaluation values of vehicles
while $\mathcal{C}$ do
  $c \leftarrow$ one of the nodes in $\mathcal{C}$
  $\mathcal{C} \leftarrow \mathcal{C} - c$
  Search the storage for the recorded trustworthiness $\tau_c^t$
  fo node $c$ at time $t$
  if $|T - t| > \Delta t$ then
    implement re-evaluation of $c$'s trustworthiness
    $t$ of $\tau_c^t \leftarrow T$
  else
    return
  end if
  $E \leftarrow E + \tau_c^t$
end while
return $E$
```

Algorithm 4 Upon the query node's acquisition of the trustworthiness list E

```
$E \leftarrow$ the trustworthiness evaluation values of the vehicles
provided by the cloud
$m \leftarrow$ the selection of relay node
$t_m \leftarrow 0$
while $m$ do
  $e \leftarrow$ one of the nodes in $E$
  $E \leftarrow E - e$
  $t \leftarrow$ the timestamp of e
  if $t > t_m$ then
    $t_m \leftarrow t$
    $m \leftarrow e$
  else
    return
  end if
end while
return $m$
```

old trustworthiness, the cloud will compare the timestamp of the trustworthiness with the timestamp of the query. If the trustworthiness is not fresh enough, the cloud will re-evaluate the vehicle.

When the query node receives trustworthiness list E, it performs Algorithm 4 to select the node that is the most suitable to be the next hop.

In Algorithm 4, E is the record of trustworthiness of candidate nodes provided by the cloud. The algorithm aims to choose the node with the largest trustworthiness value in the candidate node list.

3.3 Trustworthiness Evaluation

The principal component analysis (PCA) tries to simplify the multi-variable data table under the principle of minimizing the loss of data information. That is to say, the PCA approach can be utilized to reduce dimensions of high dimensional variable space. In our protocol, the PCA approach is utilized to evaluate the vehicle trustworthiness according to the uploaded AP list which is actually a multi-variable data table.

The trustworthiness evaluation procedure by the PCA is as follows:

1. Standardization of the attributes

According to Definition 5. Suppose that one vehicle has h valid APs recording different attributes $\varphi_1, ..., \varphi_h$. An AP list $\mathcal{P} = \{\varphi_1, ..., \varphi_h\}$. Suppose that a total of n vehicles need to be evaluated. The value of the j-th index of the i-th evaluation object is represented by φ_{ij}. Each index value φ_{ij} is converted into a standardized value $\tilde{\varphi}_{ij}$, which is calculated by Eq. 3.

$$\tilde{\varphi}_{ij} = \frac{\varphi_{ij} - \mu_j}{s_j}, \quad (i = 1, 2, ..., n; j = 1, 2, ..., h) \tag{3}$$

where $\mu_j = \frac{1}{n}\sum_{i=1}^{n}\varphi_{ij}$ and $s_j = \frac{1}{n-1}\sum_{i=1}^{n}(\varphi_{ij} - \mu_j)^2$, $(j = 1, 2, ..., h)$. Eq. 4 represents the standardized index variable.

$$\tilde{x}_i = \frac{x_i - \mu_j}{s_j}, \quad (i = 1, 2, ..., h) \tag{4}$$

where x_i is the variable of index φ_i.

2. Calculation of the correlation coefficient matrix

In this step, the correlation coefficient matrix R is calculated according to vehicles and attribute parameters. The correlation coefficient matrix $R = (r_{ij})_{h \times h}$ can be calculated by Eq. 5.

$$r_{ij} = \frac{\sum_{k=1}^{n}\tilde{\varphi}_{ki} \cdot \tilde{\varphi}_{kj}}{n - 1}, \quad (i, j = 1, 2, ..., m) \tag{5}$$

where $r_{ii} = 1$ and $r_{ij} = r_{ji}$. Note that r_{ij} is the correlation coefficient of the i-th index and the j-th index.

3. Calculation of eigenvalues and eigenvectors

This step aims at calculating the eigenvalues $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_h \geq 0$ and the corresponding eigenvectors $u_1, u_2, ..., u_h$ of the correlation coefficient matrix R obtained in step 2, where $u_j = (u_{1j}, u_{2j}, ..., u_{hj})^T$. The new m index variables are composed of the eigenvectors as shown in Eq. 6.

$$\begin{cases} y_1 = u_{11}\tilde{x}_1 + u_{21}\tilde{x}_2 + \cdots + u_{h1}\tilde{x}_h \\ y_2 = u_{12}\tilde{x}_1 + u_{22}\tilde{x}_2 + \cdots + u_{h2}\tilde{x}_h \\ \vdots \\ y_h = u_{1h}\tilde{x}_1 + u_{2h}\tilde{x}_2 + \cdots + u_{hh}\tilde{x}_h \end{cases} \tag{6}$$

where y_z(z is from 1 to h) is defined as the z-th principal component.

4. Calculation of the contribution rate

The information contribution rate of each eigenvalue λ_j $(j = 1, 2, ..., m)$ is computed by Eq. 7 and the accumulative contribution rate of that is computed by Eq. 8.

$$b_j = \frac{\lambda_j}{\sum_{k=1}^{m}\lambda_k} \quad (j = 1, 2, ..., m) \tag{7}$$

2332-7790 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBDATA.2017.2710347, IEEE Transactions on Big Data

$$\alpha_p = \frac{\sum\limits_{k=1}^{p} \lambda_k}{\sum\limits_{k=1}^{h} \lambda_k} \tag{8}$$

Eq. 8 calculates the accumulative contribution rate of principal components $y_1, y_2, ..., y_p$. When α_p is close to 1, the first p index variables are chosen as the p principal components, instead of the original h index variables. According to the p principal components, the trustworthiness evaluation can be done to make a comprehensive analysis of the corresponding vehicle.

5. Calculation of the trustworthiness value

Finally, after the above steps, we can obtain the trustworthiness value τ queried by the nodes in the network from Eq. 9.

$$\tau = \sum_{j=1}^{p} b_j y_j \tag{9}$$

where b_j is the information contribution rate of the j-th principal component.

4 Performance Evaluation

This section aims to show the performance of the proposed protocol. First, a test of the trustworthiness evaluation process is provided to show the obtain of trustworthiness value. Secondly, ns-2 is utilized to simulate the routing protocol from three different aspects, packet delivery ratio, normalized routing overhead and average end-to-end delay in different scales of networks.

4.1 Trustworthiness Evaluation Test

In order to test whether our trustworthiness evaluation scheme is feasible, an example is utilized to elaborate the evaluation process.

Given ten different vehicles and six of their associated attribute parameters, we obtain the object of the test. The six APs we selected are as follows. Vehicle mileage is used to record the total number of miles traveled since the vehicle was used. Remaining gasoline is used to record how much gasoline the vehicle currently has. Vehicle violation number is used to record the number of times a vehicle has violated traffic rules since its inception. Vehicle failure number is used to record the number of failures of the vehicle from the date of use, such as the engine failure. Vehicle accident number records the number of accidents of the vehicle since the date of use. APh is utilized to record the frequency at which the vehicle is used as a relay node which is corresponding to the package size it has transmitted. These parameters are not absolutely independent of each other. For instance, if a vehicle has run a high number of miles, the failure rate can be relatively high. If a vehicle violates more times of traffic rules, it will have a higher probability of accidents.

We choose ten vehicles $v_1, v_2, ..., v_{10}$ as an example. In the test six parameters are considered. In order to facilitate the utilization of these parameters in the evaluation, we calculated the reciprocal of some parameters. x_1 represents the reciprocal of the vehicle mileage. x_2 shows the remaining gasoline. x_3 delegates the reciprocal of the vehicle violation number. x_4 indicates the reciprocal of the vehicle failure number. x_5 implies the reciprocal of the vehicle accident number. x_6 denotes the reciprocal of APh defined in this paper. Once there is 0 in the data, the reciprocal of the parameter is recorded as 2, since the reciprocal of zero is meaningless.

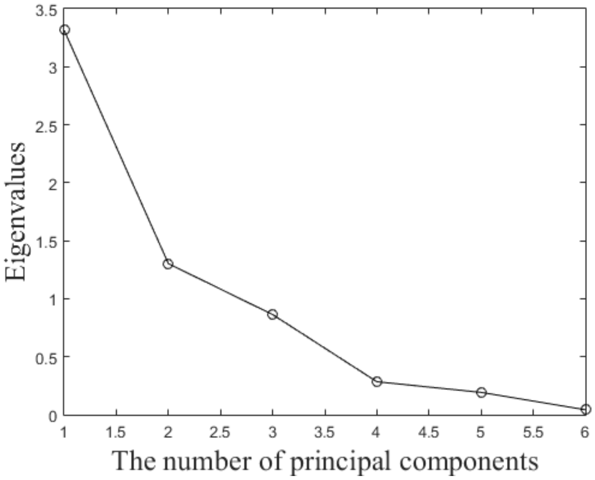


Fig. 5: The gravel map of the evaluation

The six eigenvalues of principal components are 3.3143, 1.3025, 0.8633, 0.2845, 0.1926 and 0.0428. According to the data drawn from the gravel map as shown in Fig. 5, the contribution of the first three principal components is larger. The accumulative contribution rate of the first three principal components is 91.3358%, which is sufficient to evaluate the overall condition of the vehicle.

The gravel map illustrates that the first three components can be considered as the principal components. In other words, through these three components, the cloud can basically determine the vehicle's trustworthiness which is based on the six attribute parameters in Table 3. "Ranking" represents the ranking of scores based on trustworthiness evaluation results.

From Table 3, we can figure out that the trustworthiness value τ ranks from 2.0851 to -1.0108. The vehicle v_1 achieves the highest trustworthiness value, gaining the first place. However, the 10th vehicle v_{10} receives the lowest evaluation of trustworthiness, ranking the last of the ten vehicles. It is noteworthy that the APh value of v_1 is large, which indicates that v_1 has acted as a relay node in this network for many times. The reason is that the historical big data of v_1 are excellent, which results in a high trustworthiness evaluation.

4.2 Simulation Results

Ns-2 simulator is utilized to simulate and compare among AODV, GrD-OTBR [1], DDSRP [76] and the proposed TERP. Protocols are run on Ubuntu 16.04 TLS operating system with an Intel(R) Xeon(R) CPU E5-2650 v2 of 2.60 GHz, and the reversion of network simulator is `ns-2.35`.

In order to better demonstrate the performance of the protocol, it is necessary to determine the appropriate test

2332-7790 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBDATA.2017.2710347, IEEE Transactions on Big Data

TABLE 2: Example of an AP list of ten vehicles

	x_1	x_2	x_3	x_4	x_5	x_6
v_1	0.405252067	0.96	1	2	2	0.018181818
v_2	0.2362614	0.62	0.5	0.5	2	0.024390244
v_3	0.424502271	0.21	1	0.25	2	0.02173913
v_4	0.19388863	0.55	2	0.2	1	0.018867925
v_5	0.08900202	0.89	0.2	0.5	0.5	0.019607843
v_6	0.070154269	0.75	0.5	0.1667	0.3333	0.024390244
v_7	0.080892405	0.27	0.125	0.1428	0.125	0.025
v_8	0.070347234	0.98	0.5	0.0833	0.3333	0.023809524
v_9	0.074282616	0.33	0.3333	0.1	0.5	0.027027027
v_{10}	0.065161013	0.12	0.125	0.0909	0.3333	0.038461538

TABLE 3: The result of trustworthiness evaluation

Vehicle identity	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9	v_{10}
τ	2.0851	0.6627	1.1983	0.4384	-0.4074	-0.6838	-0.8550	-0.7273	-0.7001	-1.0108
Ranking	1	3	2	4	5	6	9	8	7	10

TABLE 4: Parameters used in simulation

Parameter	**Value**
Number of nodes	20, 50, 100 and 150
Mobility model	Described by IPNs
Mac	IEEE 802.11 DCP
Traffic source	CBR for UDP-based traffic
Node speed	$0 \sim 20\text{m/s}$
Propagation model	Two-ray ground reflection
Simulation time	1000 seconds
Data transmission rate	2 Mbps

environment in our simulations. We set 20, 50, 100 and 150 nodes moving in a square area of 1000m × 1000m. These nodes move in a certain range around their basic positions, which is well defined by IPNs. Each node picks a random spot in the range around them and moves there with a speed uniformly distributed between $0 \sim 20$ meters/sec. A two-ray ground reflection propagation channel is considered. The data transmission rate is set to be 2 Mbps. Most other parameters use ns-2 defaults. Table 4 shows some significant parameters.

To analyze the proposed routing protocol, we define three different metrics as follows. The packet delivery ratio (PDR) represents the probability that a delivered packet is successfully received. Normalized routing overhead (NRO) indicates the overhead incurred by the network during the transmission of information. Average end-to-end delay (E2D) shows the average interval of sending and receiving data packets. The reason for test in different sizes of networks is to verify whether our protocol can adapt to different scenarios. Fig. 6 shows the plotted data from our simulated environment.

Based on graphs and data, we make the following brief analysis to demonstrate that our protocol is applicable to an incompletely predictable vehicle ad hoc network.

Viewed from Fig. 6(a), the PDR of the four protocols can be maintained above 80% in most cases, which indicates that the four protocols can guarantee the success of the routing. However, AODV does not perform well in small-scale networks, which is not suitable for vehicle networks that need to be partitioned into small regions for transmission. TERP's delivery ratio is higher than other protocols. Due to high speeds of vehicles in the simulated network, TERP outperforms the GrD-OTBR and DDSRP. With the assistance of the cloud, once more historical big data is available, TERP can more accurately transmit data packets to the specified terminals.

Observed from Fig. 6(b), the NRO of the four protocols are not very high. Compared with Fig. 6(a), DDSRP and GrD-OTBR performs well in terms of PDR but not well in NRO. The TERP has lower overhead than DDSRP and GrD-OTBR but more overhead than AODV. The TERP introduces the cloud to routing, but little traditional routing duties are handed over to the cloud instead of being dealt with by the nodes themselves. The TERP only requires the cloud to give a reasonable value of trustworthiness, which reduces the abuse of broadcast packets and the frequency of use of nodes. So that the overhead in the network has been reduced.

Fig. 6(c) shows that the E2D of the proposed protocol is close to that of AODV in networks with big scales. TERP is able to obtain a smaller delay because, with the help of the cloud, the node can determine the next hop node more quickly.

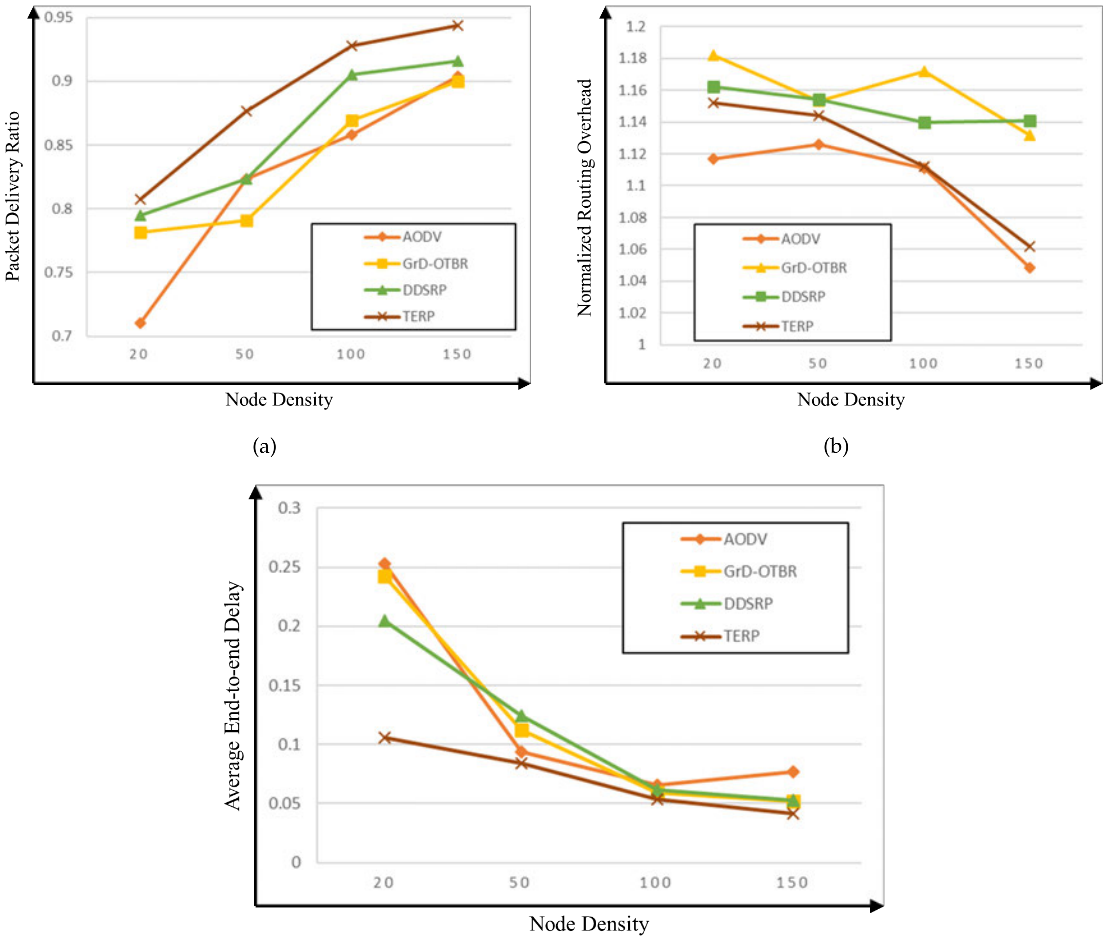
In general, due to a more simple and efficient node trustworthiness evaluation mechanism, TERP are more excellent in the above three aspects of the performance of the route.

5 Conclusions

In this paper, we proposed a novel routing protocol named trustworthiness evaluation-based routing protocol (TERP) for incompletely predictable vehicular ad hoc networks. Due to the inherent characteristics of this type of networks, it is particularly important to design a trusted route transmission method that is suitable for utilization in vehicle networks. For that reason, attribute parameters of vehicles are listed and sent to the cloud as the evaluation basis. By filtering according to transmission probabilities, candidate nodes are aggregated into a list and sent to the cloud. The cloud looks up and feeds back the latest trustworthiness evaluations of these nodes. Based on the trustworthiness

2332-7790 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBDATA.2017.2710347, IEEE Transactions on Big Data



(a)

(b)

(c)

Fig. 6: Performance evaluation: (a) packet delivery ratio vs node density; (b) normalized routing overhead vs node density; (c) average end-to end delay vs node density

provided by the cloud, nodes in the network perform the needed routing and packet delivery.

An evaluation test has been implemented and showed that our evaluation process is feasible. Moreover, the simulation of the proposed new protocol indicates that TERP is able to maintain a high packet delivery ratio, with low overhead and end-to-end delay. In order to enhance the security method of the proposed protocol, in our future work, we suppose to propose a reliable and secure trustworthiness evaluation system and further improve the trustworthiness evaluation scheme by supporting also privacy and anonymity [77], [78].

Acknowledgments

This work is supported by the National Natural Science Foundation of China under Grant No. 61672295, No. 61300237, No. U1405254, the State Key Laboratory of Information Security under Grant No. 2017-MS-10, the 2015 Project of Six Personnel in Jiangsu Province under Grant No. R2015L06, the CICAEET fund, and the PAPD fund.

References

[1] J. Shen, C. Wang, A. Wang, X. Sun, S. Moh, and P. C. Hung, "Organized topology based routing protocol in incompletely predictable ad-hoc networks," *Computer Communications*, 2016, doi: 10.1016/j.comcom.2016.07.009.

[2] G. Ping, W. Jin, H. G. Xue, S. K. Chang, and J. U. Kim, "A variable threshold-value authentication architecture for wireless mesh networks," *Journal of Internet Technology*, vol. 15, no. 6, pp. 929–935, 2014.

[3] J. Shen, H. Tan, J. Wang, J. Wang, and S. Lee, "A novel routing protocol providing good transmission reliability in underwater sensor networks," *Journal of Internet Technology*, vol. 16, no. 1, pp. 171–178, 2015.

[4] D. He, S. Zeadally, N. Kumar, and J. H. Lee, "Anonymous authentication for wireless body area networks with provable security," *IEEE Systems Journal*, 2016, doi: 10.1109/JSYST.2016.2544805.

[5] J. Shen, H. Tan, S. Moh, and I. Chung, "Enhanced secure sensor association and key management in wireless body area networks," *Journal of Communications & Networks*, vol. 17, no. 5, pp. 453–462, 2015.

[6] M. B. Abdullahi and G. Wang, "A Lightweight Anonymous On-demand Routing Scheme in Wireless Sensor Networks," in *2012 IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications*, June 2012, pp. 978–985.

[7] Q. Jiang, J. Ma, F. Wei, Y. Tian, J. Shen, and Y. Yang, "An untraceable temporal-credential-based two-factor authentication scheme using ecc for wireless sensor networks," *Journal of Network & Computer Applications*, vol. 76, pp. 37–48, 2016.

2332-7790 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBDATA.2017.2710347, IEEE Transactions on Big Data

[8] S. Xie and Y. Wang, "Construction of tree network with limited delivery latency in homogeneous wireless sensor networks," *Wireless Personal Communications*, vol. 78, no. 1, pp. 231–246, 2014.

[9] T. Ma, J. Zhou, M. Tang, Y. Tian, A. Al-Dhelaan, M. Al-Rodhaan, and S. Lee, "Social network and tag sources based augmenting collaborative recommender system," *Ieice Transactions on Information & Systems*, vol. 98, no. 4, pp. 902–910, 2015.

[10] Z. Fu, X. Wu, C. Guan, X. Sun, and K. Ren, "Towards efficient multi-keyword fuzzy search over encrypted outsourced data with accuracy improvement," *IEEE Transactions on Information Forensics and Security*, 2016, doi: 10.1109/TIFS.2016.2596138.

[11] Z. Pan, Y. Zhang, and S. Kwong, "Efficient motion and disparity estimation optimization for low complexity multiview video coding," *IEEE Transactions on Broadcasting*, vol. 61, no. 2, pp. 166–176, 2015.

[12] T. V. Le, R. Oentaryo, S. Liu, and H. C. Lau, "Local Gaussian Processes for Efficient Fine-Grained Traffic Speed Prediction," *IEEE Transactions on Big Data*, vol. PP, no. 99, pp. 1–1, 2016.

[13] W. Chen, L. Jian, and S.-Y. Kuo, "Video-based on-road driving safety system with lane detection and vehicle detection," in *2012 12th International Conference on ITS Telecommunications*, Nov 2012, pp. 537–541.

[14] D. Zhang, T. He, S. Lin, S. Munir, and J. A. Stankovic, "Taxi-Passenger-Demand Modeling Based on Big Data from a Roving Sensor Network," *IEEE Transactions on Big Data*, vol. PP, no. 99, pp. 1–1, 2016.

[15] S. Sarkar, S. Chawla, S. P. Parambath, J. Srivastava, H. Hammady, F. Filali, W. Znaidi, and J. Borge-Holthoefer, "Effective Urban Structure Inference from Traffic Flow Dynamics," *IEEE Transactions on Big Data*, vol. PP, no. 99, pp. 1–1, 2017.

[16] M. Ota, H. Vo, C. Silva, and J. Freire, "STaRS: Simulating Taxi Ride Sharing at Scale," *IEEE Transactions on Big Data*, vol. PP, no. 99, pp. 1–1, 2016.

[17] A. Castiglione, A. D. Santis, B. Masucci, and F. Palmieri, "Cryptographic hierarchical access control for dynamic structures," *IEEE Transactions on Information Forensics & Security*, vol. 11, no. 10, pp. 2349–2364, 2016.

[18] G. Carullo, A. Castiglione, G. Cattaneo, A. De Santis, U. Fiore, and F. Palmieri, "FeelTrust: Providing Trustworthy Communications in Ubiquitous Mobile Environment," in *2013 IEEE 27th International Conference on Advanced Information Networking and Applications (AINA)*, March 2013, pp. 1113–1120.

[19] X. Wen, L. Shao, Y. Xue, and W. Fang, "A rapid learning algorithm for vehicle classification," *Information Sciences*, vol. 295, no. 1, pp. 395–406, 2015.

[20] X. Chen, J. Li, J. Ma, Q. Tang, and W. Lou, "New algorithms for secure outsourcing of modular exponentiations," *IEEE Transactions on Parallel & Distributed Systems*, vol. 25, no. 9, pp. 2386–2396, 2014.

[21] Q. Jiang, J. Ma, and F. Wei, "On the security of a privacy-aware authentication scheme for distributed mobile cloud computing services," *IEEE Systems Journal*, 2016, doi: 10.1109/JSYST.2016.2574719.

[22] Y. Ren, J. Shen, J. Wang, J. Han, and S. Lee, "Mutual verifiable provable data auditing in public cloud storage," *Journal of Internet Technology*, vol. 16, no. 2, pp. 317–323, 2015.

[23] Z. Fu, K. Ren, J. Shu, and X. Sun, "Enabling personalized search over encrypted outsourced data with efficiency improvement," *IEEE Transactions on Parallel & Distributed Systems*, vol. 27, no. 9, pp. 2546–2559, 2016.

[24] B. Martini and K.-K. R. Choo, "Cloud storage forensics: ownCloud as a case study," *Digital Investigation*, vol. 10, no. 4, pp. 287 – 299, 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1742287613000911

[25] B. Martini and K. K. R. Choo, "Cloud Forensic Technical Challenges and Solutions: A Snapshot," *IEEE Cloud Computing*, vol. 1, no. 4, pp. 20–25, Nov 2014.

[26] B. Martini and K.-K. R. Choo, "Distributed filesystem forensics: XtreemFS as a case study," *Digital Investigation*, vol. 11, no. 4, pp. 295 – 313, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1742287614000942

[27] D. Quick and K.-K. R. Choo, "Google Drive: Forensic analysis of data remnants," *Journal of Network and Computer Applications*, vol. 40, pp. 179 – 193, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1084804513002051

[28] ——, "Dropbox analysis: Data remnants on user machines," *Digital Investigation*, vol. 10, no. 1, pp. 3 – 18, 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S174228761300011X

[29] Z. Fu, X. Sun, Q. Liu, L. Zhou, and J. Shu, "Achieving efficient cloud search services: Multi-keyword ranked search over encrypted cloud data supporting parallel computing," *Ieice Transactions on Communications*, vol. E98.B, no. 1, pp. 190–200, 2015.

[30] Q. Jiang, M. K. Khan, X. Lu, J. Ma, and D. He, "A privacy preserving three-factor authentication protocol for e-health clouds," *Journal of Supercomputing*, vol. 72, no. 10, pp. 3826–3849, 2016.

[31] C. F. Lai, S. Zeadally, J. Shen, and Y. X. Lai, "A cloud-integrated appliance recognition approach over internet of things," *Journal of Internet Technology*, vol. 16, no. 7, pp. 1157–1168, 2015.

[32] Z. Xia, X. Wang, X. Sun, and Q. Wang, "A secure and dynamic multi-keyword ranked search scheme over encrypted cloud data," *IEEE Transactions on Parallel & Distributed Systems*, vol. 27, no. 2, pp. 340–352, 2015.

[33] R. Ranjan, L. Wang, A. Y. Zomaya, D. Georgakopoulos, X. H. Sun, and G. Wang, "Recent advances in autonomic provisioning of big data applications on clouds," *IEEE Transactions on Cloud Computing*, vol. 3, no. 2, pp. 101–104, April 2015.

[34] P. Bedi and V. Jindal, "Use of Big Data technology in Vehicular Ad-hoc Networks," in *2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, Sept 2014, pp. 1677–1683.

[35] Y. R. B. Al-Mayouf, M. Ismail, N. F. Abdullah, A. W. A. Wahab, O. A. Mahdi, S. Khan, and K.-K. R. Choo, "Efficient and Stable Routing Algorithm Based on User Mobility and Node Density in Urban Vehicular Network," *PLOS ONE*, vol. 11, no. 11, pp. 1–24, 11 2016. [Online]. Available: https://doi.org/10.1371/journal.pone.0165966

[36] M. R. Jabbarpour, H. Zarrabi, R. H. Khokhar, S. Shamshirband, and K.-K. R. Choo, "Applications of computational intelligence in vehicle traffic congestion problem: a survey," *Soft Computing*, pp. 1–22, 2017. [Online]. Available: http://dx.doi.org/10.1007/s00500-017-2492-z

[37] D. He, C. Chen, S. Chan, and J. Bu, "A distributed trust evaluation model and its application scenarios for medical sensor networks," *IEEE Transactions on Information Technology in Biomedicine A Publication of the IEEE Engineering in Medicine & Biology Society*, vol. 16, no. 6, pp. 1164–75, 2012.

[38] A. Zhou, J. Li, Q. Sun, C. Fan, T. Lei, and F. Yang, "A security authentication method based on trust evaluation in vanets," *EURASIP Journal on Wireless Communications and Networking*, vol. 2015, no. 1, pp. 1–8, 2015.

[39] D. Filev, J. Lu, and D. Hrovat, "Future mobility: Integrating vehicle control with cloud computing," *Mechanical Engineering*, vol. 135, no. 3, pp. S18–S24, 03 2013.

[40] J. Zhang, F. Y. Wang, K. Wang, W. H. Lin, X. Xu, and C. Chen, "Data-driven intelligent transportation systems: A survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 4, pp. 1624–1639, December 2011.

[41] C. Gosman, T. Cornea, C. Dobre, F. Pop, and A. Castiglione, "Controlling and filtering users data in Intelligent Transportation System," *Future Generation Computer Systems*, p. In Press, 2016. [Online]. Available: http://dx.doi.org/10.1016/j.future.2016.12.014

[42] ——, *Putting the User in Control of the Intelligent Transportation System*. Cham: Springer International Publishing, 2016, pp. 231–246. [Online]. Available: 10.1007/978-3-319-40253-6_14

[43] J. Shen, C. Wang, A. Castiglione, D. Liu, and C. Esposito, *Greedy Probability-Based Routing Protocol for Incompletely Predictable Vehicular Ad-hoc Network*. Cham: Springer International Publishing, 2016, pp. 208–217. [Online]. Available: 10.1007/978-3-319-49145-5_21

[44] F. Palmieri and A. Castiglione, "Condensation-based routing in mobile ad-hoc networks," *Mobile Information Systems*, vol. 8, no. 3, pp. 199–211, 2012. [Online]. Available: 10.3233/MIS-2012-0140

[45] G. Theodorakopoulos and J. S. Baras, "On trust models and trust evaluation metrics for ad hoc networks," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 2, pp. 318–328, 2006.

[46] W. Jiang, G. Wang, M. Z. A. Bhuiyan, and J. Wu, "Understanding Graph-Based Trust Evaluation in Online Social Networks: Methodologies and Challenges," *ACM Comput. Surv.*, vol. 49, no. 1, pp. 10:1–10:35, May 2016. [Online]. Available: http://doi.acm.org/10.1145/2906151

[47] W. Jiang, G. Wang, and J. Wu, "Generating trusted graphs for trust

2332-7790 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBDATA.2017.2710347, IEEE Transactions on Big Data

evaluation in online social networks," *Future Generation Computer Systems*, vol. 31, no. 1, pp. 48–58, 2014.

[48] A. Castiglione, A. D. Santis, and B. Masucci, "Key indistinguishability vs.strong key indistinguishability for hierarchical key assignment schemes," *IEEE Transactions on Dependable & Secure Computing*, vol. 13, no. 4, pp. 451–460, 2015.

[49] X. Chen, F. Zhang, W. Susilo, H. Tian, J. Li, and K. Kim, "Identity-based chameleon hashing and signatures without key exposure," *Information Sciences*, vol. 265, no. 5, pp. 198–210, 2014.

[50] C. Esposito, A. Castiglione, F. Palmieri, and M. Ficco, "Trust management for distributed heterogeneous systems by using linguistic term sets and hierarchies, aggregation operators and mechanism design," *Future Generation Computer Systems*, p. In Press, 2015. [Online]. Available: http://dx.doi.org/10.1016/j.future.2015.12.004

[51] S. Chen, G. Wang, G. Yan, and D. Xie, "Multi-dimensional fuzzy trust evaluation for mobile social networks based on dynamic community structures," *Concurrency and Computation: Practice and Experience*, p. In Press, 2016, cpe.3901. [Online]. Available: http://dx.doi.org/10.1002/cpe.3901

[52] A. Mohsenzadeh, H. Motameni, and J. E. Meng, "A new trust evaluation algorithm between cloud entities based on fuzzy mathematics," *International Journal of Fuzzy Systems*, vol. 18, no. 4, pp. 659–672, 2016.

[53] V. V. Rajendran and S. Swamynathan, "Hybrid model for dynamic evaluation of trust in cloud services," *Wireless Networks*, vol. 22, no. 6, pp. 1807–1818, 2016.

[54] W. Jiang, J. Wu, F. Li, G. Wang, and H. Zheng, "Trust Evaluation in Online Social Networks Using Generalized Network Flow," *IEEE Transactions on Computers*, vol. 65, no. 3, pp. 952–963, March 2016.

[55] J. Wang, Y. Zhang, Y. Wang, and X. Gu, "Rprep: A robust and privacy-preserving reputation management scheme for pseudonym-enabled vanets," *International Journal of Distributed Sensor Networks*, vol. 2016, no. 3-4, pp. 1–15, 2016.

[56] F. Pop, C. Dobre, B. C. Mocanu, O. M. Citoteanu, and F. Xhafa, "Trust models for efficient communication in mobile cloud computing and their applications to e-commerce," *Enterprise Information Systems*, vol. 10, no. 9, pp. 982–1000, 2016.

[57] A. Aburumman, W. J. Seo, C. Esposito, A. Castiglione, R. Islam, and K.-K. R. Choo, , "A secure and resilient cross-domain SIP solution for MANETs using dynamic clustering and joint spatial and temporal redundancy," *Concurrency and Computation: Practice and Experience*, pp. n/a–n/a, 2016, cpe.3978. [Online]. Available: http://dx.doi.org/10.1002/cpe.3978

[58] A. Aburumman and K.-K. R. Choo, *A Domain-Based Multi-cluster SIP Solution for Mobile Ad Hoc Network*. Cham: Springer International Publishing, 2015, pp. 267–281. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-23802-9_21

[59] A. Aburumman, W. J. Seo, R. Islam, M. K. Khan, and K.-K. R. Choo, *A Secure Cross-Domain SIP Solution for Mobile Ad Hoc Network Using Dynamic Clustering*. Cham: Springer International Publishing, 2015, pp. 649–664. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-28865-9_43

[60] M. Ge and K.-K. R. Choo, *A Novel Hybrid Key Revocation Scheme for Wireless Sensor Networks*. Cham: Springer International Publishing, 2014, pp. 462–475. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-11698-3_35

[61] J. a. G. Filho, A. Patel, B. L. A. Batista, and J. C. Júnior, "A systematic technical survey of dtn and vdtn routing protocols," *Computer Standards & Interfaces*, vol. 48, pp. 139–159, 2016.

[62] A. Shokrollahi and M. N. Maybodi, "An energy-efficient clustering algorithm using fuzzy c-means and genetic fuzzy system for wireless sensor network," *Journal of Circuits Systems & Computers*, vol. 26, no. 1, 2017, doi:10.1142/S0218126617500049.

[63] Y. Zhang, X. Sun, and B. Wang, "Efficient algorithm for k-barrier coverage based on integer linear programming," *China Communications*, vol. 13, no. 7, pp. 16–23, 2016.

[64] Y. Ahmadi, N. Neda, and R. Ghazizadeh, "Range free localization in wireless sensor networks for homogeneous and non-homogeneous environment," *IEEE Sensors Journal*, vol. 16, no. 22, pp. 8018–8026, 2016.

[65] A. M. Krishnan and P. G. Kumar, "An effective clustering approach with data aggregation using multiple mobile sinks for heterogeneous wsn," *Wireless Personal Communications*, vol. 90, no. 2, pp. 423–434, 2016.

[66] M. Itoh, D. Yokoyama, M. Toyoda, Y. Tomita, S. Kawamura, and M. Kitsuregawa, "Visual Exploration of Changes in Passenger Flows and Tweets on Mega-City Metro Network," *IEEE Transactions on Big Data*, vol. 2, no. 1, pp. 85–99, March 2016.

[67] Y. Zheng, W. Wu, Y. Chen, H. Qu, and L. M. Ni, "Visual Analytics in Urban Computing: An Overview," *IEEE Transactions on Big Data*, vol. 2, no. 3, pp. 276–296, Sept 2016.

[68] E. Celikten, G. L. Falher, and M. Mathioudakis, "Modeling Urban Behavior by Mining Geotagged Social Data," *IEEE Transactions on Big Data*, vol. PP, no. 99, pp. 1–1, 2016.

[69] H. Nguyen, W. Liu, and F. Chen, "Discovering Congestion Propagation Patterns in Spatio-Temporal Traffic Data," *IEEE Transactions on Big Data*, vol. PP, no. 99, pp. 1–1, 2016.

[70] Y. Ding, Y. Li, K. Deng, H. Tan, M. Yuan, and L. M. Ni, "Detecting and Analyzing Urban Regions with High Impact of Weather Change on Transport," *IEEE Transactions on Big Data*, vol. PP, no. 99, pp. 1–1, 2016.

[71] T. S. Lin, C. H. Chien, T. H. Chang, and S. Y. Kuo, "Quantum signature scheme for vehicular networks using entangled states," in *2011 Carnahan Conference on Security Technology*, Oct 2011, pp. 1–6.

[72] A. M. Dziekoski and R. O. Schoeneich, "Dtn routing algorithm for networks with nodes social behavior," *International Journal of Computers Communications & Control*, vol. 11, no. 4, pp. 457–471, 2016.

[73] P. Goel, L. Kulik, and K. Ramamohanarao, "Optimal Pick up Point Selection for Effective Ride Sharing," *IEEE Transactions on Big Data*, vol. PP, no. 99, pp. 1–1, 2016.

[74] M. Gochoo, D. Bayanduuren, U. Khuchit, G. Battur, T. H. Tan, S. Y. Kuo, and S. C. Huang, "Design and application of novel morphological filter used in vehicle detection," in *2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS)*, June 2016, pp. 1–5.

[75] A. Castiglione, R. De Prisco, A. De Santis, U. Fiore, and F. Palmieri, "A botnet-based command and control approach relying on swarm intelligence," *Journal of Network and Computer Applications*, vol. 38, pp. 22–33, 2014. [Online]. Available: http://dx.doi.org/10.1016/j.jnca.2013.05.002

[76] J. Shen, W. Chen, C. Lai, A. Wang, and H. Chao, "Direction density-based secure routing protocol for healthcare data in incompletely predictable networks," *IEEE Access*, 2016, doi: 10.1109/ACCESS.2016.2637887.

[77] C. Shi, X. Luo, P. Traynor, M. H. Ammar, and E. W. Zegura, "ARDEN: Anonymous networking in delay tolerant networks," *Ad Hoc Networks*, vol. 10, no. 6, pp. 918 – 930, 2012.

[78] Z. Su, X. Luo, W. Wu, and J. Cao, "ACC: Anonymous Cooperative Caching in Wireless Ad Hoc Networks," in *2013 IEEE 10th International Conference on Mobile Ad-Hoc and Sensor Systems*, Oct 2013, pp. 64–72.

Jian Shen received the B.E. degree from Nanjing University of Information Science and Technology, Nanjing, China in 2007 and the M.E. and Ph.D. degrees in Computer Science from Chosun University, Gwangju, Korea in 2009 and 2012, respectively. Since late 2012, he has been a Full Professor in the School of Computer and Software at Nanjing University of Information Science and Technology, Nanjing, China. His research interests include computer networking, security systems, mobile computing and networking, ad-hoc networks and systems, and ubiquitous sensor networks.

Chen Wang received the B.E. degree in 2016 and is currently working toward the M.E. degree at Nanjing University of Information Science and Technology, Nanjing, China. He focuses on information security and incompletely predictable ad hoc networks. His research interests include information security, ad-hoc networks and systems, and wireless sensor networks.

2332-7790 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBDATA.2017.2710347, IEEE Transactions on Big Data

Aniello Castiglione (S'04-M'08) received the Ph.D. degree in computer science from the University of Salerno, Italy. Actually, he is an Adjunct Professor with the University of Salerno, and University of Naples "Federico II", Italy. He received the Italian national habilitation as an Associate Professor of Computer Science. He serves as a Reviewer for around 50 international journals, He acted as a Guest Editor for several journals. Ha also serves as a Managing Editor of two international journals, an Editor of several editorial boards. He served as a Program Chair and TPC Member of around 90 international conferences. He has authored more than 140 papers in international journals and conferences. He has been involved in forensic investigations, collaborating with several Law Enforcement Agencies as a Consultant. His current research interests include information forensics, digital forensics, security and privacy on cloud, communication networks, and applied cryptography.

Dengzhi Liu received the B.S. degree in 2014 and is currently working toward the M.E. degree at Nanjing University of Information Science and Technology, Nanjing, China. He focuses on the security and privacy issues in cloud environment. His current research interests are fine-grained cloud data access control, privacy preserving search schemes in cloud and auditing protocols for cloud storage.

Christian Esposito (S'06-M'09) received the Ph.D. degree in computer engineering and automation from the University of Napoli "Federico II", Italy. Actually, he is adjunct professor at the University of Naples "Federico II", Italy, and at the University of Salerno, Italy, where he is also a research fellow. He regularly serves as a reviewer and guest editor for several international journals, and conferences (with about 200 reviews being done). He has been a PC member or involved in the organization of about 40 international conferences/workshops. He also serves as guest editor for several journals, and member of two editorial boards. His research interests include reliable and secure communications, middleware, distributed systems, positioning systems, multi-objective optimization, and game theory.

2332-7790 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.