PRIF: A Privacy-Preserving Interest-Based Forwarding Scheme for Social Internet of Vehicles

Liehuang Zhu, Member, IEEE, Chuan Zhang, Chang Xu, Senior Member, IEEE, Rixin Xu, Kashif Sharif, Member, IEEE, and Mohsen Guizani, Fellow, IEEE

Abstract—Recent advances in socially aware networks (SANs) have allowed its use in many domains, out of which the Social Internet of Vehicles (SIOV) is of prime importance. SANs can provide a promising routing and forwarding paradigm for SIOV by using interest-based communication. Though able to improve the forwarding performance, existing interest-based schemes fail to consider the important issue of protecting users’ interest information. In this paper, we propose a privacy-preserving interest-based forwarding scheme (PRIF) for SIOV, which not only protects the interest information but also improves the forwarding performance. We propose a privacy-preserving authentication protocol to recognize communities among mobile nodes. During data routing and forwarding, a node can know others’ interests only if they are affiliated with the same community. Moreover, to improve forwarding performance, a new metric community energy is introduced to indicate vehicular social proximity. Community energy is generated when two nodes encounter one another and information is shared among them. PRIF considers this energy metric to select forwarders toward the destination node or the destination community. Security analysis indicates PRIF can protect nodes’ interest information. In addition, extensive simulations have been conducted to demonstrate that PRIF outperforms the existing algorithms, including the BEEINFO, Epidemic, and ProPHET.

Index Terms—Community, forwarding, interest, privacy-preserving, Social Internet of Vehicles (SIOV).

I. INTRODUCTION

The recent proliferation of mobile devices (e.g., mobile phones, vehicle onboard equipment, tablets, etc.) has changed the future of communication and services [1]. Due to the inseparable bond between mobile devices and their human carriers, social relationships and users’ mobility aspects are exploited in various research fields, such as opportunistic networks (OppNets) [2], vehicular ad-hoc networks [3], and delay tolerant networks [4]. This emerging network paradigm, also known as socially aware networking (SAN), is able to take advantage of users’ social properties, and further uses them as a main design ingredient for the Social Internet of Vehicles (SIOV), a communication network where vehicles behave “socially” [5], [6].

SIOV works similarly to OppNets and DTNs, as they all lack end-to-end fixed paths from the source to the destination, they utilize store-carry-forward paradigm for such services. When there is a need for data dissemination, a key problem for these networks is to predict the future encounter opportunity. Nevertheless, the difference between SIOV and OppNets (or DTNs) is that SIOV considers social properties of devices to solve the data routing and forwarding problems and challenges.

Fig. 1 depicts two social levels in vehicular environment. Vehicles/individuals carrying mobile devices (e.g., smartphone, smart watch, digital camera, etc.) construct SIOV when they are in communication range, and their inherent social ties determine virtual social networks. Generally, social relationships are relatively stable and change less frequently than transmission links among mobile devices. Therefore, it is crucial to take advantage of mobile devices’ social properties to make smarter forwarding decisions.

Manuscript received March 15, 2018; accepted May 29, 2018. Date of publication June 12, 2018; date of current version August 9, 2018. This work was supported by the National Natural Science Foundation of China under Grant 61402037 and Grant 61272512. (Corresponding author: Chang Xu.)

L. Zhu, C. Zhang, C. Xu, R. Xu, and K. Sharif are with the Beijing Engineering Research Center of Massive Language Information Processing and Cloud Computing Application, School of Computer Science and Technology, Beijing Institute of Technology, Beijing 100081, China (e-mail: liehuangz@bit.edu.cn; chuanz@bit.edu.cn; xuchang@bit.edu.cn; xurixin@bit.edu.cn; 7620160009@bit.edu.cn).

X. Du is with the Department of Computer and Information Sciences, Temple University, Philadelphia, PA 19122 USA (e-mail: dxj2005@gmail.com). M. Guizani is with the Department of Electrical and Computer Engineering, University of Idaho, Moscow, ID 83844 USA (e-mail: mguizani@ieee.org). Digital Object Identifier 10.1109/JIOT.2018.2846653
Recently, a series of social-based routing protocols, [7]–[17] have been proposed. Most of them adopt the notion of “community” to make forwarding decisions. Specifically, mobile nodes can be divided into different communities based on their contact frequency or social relationships. It is generally agreed that members in a same community meet each other more often than others in different communities. Thus, a forwarding decision usually relies on how to construct a community and choose suitable forwarders. For example, LocalCom [7] detects communities using neighboring graphs, while Gently [8] chooses forwards based on CAR-like [9] and Label [10] protocols. Communities in these schemes can be obtained from historical records, such as encounter frequency, encounter length, and separation time. However, they ignore nodes’ inherent social relationships, especially considering that mobile nodes are always carried and used by people. Many forwarding schemes [10]–[14] have been proposed based on social network metrics. For example, Label [13] and Group [14] deliver messages only if message carriers meet the members within the same social community of the destination node, while Bubble Rap [11] uses a hierarchical community structure and forwards data if a node holds higher centrality. These social-based works utilize social relationships to make better forwarding strategies. Nevertheless, the drawback in these schemes is that the cost to form and maintain communities. However, they ignore nodes’ inherent social relationships, especially considering that mobile nodes are always carried and used by people. Many forwarding schemes [10]–[14] have been proposed based on social network metrics. For example, Label [13] and Group [14] deliver messages only if message carriers meet the members within the same social community of the destination node, while Bubble Rap [11] uses a hierarchical community structure and forwards data if a node holds higher centrality. These social-based works utilize social relationships to make better forwarding strategies. Nevertheless, the drawback in these schemes is that the cost to form and maintain communities.

In this paper, we propose PRIF: a privacy-preserving interest-based forwarding scheme to protect the sensitive interest information and improve the forwarding efficiency for SIOV. To summarize, the contributions of PRIF include the following.

1) We classify communities based on personal interests. Inspired by BEEINFO-D&S and general laws in practical physics, a novel social metric community energy is introduced to measure the social ability of a mobile node to forward messages to others. Generally, community energy is generated by node encounters. Specifically, a node will establish inter-community energy to the encountering node, if it is within the same community. Otherwise, intracommunity energy will be built toward the encountering node community. Therefore, a better forwarder should be a node with higher intercommu-

2) The interest information is private, and it is dangerous to deliver it to others. Thus, we take advantage of signature-based envelopes and design a privacy-preserving authentication protocol. In this way, a node can recognize the members coming from which communities. However, it cannot know the interests of the members unless they are affiliated to the same group with it.

3) Extensive simulation analysis has been conducted to compare PRIF with several existing schemes. Specifically, compared with two representative schemes, i.e., Epidemic [18] and PRoPHET [19], PRIF performs better in message delivery, overhead, and hop counts. In addition, the proposed scheme outperforms the existing interest-based scheme BEEINFO-D&S in delivery ratio and overhead.

The rest of this paper is organized as follows. In Section II, we review existing social-based data forwarding protocols. Section III describes the detailed design of PRIF for SIOV. In Section IV, we give the security analysis of PRIF and in Section V, we analyze the performance. Finally, we conclude this paper in Section VI.

II. RELATED WORKS

In recent years, feature extraction has received considerable attention in various fields [20]–[24]. In VANETs, to adapt to the frequently changing topology and high-speed mobility [25]–[27], social property as a special feature among people, plays an important role in designing routing algorithms. Many social-based routing algorithms [7]–[17] have been proposed which are roughly based on two main aspects: 1) behavior regularity and 2) community information.

Behavior regularity focuses on individuals’ behaviors. It relies on the principle that people usually have repeated mobility patterns. In real world, people often hold similar mobility patterns. For instance, they usually follow similar paths from their home to offices during weekdays. Accordingly, regular behaviors can be used to predict the future encounter probability, and works in [12], [16], and [17] have proposed algorithms based on this metric. For example, SimBet [12] constructs a utility function by exploiting the similarity and betweenness centrality to the destination with the help of an ego network. To describe nodes’ relationships, SimBetTS [16] considers another important factor (i.e., social tie strength) to choose more suitable forwarders. Moreover, HiBOp [17] can automatically learn users’ behaviors and social relations to execute the forwarding process.

Another important basis to support social-based routing is community information. Generally, communities can be constructed by individuals’ interests or encounter frequency, and it is generally agreed that nodes coming from the same community will meet each other more frequently. A series of routing algorithms have been proposed based on this metric [7], [8], [10], [11], [13]–[15]. The simplest community-based routing method is LABEL [10], in which messages are only delivered to the nodes in the destination community. Similar to the scheme in [10], Li et al. [14] proposed a new community-based scheme. However, [10] and [14] do not take nodes’ relationships into consideration. To solve the problem, Bubble Rap [11] and Friendship-based [13] were proposed to select nodes with higher social centrality as relay nodes. Nevertheless, these schemes suffer a common drawback that the cost to form and maintain a socially aware
overlay is extremely high. Besides the intracommunity routing, intercommunity routings were considered in [7] and [8]. LocalCom [7] utilized the encounter history such as encounter frequency, encounter period and separation time to construct a neighboring graph which was further utilized to detect communities, represent nodes’ similarity for intracommunity, and design routing strategies for intercommunity communication, while Gently [8] adopted a context-aware adaptive routing (CAR) algorithm and LABEL protocol. When no nodes within the destination community are in reach, Gently adopts the CAR-like routing algorithm. When the message carrier encounters a node coming from the destination community, it utilizes a LABEL-based protocol. Finally, a CAR-like routing strategy is used again to transmit the messages to the destination node in the destination community. These schemes predict future encounter probability by using historical records. However, nodes’ group identities are ignored. Xia et al. [15] recently proposed an interest-based routing algorithm, called BEEINFO. This scheme is constructed based on a fact that people usually gather together to obtain and share their interest information. Therefore, they utilize interests to form communities, and design different forwarding strategies. However, they fail to protect the sensitive interest information.

The motivation for the proposed PRIF approach is to protect nodes’ interests and make forwarding decisions. Specifically, to detect and maintain communities, similar to BEEINFO, PRIF uses interests to construct communities. To protect nodes’ interests, a privacy-preserving authentication protocol is designed. Considering people’s regular behaviors, PRIF gathers nodes’ community information, and predicts the future destination community or destination node encounter probability.

III. PRIF

This section elaborates the design details of PRIF. We first give an overview of the whole scheme, and introduce the community detection method followed by the key concept of PRIF: a new social metric as community energy. Then, we introduce the PRIF scheme.

A. Overview

The system model considered in this paper is a typical SIOV application scenario. There are potentially three kinds of mobile objects in the street: 1) cars; 2) buses; and 3) pedestrians. Each car or bus owns a vehicle device, and each pedestrian carries a mobile device. They communicate with each other through wireless interface, such as Wi-Fi or Bluetooth. In this paper, we use nodes to represent these devices. The aim of the application is to design an effective forwarding strategy by using these mobile nodes, without disclosing their interest information.

Since mobile nodes are used and controlled by people, the carriers' behavior can be an exact indicator of the nodes behavior. Thus, we take advantage of the human social property (i.e., interests) to make forwarding decisions. Normally, people have different interests. Although interests usually change over time, they can be considered relatively stable in a given time period. For example, some people are interested in reading in day-to-day activities, but during the World Cup, they may pay more attention to football. People with the same interest get together more often than others to obtain and share their interest information. For instance, people who share the interest of shopping appear frequently in the shopping malls but they nearly have no interaction with those who are interested in fishing. We assume a community is only related to one interest. If a person has more than one interest, he/she will belong to several communities simultaneously. In order to make our scheme easily understandable, similar to [15], each node is assumed to only hold one interest. During the forwarding process, node’s interest information should be protected. We summarize the assumptions of the system below.

1) There are three types of mobile nodes (cars, buses, and pedestrians) in the application scenario, which forms a typical SIOV.
2) There are no malicious nodes, and nodes are fully cooperative when forwarding messages.
3) Each node only has one interest, and nodes with the same interest form a community.
4) Each node must register itself with trust authority (TA).

The notations used in this paper are listed in Table I.

| Symbol     | Definition                                                                 |
|------------|-----------------------------------------------------------------------------|
| M          | Messages to deliver                                                         |
| N_s        | Source node                                                                 |
| N_d        | Destination node                                                            |
| N_i        | Intermediate node                                                           |
| I_s        | Interest of the source node                                                 |
| I_d        | Interest of the destination node                                            |
| I_n        | Interest of an intermediate node                                            |
| E_{I(a,b)} | Inter-community energy between a and b                                      |
| E_{I_C(a,i)} | Intra-community energy between a and the community i                        |
| \alpha     | Inter-community energy prediction factor                                     |
| \beta      | Intra-community energy prediction factor                                     |

B. Community Energy

In this section, we will introduce the concept of community energy which is inspired by molecular chemistry.

1) Intercommunity Energy: In reality, molecules are composed of atoms and there exist forces among atoms. Similarly, we assume a force is generated when two nodes encounter one another. The force, called intercommunity energy, represents their social tie and determines their contact strength. The stronger energy a node has, the more opportunities it has to successfully deliver messages. Note that the intercommunity energy is only generated among nodes of the same community. We use (1), shown below, to define the intercommunity energy between the nodes a and b.

\[
E_{I(a,b)}(N) = \frac{d_{i(a,b)}(N)}{I_{t(N-1,N)}}
\]
where $d_{(a,b)}(N)$ is the contact duration between $a$ and $b$ in the $N$th encounter, and $t_{(N-1,N)}$ represents the duration that has elapsed from $(N-1)$th encounter end to $N$th encounter end.

The intercommunity energy has a transitive property, which is based on an observation in reality. For example, if a person A frequently meets B, and meanwhile B frequently meets C, then A is also considered as a good forwarder to deliver messages to C though they may not encounter one another. Similarly, as shown in Fig. 2(a), a node $a$ establishes an energy to a node $b$, and $b$ builds an energy to a node $c$. Then, an indirect energy between $a$ and $c$ is generated as in (2), which is similar to [19]

$$E_{I(a,c)} = E_{I(a,c)}^{(old)} + (1 - E_{I(a,c)}^{(old)}) \times E_{I(a,b)} \times E_{I(b,c)}. \tag{2}$$

The nodes with high energy in the past usually are good forwarders in the future. Therefore, we define the energy prediction as in (3) using an exponential weighted moving average [28]

$$E_{I(a,b)}(N + 1) = \alpha \times E_{I(a,b)}(N - 1) + (1 - \alpha) \times E_{I(a,b)}(N) \tag{3}$$

where $\alpha$ is the intercommunity energy prediction factor.

2) Intracommunity Energy: In social networks, if a person encounters others from the same community frequently, the person can be considered as a good choice to forward messages destined for this community. We utilize degree centrality, which is the number of community nodes that a node encounters, to measure the community strength of the node, as shown in Fig. 2(b). However, considering the fast movement of mobile nodes, it may not be reasonable to directly use degree centrality. For example, a car may encounter many nodes interested in shopping around a shopping mall but fewer such nodes will be encountered after passing by the mall. Thus, we use average degree centrality to represent the intracommunity energy between the node $a$ and the community $i$, as in

$$E_{C(a,i)}(N) = \sum_{k=1}^{N} n_{i_k} / t_N. \tag{4}$$

In (4), $\sum_{k=1}^{N} n_{i_k}$ is the total number of nodes belonging to the same community $i$ that a node encounters from the first encounter to the $N$th encounter, and $t_N$ is the duration time. If $a$ does not encounter members from the community $i$ for a long time, its intracommunity energy $E_{C(a,i)}$ will decrease sharply. In addition, we use (5) to combine the past and present observations to predict the future intracommunity energy. $\beta$ is the intracommunity energy prediction factor, which is similar to $\alpha$ in (3)

$$E_{C(a,i)}(N + 1) = \beta \times E_{C(a,i)}(N - 1) + (1 - \beta) \times E_{C(a,i)}(N). \tag{5}$$

3) Energy Decay: Finally, we consider the fact that if nodes do not encounter each other in a period of time, they may not remain good forwarders for each other. Thus, an evaporation/aging process is necessary. We use (6) and (7) to decay the community energy

$$E_{I_{new}} = E_{I_{old}} \times \gamma^k \tag{6}$$

$$E_{C_{new}} = E_{C_{old}} \times \gamma^k \tag{7}$$

where $\gamma$ is the aging factor, and $k$ is the number of time intervals since the last time energy was aged.

When nodes move around, they share and gather interest information, and further update the above community energy information.

C. Privacy-Preserving Interest-Based Forwarding

In this section, we introduce the PRIF scheme, including system initialization, privacy-preserving authentication, forwarding process, message scheduling, and buffer management strategies.

1) System Initialization: Let $p$ be a large prime, $\alpha \in Z_p^*$, and the order of $\alpha$ be $q$, where $q$ is a large prime factor of $p - 1$. $H_1 : \{0,1\}^* \rightarrow Z_q^*$ and $H_2 : \{0,1\}^* \rightarrow \{0,1\}^k$ are cryptographic hash functions.

TA generates a certificate revocation list $RL$, which is originally empty and public. In order to create the group $G_l$ (i.e., the community $G_l$), $l \in [1, L]$. TA randomly chooses $a_l \in Z_q^*$ and computes $\gamma_l = \alpha^{a_l} \mod p$. Then, TA sets the group secret key $msk$ for $G_l$ as $a_l$. In addition, TA generates a group ID $GID_l$ for $G_l$. 

![Fig. 2. Community energy. (a) Intercommunity energy. (b) Intracommunity energy.](image)
When a node $U_i$ wants to join the group $G_i$, TA registers it. TA generates a certificate and sends it to $U_i$ over an authenticated private channel. TA randomly selects a string $id_i$ and $k_i \in Z_q^*$, and generates a Schnorr signature $s_i = (e_i, s_i)$, where $e_i = H_i(id_i, a^k_i \mod p)$ and $s_i = a e_i + k_i \mod q$. $U_i$'s certificate is $cert_i = (id_i, e_i, s_i, y_i)$. Note that, $y_i$ is known by the members of all the groups and TA, while the certificate $cert_i$ is only known by $U_i$ itself. If $U_i$ wants to leave the group, TA inserts $id_i$ into $\mathcal{R}_L$.

2) Privacy-Preserving Authentication: Assume that $U_i$ claims that it is affiliated to the group $G_i$, and $U_j$ claims that it belongs in the group $G_j$. After executing privacy-preserving authentication, $U_i$ can identify if $U_j$ belongs in $G_j$ and $U_j$ can identify if $U_i$ is affiliated to $G_i$. If they are in the same group, we can conclude that they have the same interest.

We assume that a node $U_i$ with $(id_i, e_i, s_i, y_i)$ belongs to $G_i$, and $G_j$'s group ID is $GID_j$. Assume $U_i$ encounters another node $U_j$, where $U_j$ claims that it is affiliated to $G_j$. $U_i$ can communicate with $U_j$ to check if $U_j$ is affiliated to $G_j$. Specifically, $U_i$ performs the following steps.

1) $U_i$ randomly selects $b_i \in Z_q^*$. Here, $b_i \mod q \neq 0$.
2) $U_i$ calculates $(y_i = a^{k_i} \cdot y_i^{-e_i} \mod p)$ and $B_i = a^{b_i} \mod p$.
3) $U_i$ sends $M_i = (GID_i, id_j, Y_j, B_i)$ to $U_j$.

Similarly, $U_j$ generates $M_j = (GID_j, id_i, Y_i, B_j)$, and sends it to $U_i$.

If $id_j$ is not listed in $\mathcal{R}_L$ and $(Y_j)^{(p-1)/q} \notin \{0, 1\}$, $U_i$ computes $K_{i,j} = B_j^{b_i} \mod p$ and sets $v_j = (h_{i,j}, sid_j)$, where $h_{i,j} = H_2(K_{i,j}, sid_j)$ and $sid_j = [M_i||M_j]$. Otherwise, $U_i$ randomly selects $h'_{i,j} \sim [0, 1]^k$, then sets $v_j = (h'_{i,j}, sid_j)$ and rejects $= T$. $U_i$ sends $v_i$ to $U_j$. Similarly, $U_j$ sends $v_j = (h_{j,i}, sid_j)$ to $U_i$, where

$$h_{i,j} = H_2(B_j^{b_i} \mod p, sid_j)$$

$$= H_2(a^{b_i} \cdot e_i \cdot k_j \mod p, sid_j).$$

If reject $= T$, then $U_i$ rejects communication. Otherwise, $U_i$ performs the following steps.

1) After $U_i$ receives $v_j$, $U_i$ computes $h'_{j,i} = H_2((y_j \cdot H_j(id_j,Y_j))^{b_i} \mod p, sid_j)$.
2) $U_i$ checks whether $h'_{j,i}$ equals to $h_{j,i}$. Since

$$h'_{j,i} = H_2((y_j \cdot H_j(id_j,Y_j))^{b_i} \mod p, sid_j)$$

$$= H_2(a^{y_j \cdot (e_j + k_j)} \mod p, sid_j)$$

Thus, if $h'_{j,i} \neq h_{j,i}$, $U_i$ can conclude that $U_j$ is an invalid participant. Otherwise, $U_i$ can conclude that the group ID of $U_j$ is $GID_j$.

According to $GID_j$, $U_i$ can conclude whether $U_j$ belongs in the same group with $U_i$.

3) Forwarding Process: When mobile nodes are in communication range, they will communicate with each other. The forwarding process consists of two parts: 1) community energy awareness and 2) message forwarding strategy.

---

**Algorithm 1:** Pseudocode for Community Energy Awareness

```plaintext
for all intermediate nodes $N_i$ connected to $N_s$
  if $i == s_i$ then
    // Update the direct inter-community energy;
    if $N_i$ has the inter-community energy record of $N_s$ then
      Update the inter-community energy with Eq. (1);
    else
      Initialize the inter-community energy between $N_i$ and $N_s$;
      // Update the indirect inter-community energy;
      for all connected nodes of $N_i$ and $N_s$ do
        Update the indirect inter-community energy with Eq. (2);
      end
    end
  else
    // Update intra-community energy;
    if $N_s$ has intra-community energy record of $N_i$ then
      Update intra-community energy with Eq. (4);
    else
      Initialize intra-community energy;
      Predict inter-community energy with Eq. (3);
      Predict intra-community energy with Eq. (5);
  end
end
```

**Community Energy Awareness:** When two nodes (for example $N_i$ and $N_j$) encounter, they first check if they are in the same community in a privacy-preserving way, and then update the community energy. If they are affiliated to different communities, they accumulate the community number and update their intracommunity energy. Otherwise, they compute the connection time and update their intercommunity energy. We give the pseudocode of community energy awareness in Algorithm 1.

**Message Forwarding Strategy:** The message forwarding strategy is the core of PRIF. By using community energy, the best forwarders can be chosen for the destination. According to the communities of $N_s$, $N_i$, and $N_d$, PRIF uses different message forwarding strategies.

Assume that a node $N_i$ with a message $M$ destined for $N_d$ meets another node $N_i$. If $N_i$ is not the destination node and $N_i, N_s, N_d$ belong in the same community, intercommunity energy will be used to make forwarding decisions. If $N_i$ has higher intercommunity energy to the destination, it will be selected as a better forwarder. Otherwise, $N_i$ will stop forwarding and wait for a better opportunity. If $N_i$ does not share the same interest with $N_d$, then there are only two cases where the forwarding process can occur: 1) $I_i == I_d$. In this case, $N_i$ belongs to the destination community and 2) $N_i$ does not belong to $N_d$’s community and $E_{C_{(N_i,I_d)}} < E_{C_{(N_d,I_d)}}$. Otherwise, $N_i$ will continue holding the message $M$.

As a whole, PRIF looks for active intermediate nodes (with higher intercommunity or intracommunity energy) which will allow fast transfer of message to the destination node.
Algorithm 2: Pseudocode for Message Forwarding Strategy

When $N_i$ with a message $M$ destined for $N_d$ encounters a node $N_i$.

if $N_i$ is $N_d$ then
  Deliver $E_{ID_d}(M)$ from $N_i$ to $N_d$
else
  if $I_i == I_d$ then
    // $N_i$ belongs to the destination community;
    if $I_i == I_d$ then
      // $N_i$ belongs to the destination community;
      if $E_{I_i(N_i,N_d)} < E_{I_i(N_i,N_d)}$ then
        // $N_i$ has higher inter-community energy;
        Deliver $E_{ID_d}(M)$ from $N_i$ to $N_i$;
    else
      // $N_i$ does not belong to the destination community;
      if $I_i == I_d$ then
        // $N_i$ belongs to the destination community;
        Deliver $E_{ID_d}(M)$ from $N_i$ to $N_d$;
      else
        // $N_i$ does not belong to the destination community;
        if $E_{C(N_i,I_d)} < E_{C(N_i,I_d)}$ then
          // $N_i$ has higher intra-community energy;
          Deliver $E_{ID_d}(M)$ from $N_i$ to $N_i$;
  end
  // $N_i$ does not belong to the destination community;
end

or destination community. When the message has reached the destination, it broadcasts a response message to inform all nodes which still maintain the message to discard it. We give the pseudocode of message forwarding strategy in Algorithm 2. In Algorithm 2, $E$ denotes a secure identity-based encryption algorithm [29], $ID_d$ presents the pseudo identity of $N_d$, and $E_{ID_d}(M)$ is the ciphertext of the message $M$.

4) Message Scheduling and Buffer Management: Since all mobile nodes have limited resources (i.e., battery power and buffer size), it is necessary to design message scheduling and buffer management strategies to improve the forwarding efficiency. The message scheduling policy decides in what order to deliver messages so as to ensure messages can be delivered to the destination node with higher delivery opportunities. In PRIF, we design strategies based on their communities. The message whose interest is the same as that of the current node will have a priority. When the buffer size reaches its capacity, the buffer management strategy decides which messages will be discarded if new messages arrive. Moreover, similar to message scheduling, the buffer management scheme is also based on communities. Details of both are given below.

Message Scheduling Algorithm: When $N_i$ is selected as a message forwarder and it has a set of messages to be delivered, then the relation between $N_i$ and $N_d$ is a major factor that needs to be considered. Specifically, the algorithm orders messages with the following priority rules.

1) The messages satisfying $I_i == I_d$ will have priority. The messages satisfying this condition will be ordered according to intercommunity energy. If the intercommunity energy is equal, the newer message will be transmitted first.
2) For the messages which do not satisfy $I_i == I_d$, it suits intracommunity transmission, hence intracommunity energy of $N_i$ is considered. The messages with higher intracommunity energy will have higher priority. If intracommunity energy is equal, then the newer one will be transmitted first.

Buffer Management Algorithm: The buffer management algorithm relies on the relation between the source node $N_i$ and the message. It discards the messages following the reverse order as that of the message scheduling sequence.

1) The messages which have different interests with the destination nodes will be discarded first. In this condition, the messages with lower intracommunity energy will be replaced first. In the case that intracommunity energy is equal, the older one will be discarded.
2) We then consider the messages in which $I_i == I_d$. The messages with lower intercommunity energy will be replaced first. If the intercommunity energy is equal, the message coming later will be discarded.

IV. SECURITY ANALYSIS

In this section, we will introduce the security model and prove that our scheme is privacy-preserving by showing that the adversary cannot determine the participants’ interest information.

A. Security Model

In this security model, the privacy property is defined by using a game between an adversary $A$ and a challenger $C$. The adversary $A$’s goal is to learn about the players’ interest information. The adversary cannot learn the interest information unless it can distinguish the two executions: one where the challenger $C$ executes the protocols as honest players, while the other where the adversary $A$ runs the protocol with a simulator.

First, the challenger $C$ creates a group in which $m$ members are included. Specifically, $C$ generates $msk$ of the group. Besides, $C$ generates a certificate $cert_i$ for the user $U_i$, where $i \in [1, m]$. Then, $C$ chooses corrupted players and gives the certificates of the corrupted players to $A$. Subsequently, $C$ updates $RL$.

Afterwards, $A$ sends a polynomial number of $Start(\Pi_i^1, G)$, $Send(\Pi_i^1, \Delta)$, and $Corrupt(U_i)$ queries adaptively. $C$ picks at random a bit $b$ uniformly, where $b \in [0, 1]$. If $b = 1$, $C$ answers $A$’s requests as honest players. Otherwise, $C$ responds to $A$ by using the simulator. When $b = 0$, $C$ replies to the queries as follows.

1) $Start(\Pi_i^1)$ and $Send(\Pi_i^1, \Delta)$ queries: $C$ answers the queries with the messages generated by the simulator. $C$ will set reject as $T$ and return null, if $\Delta$ is incorrect.
2) $Corrupt(U_i)$: $C$ gives $cert_i$ to $A$ and updates the revocation list.
Finally, $A$ outputs a bit $b'$. $A$ wins the game, if $b' = b$ holds. The advantage with which $A$ wins the game is defined to be

$$\mathsf{Adv}(A) = |2 \cdot \Pr[b = b'] - 1|.$$

Definition 1: The proposed protocol is said to be privacy-preserving, if for any probabilistic polynomial time adversary $A$, $\mathsf{Adv}(A)$ is negligible.

### B. Security Proof

**Theorem 1:** Assume $A$ can ask the $H_2$ random oracle at most $qH_2$ times. The proposed scheme is secure with the probability $1 - 2qH_2/(q-1)$, where $q$ is the order of group $Z_q^*$ and $q$ is a large prime.

**Proof:** In order to prove that our protocol is privacy-preserving, we design two games Game0 and Game1, where Game1 denotes a simulation, and Game0 represents the real game.

In order to show that the adversary $A$ cannot distinguish its view in Game1 and its view in Game0, Game1 is constructed as follows.

**Simulation:** Assume pid$_i^0 = \{U_i, U_j\}$.

1. **Start($\Pi_i^0$):** $C$ randomly chooses id, corresponding to $U_i$, selects randomly $k \in Z_q^*$ and calculates $Y_i = g^k \mod p$, then picks randomly $b_i \in Z_q^*$ and calculates $B_i = a^{b_i} \mod p$. $C$ replies with $M_i = (\text{id}_i, Y_i, B_i)$.

2. **Send($\Pi_i^1, A$):** $C$ randomly selects $h_j^i \leftarrow \{0, 1\}^q$, then sets $v_i = (h_j^i, \text{sid}_i^0)$. Output $v_i$.

3. **Corrupt($\text{id}_i$):** $C$ gives cert, to $A$ and inserts $\text{id}_i$ to the revocation list.

For any $\Pi_i^0$ and any $j \in D$, $K_{i,j}$ and $K_{j,i}$ is defined via the messages ($\text{id}_j, Y_j, B_j$) and ($\text{id}_j, Y_j, B_j$), where ($\text{id}_j, Y_j, B_j$) is sent by $A$ to $\Pi_i^1$. That is, $K_{i,j} = (y^{(\text{id}_j, Y_j)} Y_j)^{b_i} \mod p$ and $K_{j,i} = (y^{(\text{id}_j, Y_j)} Y_j)^{b_j} \mod p$. Let $\mathcal{E}_{H_2}(i)$ denote the event that $A$ sends $H_2$ query on ($\text{sid}_j^0$) or ($\text{sid}_i^0$).

We can observe the difference between the two games. That is, $h_j^i$ is randomly selected in Game1, while $k \in D$. Thus, $A$ cannot distinguish its view in Game1 from its view in Game0 unless $\mathcal{E}_{H_2}(i)$ happens. In the proposed scheme, $K_{i,j} = (y^{(\text{id}_j, Y_j)} Y_j)^{b_i} \mod p$, $K_{j,i} = (y^{(\text{id}_j, Y_j)} Y_j)^{b_j} \mod p$, and $A$ cannot know $y$. Therefore, $K_{i,j}$ and $K_{j,i}$ can be considered as $a^{x_1} \mod p$ and $a^{x_2} \mod p$, respectively, for some unknown $x_1 \in Z_q^*$ and some unknown $x_2 \in Z_q^*$ for $A$. Then, the probability with which $A$ can query $K_{i,j}$ or $K_{j,i}$ is $2qH_2/(q-1)$ at most. That is, the probability with which event $\mathcal{E}_{H_2}(i)$ happens is $2qH_2/(q-1)$ at most. Moreover, $A$ can distinguish its view in Game1 from that in Game0 with $2qH_2/(q-1)$ at most. $2qH_2/(q-1)$ becomes negligible, since $q$ is a large prime. That is, $A$ cannot distinguish its view in Game1 from that in Game0. Therefore, our protocol captures the privacy-preserving property.

### V. PERFORMANCE EVALUATION

We have conducted extensive experiments to evaluate the performance of the proposed PRIF and compared it with the following routing and forwarding methods, i.e., BEEINFO-D&S [15], Epidemic [18], and PRoPHET [19].

The following metrics have been used for performance comparison.

1. **Delivery Ratio:** The average ratio of successfully delivered messages to all created messages from the sources to the destinations.
2. **Overhead:** The percentage of relayed messages which excludes the delivered messages.
3. **Average Hop Count:** The average number of hops when messages are delivered successfully.

### A. Simulation Settings

Similar to BEEINFO-D&S [15] and other DTN routing schemes, such as [28] and [30]–[32]. The OppNet Environment Simulator [33] is used to evaluate the performance of the PRIF.

In our experiments, five groups of nodes are considered, including two pedestrian groups, two car groups, and one bus group. All groups consist of 40 nodes except the bus group which has six nodes. There are two kinds of Bluetooth interface to realize wireless transmission: one is used for cars and pedestrians, where the communication range is 10 m and transmission rate is 2 Mb/s, and the other one is used for buses with a higher communication range and transmission rate (i.e., 100 m and 10 Mb/s). Messages are only generated by nodes of cars and pedestrians groups, every 50–90 s. The size of message is set as 0.5–1 MB. We implement the experiments by varying the values of two important factors: 1) buffer size (10–50 MB) and 2) TTL (600–3600 min). Detailed simulation parameters are listed in Table II.

| Parameter | Value or Range |
|-----------|----------------|
| Simulation time | 400000 s |
| Time window | 30 s |
| Warm up time | 5000 s |
| Area | $4500 \times 3400$ m² |
| Speed of pedestrians | $0.5 \sim 1.5$ m/s |
| Speed of cars | $2.7 \sim 13.9$ m/s |
| Speed of buses | $7 \sim 10$ m/s |
| Wait time at destination | $100 \sim 200$ s |
| Message TTL | 600 min |
| Event interval | 50 – 90 s |
| Message size | 500 – 1024 KB |
| Number of nodes in each car/pedestrian group | 40 |
| Number of nodes in each bus group | 6 |
| $\alpha, \beta$ | 0.3 |
| $\gamma$ | 0.98 |

### B. Simulation Results and Analysis

The performance of the proposed PRIF is evaluated over different buffer sizes, message’s TTL, and simulation time. Each experiment runs 30 times and we compute the average result. In Figs. 3–5, we show the results of simulation experiments for delivery ratio, overhead, and hop count, respectively.

In Fig. 3, we compare the proposed PRIF scheme with other three schemes when buffer size ranges from 10 to 50 MB.
It can be observed that, with larger buffer size, more messages will be delivered to the destinations, less overhead will be generated, and fewer hops are required. The proposed PRIF performs best in terms of the delivery ratio and overhead. For example, when the buffer size is set as 50 MB, PRIF delivers 65.29% messages (compared with 62.57% for BEEINFO-D&S) with message overhead of 1146.5474 (compared with 1365.5567 for BEEINFO-D&S), and hop count of 2.8679 (similar to 2.7730 for BEEINFO-D&S). By comparison, Epidemic and PRoPHET perform worse with 49.28% and 43.84% in delivery ratio, 1365.5567 and 2276.1788 in overhead, and 4.6457 and 3.2951 in hop count experiments, respectively.

In Fig. 4, we show the performance of the four schemes with varying TTL, where the simulation time is 400 000 s and the buffer size is 10 MB. It can be seen that as the TTL increases, message delivery ratio of all schemes decreases, and PRIF exhibits best performance. When the TTL is set as 3600, PRIF delivers 42.75% messages, which is 16.84% higher than BEEINFO-D&S, 75.42% higher than Epidemic and 66.08% higher than PRoPHET. For the overhead, PRIF also outperforms the rest. In terms of the hop count, the four schemes are comparable in performance, and PRIF is at similar level with that of BEEINFO-D&S.

Fig. 5 shows the performance of all these schemes varying with simulation time (using 10 MB of buffer size and 600 min of message TTL). When the simulation time increases from 100 000 s to 500 000 s, PRIF can gather more community energy information, which helps nodes to select better forwarders. The trend of PRIF is similar to those of other schemes, but proves the overall benefit as shown in Figs. 3 and 4. Hence, it can be concluded that, with long lifetime of networks, PRIF can give better performance while preserving the privacy of interests.

In summary, PRIF achieves better delivery ratio and overhead compared with the other three schemes, and gives comparable results with BEEINFO-D&S for the hop count metric. These observations confirm the efficiency of introducing community energy in the design of social-based forwarding for SIOV.

VI. CONCLUSION

In this paper, we propose a PRIF scheme for SIOV, which not only protects nodes’ interests but also improves
the forwarding performance. We have designed a privacy-preserving authentication protocol to recognize communities among mobile nodes. Moreover, we classify communities based on nodes’ interests and present detailed methods to calculate community energy, including intercommunity energy and intracommunity energy based on their interests. Extensive simulations have been conducted, which demonstrate the efficiency and effectiveness of the proposed scheme.

REFERENCES

[1] W. Yuan, P. Wang, W. Liu, and W. Cheng, “Variable-width channel allocation for access points: A game-theoretic perspective,” IEEE Trans. Mobile Comput., vol. 12, no. 7, pp. 1428–1442, Jul. 2013.
[2] L. Pelusi, A. Passarella, and M. Conti, “Opportunistic networking: Data forwarding in disconnected mobile ad hoc networks,” IEEE Commun. Mag., vol. 44, no. 11, pp. 134–141, Nov. 2006.
[3] X. Lin and R. Lu, Vehicular Ad Hoc Network Security and Privacy. Hoboken, NJ, USA: Wiley, 2015.
[4] F. Warthman, Delay-Tolerant Networks: A Tutorial. Palo Alto, CA, USA: Morgan & Claypool, 2009.
[5] A. M. Vegni and V. Loscrì, “A survey on vehicular social networks,” IEEE Commun. Surveys Tuts., vol. 17, no. 4, pp. 2397–2419, 4th Quart., 2015.
[6] X. Hu et al., “S-Aframe: Agent-based multilayer framework with context-aware semantic service for vehicular social networks,” IEEE Trans. Emerg. Topics Comput., vol. 3, no. 1, pp. 44–63, Mar. 2013.
[7] X. Yu, X. You, B. Fang, and Y. Y. Tang, “Thinning character using modulus minima of wavelet transform,” Int. J. Pattern Recognit. Artif. Intell., vol. 20, no. 3, pp. 361–376, 2006.
[8] Z. He, X. Li, X. You, D. Tao, and Y. Y. Tang, “Connected component model for multi-object tracking,” IEEE Trans. Image Process., vol. 25, no. 8, pp. 3698–3711, Aug. 2016.
[9] Z. He, S. Yi, Y.-M. Cheung, X. You, and Y. Y. Tang, “Robust object tracking via key patch sparse representation,” IEEE Trans. Cybern., vol. 47, no. 2, pp. 354–364, Feb. 2017.
[10] P. Hui, J. Crowcroft, and E. Yoneki, “BUBBLE rap: Social-based context-aware semantic service for vehicular social networks,” in Proc. IEEE Int. Conf. Commun. (ICC), Paris, France, May 2017, pp. 1–6.
[11] D. Boneh and M. K. Franklin, “Identity-based encryption from the weil pairing,” in Proc. Adv. Cryptol. CRYPTO 21st Annu. Int. Cryptol. Conf., Heidelberg, Germany, Aug. 2001, pp. 213–229.
[12] Z. He, S. Yi, Y.-M. Cheung, X. You, and Y. Y. Tang, “Robust object tracking via key patch sparse representation,” IEEE Trans. Cybern., vol. 47, no. 2, pp. 354–364, Feb. 2017.
[13] P. Dong, X. Du, J. Sun, and H. Zhang, “Energy-efficient cluster management in heterogeneous vehicular networks,” in Proc. IEEE Conf. Comput. Commun. Workshops (INFOCOM Workshops), San Francisco, CA, USA, Apr. 2016, pp. 644–649.
[14] X. Du and H. Chen, “Security in wireless sensor networks,” IEEE Wireless Commun., vol. 15, no. 4, pp. 60–66, Aug. 2008.
[15] A. C. K. Vendramini, A. Munaretto, M. R. Delgado, and A. C. Viana, “Grant: Inferring best forwarders from complex networks’ dynamics through a greedy ant colony optimization,” Comput. Netw., vol. 56, no. 3, pp. 997–1015, 2012.
[16] D. Boneh and M. K. Franklin, “Identity-based encryption from the weil pairing,” in Proc. Adv. Cryptol. CRYPTO 21st Annu. Int. Cryptol. Conf., Heidelberg, Germany, Aug. 2001, pp. 213–229.
[17] J. Mia, O. Hasan, S. B. Mokhtar, L. Brunie, and G. Gianini, “A delay and cost balancing protocol for message routing in mobile delay tolerant networks,” Ad Hoc Netw., vol. 25, pp. 383–392, Feb. 2015.
[18] A. Keränen, T. Kärkkäinen, and J. Ott, “Simulating mobility and DTNs with the ONE (invited paper),” J. Commun., vol. 5, no. 2, pp. 92–105, 2010.

Lieuhang Zhu received the Ph.D. degree from the Beijing Institute of Technology, Beijing, China, in 2009. He is currently a Professor with the School of Computer Science and Technology, Beijing Institute of Technology. His current research interests include secure data protocols and design, group key exchange protocols, wireless sensor networks, and cloud computing.

Chuan Zhang received the bachelor’s degree in network engineering from the Dalian University of Technology, Dalian, China, in 2015. He is currently pursuing the Ph.D. degree at the School of Computer Science and Technology, Beijing Institute of Technology, Beijing, China. His current research interests include secure data services in cloud computing, security and privacy in IoT, and big data security.

Chang Xu received the Ph.D. degree in computer science from Beihang University, Beijing, China, in 2013. She is currently an Assistant Professor with the School of Computer Science and Technology, Beijing Institute of Technology, Beijing. Her current research interests include security and privacy in VANET and big data security.
Xiaojiang Du (M’04–SM’09) is currently a Professor with the Department of Computer and Information Sciences, Temple University, Philadelphia, PA, USA. He has authored or co-authored over 200 journals and conference papers and has been awarded over $5M research grants from the U.S. National Science Foundation and Army Research Office. His current research interests include security, systems, wireless networks, and computer networks.

Rixin Xu received the bachelor’s degree in software engineering from the Harbin Institute of Technology, Harbin, China, and the master’s degree in software engineering from Peking University, Beijing, China. He is currently pursuing the Ph.D. degree at the School of Computer Science and Technology, Beijing Institute of Technology, Beijing. His current research interests include security of Internet of Things and side-channel attacks.

Kashif Sharif (M’08) received the M.S. degree in information technology and Ph.D. degree in computing and informatics from the University of North Carolina at Charlotte, Charlotte, NC, USA, in 2004 and 2012, respectively. He is currently an Associate Professor with the Beijing Institute of Technology, Beijing, China. His current research interests include wireless and sensor networks, network simulation systems, software defined and data center networking, ICN, and the Internet of Things.

Dr. Sharif is a member of the ACM.

Mohsen Guizani (S’87–M’90–SM’98–F’09) received the B.S., M.S., and Ph.D. degrees from Syracuse University, Syracuse, NY, USA. He is currently a Professor and the Electrical and Computer Engineering Department Chair with the University of Idaho, Moscow, ID, USA. He has authored 9 books and over 400 publications. His current research interests include wireless communications, mobile cloud computing, computer networks, security, and smart grid.

Dr. Guizani was the Chair of the IEEE Communications Society Wireless Technical Committee. He served as an IEEE Computer Society Distinguished Speaker.