Biometrics-based Data Link Layer Anonymous Authentication in VANETs

Lin Yao*, Chi Lin*, Jing Deng†, Fangyu Deng*, Jingwei Miao ‡, and Kangbin Yim §
*School of Software Technology, Dalian University of Technology, Dalian 116023, China
Email: yaolin@dlut.edu.cn
†Department of Computer Science, University of North Carolina at Greensboro NC 27412, U.S.A.
‡University of Lyon, CNRS, INSA-Lyon, LIRIS, UMR5205, F-69621, France
§Department of Information Security Engineering, Soonchunhyang University, Asan, Republic of Korea
Email: yim@sch.ac.kr

***Abstract*—The vehicular ad-hoc network (VANET) aims to make road traffic safer and more efficient. Security and privacy are two important but somewhat contradictory objectives. On the one hand, vehicle-to-vehicle (V2V) authentication requires drivers to present some individual credentials. On the other hand, drivers need privacy protections on their identities, physical locations, and other contextual information. No sensitive information should be unnecessarily exposed during the authentication phase. Therefore, anonymous authentication is expected to play a critical role in VANET. In this paper, we propose a novel biometrics-based anonymous mutual authentication with provable link-layer location privacy preservation. During its authentication phase, two vehicles negotiate their temporary session key and generate two temporary MAC addresses. These two addresses, instead of the real ones, are used in all future communication frames. We further protect the biometric privacy with the help of a biometric encryption technique. Our analysis and simulation results show that the protocol is superior and lightweight with strong security and privacy protection.**

I. Introduction

In Vehicular Ad-hoc Networks (VANETs), moving cars are treated as nodes to create mobile networks. As a special kind of mobile ad-hoc network, it is a promising network scenario for facilitating road safety, traffic management, as well as others, in transportation system. In VANET, vehicles will establish a multi-hop and self-organized network without a predefined or centralized infrastructure. There are two kinds of communications in VANETs [1]–[3]: vehicle-to-vehicle communication and vehicle-to-Road Side unit (RSU) communication (see Fig. 1).

Similar to other wireless networks, a VANET is susceptible to passive eavesdropping and active attacks. Authentication is the fundamental procedure to protect the security of VANET. Mutual authentication helps to prevent information leakage, to prevent service abuse, and to mitigate malicious attacks. However, traditional authentication mechanisms designed for static networks or closed systems with central controls are unlikely to succeed in VANETs because of the highly dynamic network conditions.

Moreover, privacy protection for VANETs is especially important. Firstly, the communication contains a number of sensitive information, such as vehicle status, vehicle location, network topology information and so on. Such information should be protected from outsiders. Secondly, if the identity of a vehicle is leaked, an attacker will successfully track it by collecting the route traffic messages. Therefore, how to achieve anonymous authentication has become a fundamental requirement for securing VANETs.

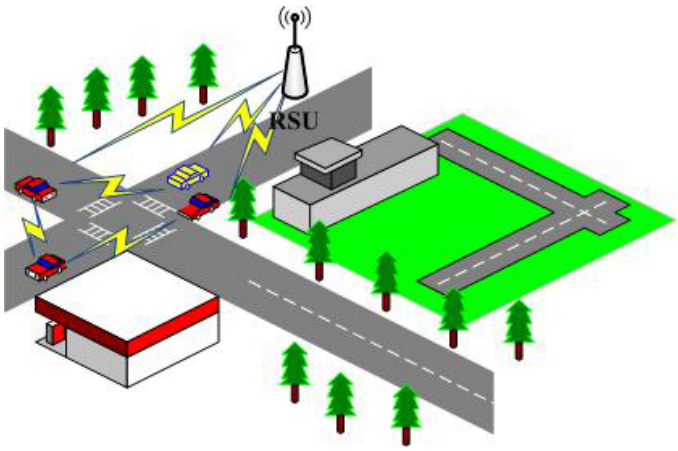

Fig. 1. The Architecture of VANET. Two types of communications are illustrated: vehicle-to-vehicle communication and vehicle-to-RSU communication. RSU is Road-Side Unit.

Anonymous authentication technology can provide effective solutions to protect VANET privacy [4]. In recent years, many anonymous authentication schemes have been proposed in VANET, including group-signature scheme [5]–[8], pseudonymous authentication [9]–[13], random silence [14], [15] and blind signature [16]. Group-signature means that vehicles hide in a group and use the group's public key to sign on behalf of the group. When a node receives a signature, it will use the group public key to verify the signature's legitimacy. Thus, the real identity of the signer cannot be known by the verifier. A pseudonym is a name that a person or group assumes for a particular purpose, which differs from his or her original or true name. In VANET, a pseudonym is the identifier of a vehicle entity and is used to hide the node's true identity. A vehicle can change its pseudonyms in different applications. Random Silence [14], [15] means a vehicle can choose the silent time and silent place randomly. During that

period, the node does not do any operation and only replaces its pseudonym. Blind Signature is used to ensure that the messages signed will not be modified [16].

However, in most of these protocols, the source and destination MAC addresses are sent in plaintext as required by communication protocols. All of the transmissions coming from the same mobile user bear the same MAC address, making it easy for adversaries to track the user. Therefore, user anonymity can be easily compromised.

In order to address this issue, we design in this work a novel anonymous authentication scheme based on biometric encryption at the data link layer. During the authentication process, a user's biometric is matched against the biometric template stored in the database through field sampling. This match will prove the user's identity. In order to protect MAC addresses from being eavesdropped, our scheme generates two temporary MAC addresses for the communicating parties. A unique session key is also generated for communication message encryption. Furthermore, biometric encryption is adopted in our scheme in order to protect biometric template from tampering [17], [18].

The rest of this paper is organized as follows. Section II highlights the related work. We describe the preliminaries in Section III. Section IV presents our algorithm. Security and privacy analyses are given in Section V. Finally, we conclude the paper in Section VI.

II. Related Work

Anonymous authentication is a very active research for securing the VANET. Because a VANET is a self-organized network without a fixed boundary, the traditional PKI-based approach is not suitable. Anonymous authentication scheme can be divided into four types [4], namely, group-signature scheme, pseudonymous authentication, random silence and blind signature.

Group-signature-based schemes [5]–[8] are utilized to protect the privacy of the nodes because no public entity will reveal the originator identity of a routine traffic message. Firstly, a short group signature scheme that supports Verifier-Local Revocation (VLR) was constructed in [5]. In this model, revocation messages were only sent to signature verifiers (as opposed to both signers and verifiers). Other researchers extended the group-signature-based anonymous authentication scheme to achieve that any public entity will not reveal the originator identity of a routine traffic message [6], [7]. However, the cost for signing and verifying were quite high. The VLR scheme was efficient when there were only a few revoked vehicles. Nevertheless, with the number of revoked vehicles, the vehicle revocation verification procedure became very time-consuming and inefficient [8]. To reduce the limitations, Giorgio proposed a group-signature-based scheme which enables vehicle on-board units to generate their own pseudonyms [8]. The group-signature-based scheme had an advantage that the revocation cost is linear with the number of revoked vehicles, but the checking operation involved two pairing calculations [19].

The pseudonymous authentication schemes [9]–[13] are designed based on traditional public-key digital signature to achieve anonymous authentication. The basic idea of the pseudonym was proposed by Raya and Hubaux [9]. Each vehicle must store a large set of pseudonymous certificates and will choose a random certificate to sign a message at one time. However, the certificate revocation list (CRL) increased quickly as a vehicle is revoked, because all the pseudonyms corresponding to the vehicle must be added into the CRL. To solve this problem, a novel authentication scheme was proposed [10]. In this scheme, a country was divided into a number of sub-regions. By signing region-specific certificates to a vehicular, the CRL size can be reduced. To improve the authentication efficiency, R.Lu introduced an efficient conditional privacy preservation (ECPP) protocol in VANETs [11]. This proposed protocol was characterized by the generation of on-the fly short-time anonymous keys between vehicle and RSU. The ECPP protocol could improve efficiency in terms of the minimized anonymous keys storage at each vehicle, fast verification on safety messages and an efficient conditional privacy tracking mechanism. To extend the RSU-aided distribute certificate service, an efficient distributed-certificate-service (DCS) scheme was proposed [12]. DCS protocol offered flexible interoperablility for certificate service in heterogeneous administrative authorities and an efficient way for any onboard units to update its certificate from the available RSUs. Verifying a mass of signatures within a rigorously required interval may cause the performance bottleneck. Therefore, a robust and efficient signature scheme for VANETs was proposed [13]. The scheme could also be gracefully transplanted to other similar batch signature schemes. However, all these pseudonyme schemes depended on the RSU density. The revocation cost was inversely proportional to the larger number of RSUs. But in many scenarios we could not decide the density of RSUs. Thus, their applications were restricted. Sun proposed an efficient pseudonymous authentication scheme with strong privacy preservation (PASS) for VANETs [19]. PASS supports the RSU-aided distributed certificate service that allowed the vehicles to update certificates on road, but the service overhead was almost unrelated to the number of updated certificates. Furthermore, it provided strong privacy preservation to the vehicles so that the adversaries could not trace the legitimate vehicles.

Random silence means that a vehicle can change the pseudonym in a specific time without being exposed by the adversary. At the end of the time interval, the vehicle will use a new pseudonym to replace the original one. The concept of a silent period was first proposed in [14] to protect a mobile user's location privacy. To avoid the linkability between two users, a mix-zone was formed. If the silent period was too long, the quality of service may have been affected. A scheme called silent cascade to enhance location privacy was proposed to trade off delay in silent cascade for anonymity [15].

Blind signature ensures the invisibility of the content of the signature to the signer. Moreover, the messages signed cannot be modified. Blind signature in VANETs was first proposed by

Zhang et al. [16] which was based on a blind signature scheme over ECC. The proposed protocol could not only protect the identity of each vehicular, but also dramatically reduce the movement tracking probability when the vehicles were handed over through a number of RSUs.

Besides those schemes mentioned above, there are other authentication methods. A social-based privacy-preserving packet forwarding (SPRING) protocol was proposed [19]. SPRING protocol was characterized by deploying RSUs at high social intersections to secure the vehicle-to-RSU communication. It could also achieve the privacy preservation and resist the black hole attacks. But the computational complexity of the intermediate nodes was high. In [20], some other casual information, such as knowledge of current and previous affiliations and some social contacts of peers, was introduced to establish an initial security context between nodes. However, it could not guarantee the anonymous communication between source and destination nodes. In [21], an anonymous communication solution was presented, which made a distinction between local and long distance communication. Though efficient mutual authentication was achieved, the scheme did not consider the internet structure of a VANET.

III. Preliminaries

In this section, we introduce several preliminaries that are foundations of our work.

A. Biometric Encryption Algorithm

Compared with token and knowledge, biometrics can represent a unique user. Furthermore, it can provide an unobtrusive and convenient authentication mechanism against fraud. Consequently, biometric authentication is used widely. But biometric templates are usually stored in plaintext, so it is easy to be compromised [18]. To secure the users' biometric templates, Biometric Encryption is proposed. Biometric Encryption solution has a two-stage process: the enrollment stage and the verification stage [17].

During the enrollment stage, the biometric image is bounded with a cryptographic key to create data as Bioscrypt. During the verification stage, the biometric image on the spot is combined with the Bioscrypt to recover the key. Bioscrypt does not reveal any information about the key or biometric feature, i.e. it is computationally hard to decode the key without any knowledge of the user's biometrics, and vice versa. Consequently, Bioscrypt provides an excellent privacy protection. The key itself is completely independent from biometrics and, therefore, can always be changed or updated. Even if the key is ever compromised, the biometrics cannot be leaked. Moreover, the key can be easily updated.

Using the Bioscrypt as the certificate reduces the storage request for the mobile users and the management request for certificates.

B. AMP+

Because of the low entropy of the password and the vulnerable quality of the password file, password authentication is not secure enough. To solve it, new password authentication and key agreement protocol (AMP), based on amplified password idea, was presented [22]. AMP can be generalized in any other cyclic groups easily; it mainly provides the password-verifier based authentication and the Diffie-Hellman based key agreement security. In the AMP protocol, an attacker who has no knowledge of the password is unable to mount any active guessing attack. Moreover, it cannot compute the shared key established between two genuine participating principals. AMP+ protocol is based on AMP and can achieve better security than AMP.

The AMP+ protocol has been proved that it can provide perfect forward security against Denning-Sacco attack, replay attack, on-line guessing attack, off-line guessing attack, small subgroup confinement and also password-file compromise.

TABLE I
Related Notations

Symbol	Meaning
U	A user
ID_x	Identity of entity X
$\mathcal{B}_x$	The Bioscrypt of entity x
$\mathcal{F}_x$	The eigen vector of entity x's face
$\mathcal{K}_x$	The key which corresponds to $\mathcal{B}_x$
K_A, K_A^{-1}	The public key and private key of entity A
$\{m\}K$	m is encrypted by K Hash function
$h()$	Hash function
$\mathcal{R}_{x(n)}$	The n-th random number generated by entity X
G	The ellipse curve
$X \rightarrow Y : \{m\}$	Entity X sends a message m to Y
$\|\|$	Message concatenation

IV. Proposed Scheme

In this section, we will present our pseudonym authentication in detail. In Table I, we list the frequently-used symbols.

A. System Model

Fig. 2 shows our system architecture. We consider a typical VANET, which consists of an authentication server (AS). Some stationary RSUs are deployed at the roadsides and a large number of vehicles are moving on the road.

(1) AS is fully trusted by all parties and is in charge of the registration of RSUs and vehicles. AS can divide the whole region into several domains.
(2) A RSU is deployed at the boundary between two domains. Each RSU is responsible for managing its domain. It connects with the AS by wired or wireless links in the system. If two vehicles locate in the same domain, they can communication through the RSUs. If two vehicles locate in the different domains, the packet will be relayed by the RSUs by multi hops.
(3) At the vehicle layer, the mobile users can communicate with each other to share local traffic information.

Our system aims at providing the anonymous authentication between two vehicles. We model our system into two logical layers, the security layer and the network layer in Fig. 3. The

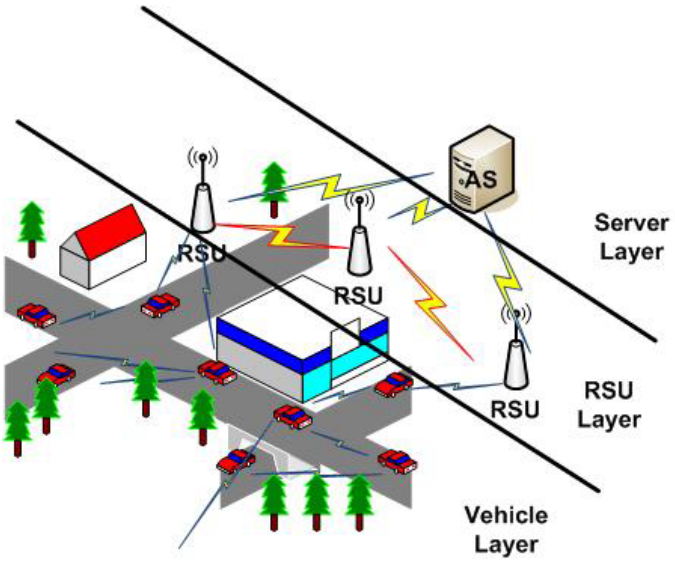

Fig. 2. System Model

network layer can provide routing services for the security layer and is transparent to the security layer. Therefore, in our scheme, we only focus on the security layer. The security layer can achieve anonymous authentication based on the pseudonym scheme. The working process of the system is summarized as follows:

(1) System initialization: Before a communication begins, every entity must go to register as a legal entity.
(2) Anonymous mutual authentication: Two vehicles can achieve anonymous authentication with the help of AS.
(3) Pseudonym generation: After mutual authentication, a session key will be established and a pseudonym for every user will also be negotiated.

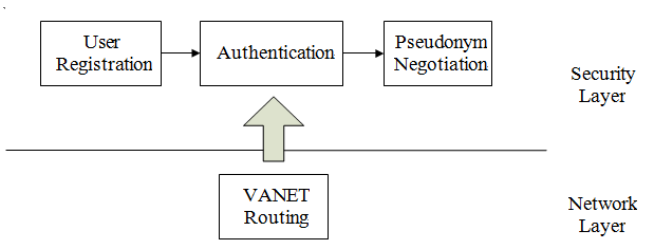

Fig. 3. Logical Layers

B. System Initialization

(1) AS generates a pair of public and private keys for U.
(2) Based on Biometric Encryption technique, AS helps U to generate his Bioscrypt with his face biometrics. AS stores U's Bioscrypt and discards U's biometrics. Bioscrypt, similar to a specific credential, represents U's identity.
(3) AS generates a unique number ID_U for each U and stores it.
(4) AS generates a random pseudonym for U.

C. Authentication Phase

In this subsection, we discuss how to achieve anonymous authentication with the help of AS. Since the opportunistic routing service in the network layer is transparent to the security layer, we only focus on the communication between the source user U_1 and the destination user U_2 while ignoring the transmissions of intermediate nodes and RSUs.

(1) $U_1 \rightarrow U_2 : h(ID_{U_1}) \parallel \{\mathcal{B}_1, \mathcal{F}_1, \mathcal{R}_{1(0)}\}\ K_{AS}$.
(2) $U_2 \rightarrow$ AS : $h(ID_{U_1}) \parallel \{\mathcal{B}_1, \mathcal{F}_1, \mathcal{R}_{1(0)}\}\ K_{AS} \parallel h(ID_{U_2}) \parallel \{\mathcal{B}_2, \mathcal{F}_2, \mathcal{R}_{2(0)}\}\ K_{AS}$.
(3) When AS receives the message from (2), U_1 will be proved legitimate if the key hash value produced by $\mathcal{B}_1$ and $\mathcal{F}_1$ is equal to $h(\mathcal{K}_1)$. Similarly, U_2 is proved legitimate.
(4) AS sends message to U_1 and U_2.
AS $\rightarrow U_1 : h(h(ID_{U_1})) \parallel \{\mathcal{R}_{1(1)}, \mathcal{R}_{2(1)}\}\ K_{U_1}$. AS $\rightarrow U_2 : h(h(ID_{U_2})) \parallel \{\mathcal{R}_{1(1)}, \mathcal{R}_{2(1)}\}\ K_{U_2}$.

D. Pseudonym Generation

After mutual authentication, a new session key K to secure the traffic between U_1 and U_2 will be established based on AMP+. Followed by this, U_1 and U_2 will negotiate their pseudonyms for subsequent communication.

(1) U_2 replaces $h(ID_{U_2})$ with ID_{U_2} and computes $S = \mathcal{R}_{1(1)} \times \mathcal{R}_{2(1)}$, $V = S \times G$ and $Q_{sp} = G \times r_{U_2}$. r_{U_2} is a random number. Then, U_2 sends the following message: $U_2 \rightarrow U_1 : h(S) \parallel \{Q_{U_2}, \mathcal{R}_{2(2)}\}V$.
(2) U_2 computes $S = \mathcal{R}_{1(1)} \times \mathcal{R}_{2(1)}$ and $V = S \times G$, and then get Q_{U_2} by decrypting message (2). $e_1 = h(Q_{U_2})$ and $Q_{U_1} = (Q_{U_2} \times e_1 + V) \times \mathcal{R}_{1(2)}$ is computed, and the following message is sent:
$U_1 \rightarrow U_2 : h(S) \parallel \{Q_{U_1}, \mathcal{R}_{1(2)}\}V$.
(3) U_2 computes $e_2 = h(Q_{U_2}, Q_{U_1})$, $\omega = (\mathcal{R}_{2(2)} \times e_1 + S)^{-1}(\mathcal{R}_{2(2)} + e_2)$, $K = h(Q_{U_1}, \omega)$, and $M_1 = h(Q_{U_2}, K)$. Then the message is sent:
$U_2 \rightarrow U_1 : h(S) \oplus S \parallel \{M_1, \mathcal{R}_{1(2)}\}V$.
(4) U_1 computes $e_2 = h(Q_{U_2}, Q_{U_1})$, $K' = h((Q_{U_2} + G \times e_2) \times \mathcal{R}_{1(2)})$, and $M_1' = h(Q_2, K)$. If M_1' = M_1 is true, U_1 will know that K' is equal to K and compute $M_2 = h(Q_{U_1}, K')$ and $Link = h(\mathcal{R}_1 \| \mathcal{R}_2 \| K)$. The previous 48 bits will be used as the subsequent pseudonym of U_1. Then the following message is sent:
$U_1 \rightarrow U_2 : h(h(S) \oplus S) \| M_2$.
(5) After receiving the message (4), U_2 will compute $M_2{}' = h(Q_{U_1}, K)$. If $M_2{}' = M_2$ is true, U_2 will know that K' is equal to K and compute $Link = h(\mathcal{R}_1 \| \mathcal{R}_2 \| K)$. Similarly, the last 48 bits will be used as the subsequent pseudonym of U_2.

U_1 and U_2 can judge that the other corresponding party knows the shared key K in the steps of (4) and (5). Moreover, each party has obtained a new pseudonym in the steps of (4) and (5).

V. Security and Privacy Analysis

In this section, we will analyze the performance of our scheme in terms of security and privacy.

A. Security Analysis and Discussions

1) Security Analysis: Mutual Authentication: According to the section IV-C and IV-D, it can be seen that U_1 and U_2 can achieve their mutual authentication with the help of *AS*.

Anonymous Authentication: AS generates a random pseudonym for every user in the system initialization. During authentication, U_2 cannot get $\mathcal{B}_1$ and $\mathcal{F}_1$ encrypted by K_{AS}. Therefore, the user can complete the anonymous communication. After authentication, a new pseudonym is generated for each user.

Nonlinkability: Nonlinkability means that, for both insiders and outsiders, (1) neither of them could ascribe any session to a particular user, and (2) neither of them could link two different sessions to the same user [23], [24]. In our proposed scheme, U_2 cannot ascribe some sessions to U_1. Firstly, because $\mathcal{B}_1$ and $\mathcal{F}_1$ are always combined with a fresh nonce and encrypted by K_{AS}, U_2 cannot decrypt them. Secondly, anonymous authentication between U_1 and U_2 can be achieved, because they communicate with pseudonyms.

Fresh session keys: How to generate fresh session keys has been given in IV-D. Data traffic is secured, and data confidentiality and integrity can be provided using symmetric cryptography.

Backward Security: Backward secrecy guarantees that a passive adversary who knows a contiguous subset of session keys cannot discover the preceding session keys. Our scheme provides perfect backward secrecy due to the hard discrete logarithm problem in AMP+.

Reply Attack: Each principal can believe the freshness of the random number contained in the messages. The random number helps the principals to deny replay attack.

DOS Attack: Resistance to DOS attack is achieved by the two ways. Firstly, a user's pseudonym always changes as a new session begins, so it is difficult for attackers to associate different pseudonyms with the same user. Secondly, it can be seen that the first part of every message is a hash output in section IV-C and IV-D, which changes regularly after each round. For example, the first part is $h(S)$ in the message of (5) and (6), however, the first part is $h(S)$ in the message of (5) and (6), yet the first part turns into $h(h(S) \oplus S)$ in the message of (7) and (8).

On-line Guessing Attack: Resistance to online guessing attack is an inherent characteristic of AMP+. If the adversary is an active participant in a protocol, the success probability of a guessing attack against the password will be negligible. Our AMP+ protocol can defend against such type of attack because it is related to the hard discrete logarithm problem (DLP). It's clearly that a passive eavesdropper would not be able to compute the shared session key, unless he knows both $\mathcal{R}_1$ and $\mathcal{R}_2$. If an active adversary masquerades as one of the principals, such as U_1, he must guess the correct *S*, else U_2 will abort the protocol immediately without revealing any message.

Off-line Guessing Attack: Similarly, as the analysis of the on-line guessing attack discussed above, the adversary must know $\mathcal{R}_{1(2)}$ and $\mathcal{R}_{2(2)}$ to compute the correct *K*, which is impossible because of the hard discrete logarithm problem in AMP+.

Biometric Privacy: When U_1 registers with *AS*, *AS* stores $\mathcal{B}_1$ instead of his biometric plaintext. $\mathcal{B}_1$ does not reveal any information about $U_1's$ biometrics and provides an excellent privacy protection. Even if the key used to generate $\mathcal{B}_1$ is compromised, the biometrics cannot be leaked. Therefore, Biometric Encryption technique can protect the privacy of U_1's biometric template.

Differentiated Service Access Control: Differentiated service access control can be achieved by classifying mobile users into different service types. Bioscrypt is generated by binding biometrics with the key, so it varies with the different keys or the different biometrics. It is possible for a user to obtain multiple certificates with his single biometrics. Differentiated access control policies can be achieved for different services or different users requesting the same service.

Flexibility and Scalability: If service types are carefully defined and the service providers are well classified, higher differentiated service access control can be achieved, which will improve the flexibility and scalability of our proposed scheme.

Multiple/Cancelable/Revocable identifiers: Bioscrypt, similar to a digital certificate, can be seen as a user's identifier. By binding with different keys, a user can get multiple identifiers. Different identifiers are independent. Even if a single account identifier is compromised, other identifiers cannot be comprised. The comprised identifier may be revoked or replaced by a newly generated one.

We compared our scheme with several related schemes in terms of storage, protocol comparisons. We evaluated the several metrics using probability analysis in Section V-A1. We simulated our scheme in a realistic network setting and present our results in Section V-B.

The used notations are listed in Table II.

TABLE II
LENGTH OF PARAMETERS

Notation	Meaning
l_b	Length of Bioscrypt
l_r	Length of random number
l_h	Length of hash value
l_k	Length of key
l_{id}	Length of ID
m	The number of mobile users
n	The number of service providers

2) Storage Analysis: While the protocol is running, every mobile user only stores his Bioscrypt, ID_u, K_{AS}, K_A and K_A^{-1} as fixed parameters. Some random numbers and the session key are stored as temporary parameters. If the key length is 1024bit and the hash function is MD5, the whole storage is less than 1M, which suits the mobile devices with limited resources. The storage requirement is shown in Table III. U_1 and U_2 stand for the source and the destination respectively.

3) Protocol Comparison: In Table IV, we compare our scheme with anonymous authentication protocol proposed by

TABLE III
STORAGE PERFORMANCE

	Permanent storage	Temporary storage
U_1	$l_b + l_k + l_{id}$	$2l_r + l_k$
U_2	$l_h + l_k + l_{id}$	$2l_r + l_k$
AS	$m(l_h + l_k) + n(l_k + l_{id}) + 2(l_h)$	$2l_r$

He et al. [25] and Ren et al. [23]. Please refer to Section IV-B for detailed discussions on our scheme's security properties.

TABLE IV
PROTOCOL COMPARISON

	He [25]	Ren [23]	Ours
Anonymous authentication mechanism	BS	BS	BioS
MAC address protection	N	N	Y
Correctness proof	N	N	Y
Nonlinkability	N	Y	Y
Differentiated service access control	N	Y	Y
DOS resilience	N	Y	Y

[1] BS means Blind Signature and Bios means Bioscrypt.

4) Probability Analysis and Discussion: The probability of K-anonymity: A subject is considered as *K*-anonymity with respect to location information, if and only if the location information sent from one mobile user is indistinguishable from the location information of at least $K-1$ other mobile users [26]. In our scheme, we use *K*-anonymity to evaluate the privacy protection degree of a single user.

Firstly, assume that the arrival rate of mobile users is to follow a Poisson distribution, the probability that there are *n* users is:

$$P(n) = (\frac{\lambda}{\mu})^n \frac{e^{-\frac{\lambda}{\mu}}}{n!} \tag{1}$$

Each mobile user will stay in the system for an average of $\frac{1}{\mu}$ time. If *K*-anonymity is met, there are at least *k* users staying in the system. The success probability of *K*-anonymity follows:

$$p(x > k) = 1 - p(x \leq k) = 1 - \sum_{i=0}^{k-1} p(i) = 1 - \sum_{i=0}^{k-1} (\frac{\lambda}{\mu})^i \frac{e^{-\frac{\lambda}{\mu}}}{i!} \tag{2}$$

From the equation 2, we can see that the probability of *K*-anonymity is relevant to μ, λ and *K*. Fig. 4 shows that the *K*-anonymity probability is inversely proportional to *K* and μ. In Fig. 5, the *K*-anonymity probability is inversely proportional to λ.

In the worst case, when there are two users in the system, it is easy for U_2 to infer U_1's identity. However, the probability is fairly small in this case. Suppose that the time during which U_1 stays with U_2 obeys exponential distribution:

$$P(x) = \delta e^{-\delta x} \tag{3}$$

t_{int} is defined as the time interval that U_1 changes his pseudonym. Therefore, the probability that U_1 is tracked out

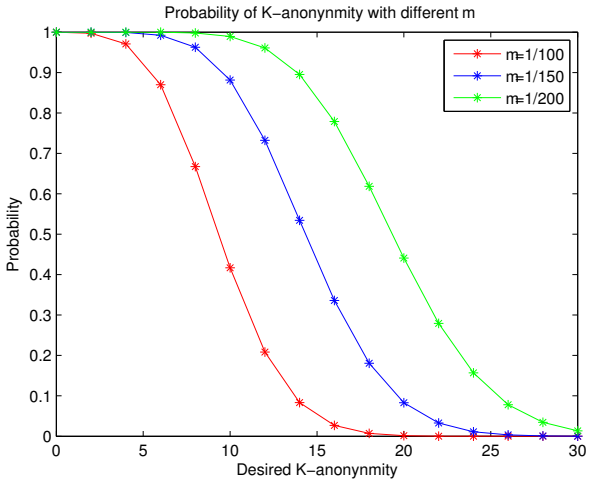

Fig. 4. The Relationship between *K*-anonymity and μ

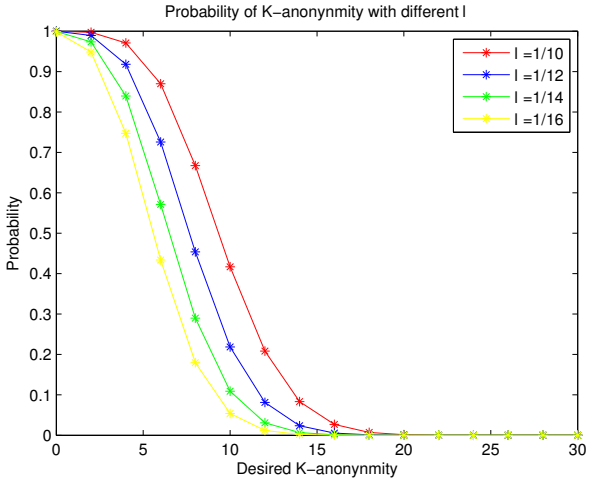

Fig. 5. The Relationship between *K*-anonymity and λ

can be deduced:

$$P_{track(x)} = \delta e^{-\delta x}((\frac{x}{\mu})e^{-\frac{\lambda}{\mu}})^{\frac{x}{t_{int}}} \tag{4}$$

From Fig. 6, we note the probability of being traced is around 10^{-3} magnitude. In Fig. 6, the probability is inversely proportional to t_{int}.

Fig. 4 to 6 show that longer sojourn time and frequently changing pseudonyms can bring higher probability of *K*-anonymity.

B. Simulations

Our simulations were constructed based on the Simulation for Urban Mobility platform [27]. SUMO is an open source traffic simulation package including net import and demand modeling components. A region in the city of Beijing, China was downloaded from the OpenStreetMap [28] application, then it is abstracted into Fig. 7. In our experiments, we selected a rectangular area of 1500 by 1000 meters. We randomly

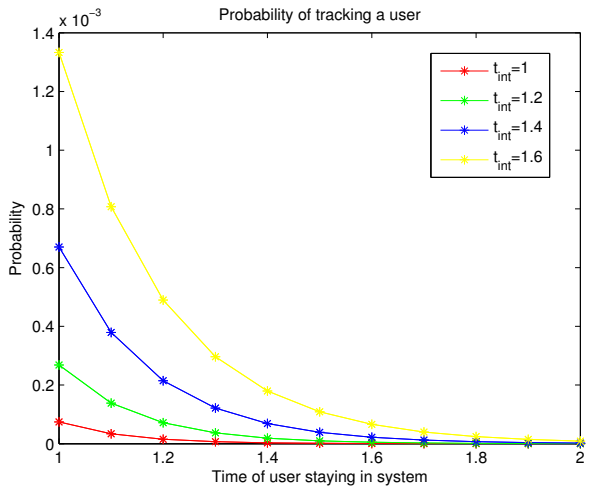

Fig. 6. The Relationship between Tracking U_1 and t_{int}

generated three sets of vehicles with different average numbers (50, 60, and 70) and an average vehicle speed of 70 m/s. We considered two vehicles providing anonymity to each other if their distance is shorter than 100 meters. In the initial stage, in every second, a new vehicle will be added into the system. The departure place and the destination of each vehicle are randomly selected from the map. When calculating the trip to the destination, the vehicles adopt the Dijkstra algorithm to find the nearest way.

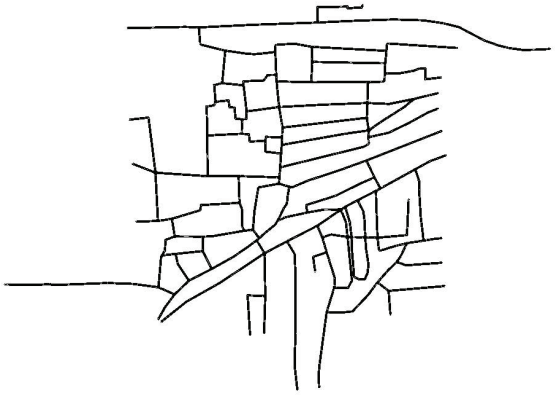

Fig. 7. The abstracted road map

1) Average Anonymity: In our experiments, the user anonymous number is defined as the number of other users to protect the true user's location information. Thus, the location information of the protected user is indistinguishable from the other users. When more than one user communicates with the server, the attackers cannot successful trace a user based on the user's MAC address, because every user's MAC address changes frequently. If the distance between two users u_i and u_j is less than the length D, u_i or u_j can treat the other side as the anonymity member. At a certain moment, all the anonymity member of u_i constitutes u_i's anonymous set. The average degree of anonymity is defined as the average size of all the users' anonymous sets at a certain moment. By experiments, we can draw the data in Fig. 8.

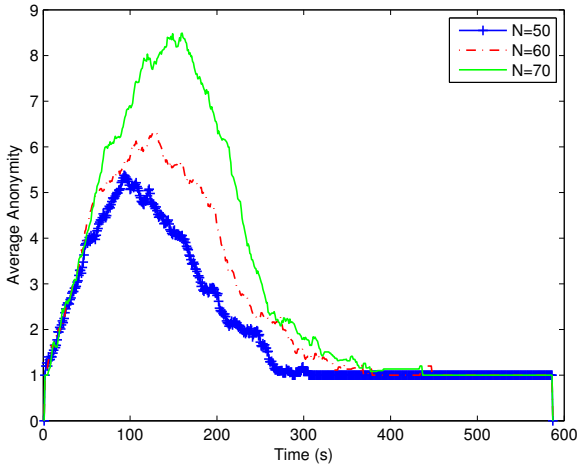

Fig. 8. The average anonymity

In Fig. 8, x-axis represents time, and y-axis represents the average anonymity. The value of the average anonymity decreases after the initial increase. During the initialization, there are fewer users, so the user cannot find other users in his anonymity set and the value of y-axis is one. With the simulation running on, more and more vehicles are added into the system. Moreover, due to the dynamicity of the users, each users can establish the anonymity set with other users more easily. Therefore, the average anonymity set of the system increases. But with the simulation running on, when some of the users arrive at the destination, they will leave the system. This phenomenon causes the reduction of the number of the users in the system, which reduce the average anonymity. At last, when all the users leave the system, the average anonymity becomes 0.

2) Attack Probability: In our scheme, every user can protect his identity by changing his MAC. It is difficult for attackers to correlate the same user by analyzing the MAC addresses. Our experiments focus on the success rate of resisting the correlation attack. Correlation attack means that attackers can correlate the same user by his location information, while not his MAC address. For example, if there is only one user in a region, attackers can trace him successfully whatever MAC address he adopts and wherever he is. We have conducted some experiments on the average probability of the success correlation attack. The experiments results are listed in Fig. 9. x-axis represents time, and y-axis represents the average attack probability. Initially, when the number of the users is small, the probability of linking the identity of the users with the MAC addresses is high. Because, in such a circumstance, the users are sparsely distributed in the system. The distances between users are far, although the users can conceal the MAC address when communicating, the attacker can still mount a correlation attack to figure out the real identity of the users.

When the number of the users increases, the density of the users becomes high, which thus causes the decrease of the probability of correlation attack. Therefore, the probability of attack gradually decreases. After a period of time, some of the users arrive at the destination and leave the system, which causes the reduction of the number of the users in the system. This causes the probability of figuring out the identity of the user to grow up again. At last, the probability reaches 1.

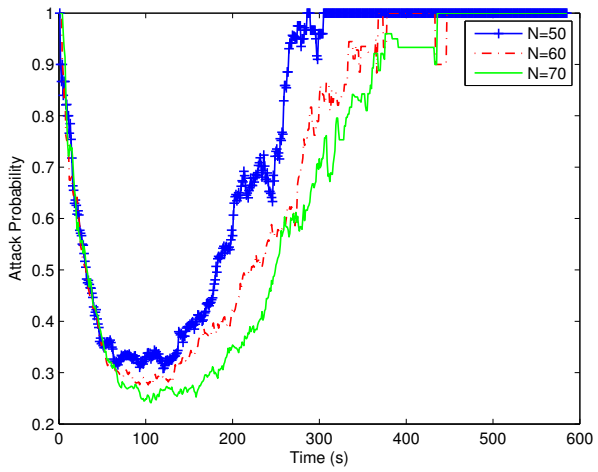

Fig. 9. The attack probability

During the initialization, there are fewer users, so the user can trace the user at the probability of 100%. When the number of users gradually increases, the value of y-axis will increase. When users gradually leave, the value will reduce to one. Fig. 9 shows that more users can bring better security.

3) Entropy: In information theory, entropy is a measure of the uncertainty associated with a random variable. The entropy can be defined as $H = -\sum_{i=0}^{n} p_i \log_2 p_i$, where represents the successful attack at one moment. X-axis represents time, and y-axis represents the entropy value. First, the probability of success correlation attack for every user is calculated. Then, the system entropy is calculated. The experiment result is shown in Fig. 10.

When the number of gradually increases, the value of y-axis will increase. When users gradually leave, the value will reduce to zero. Fig. 10 shows that more users can bring better security.

VI. Conclusion

In this paper, we have proposed an anonymous authentication scheme based on Biometric Encryption in VANETs. Biometric Encryption not only ensures authentication convenience, but also secures the biometric template. A shared session key is derived between two vehicles, and besides that, a pseudonym for every vehicle is negotiated. Security analysis proves that our scheme can resist multiple attacks. Privacy analysis shows that K-anonymity is achieved to protect the identity for every user.

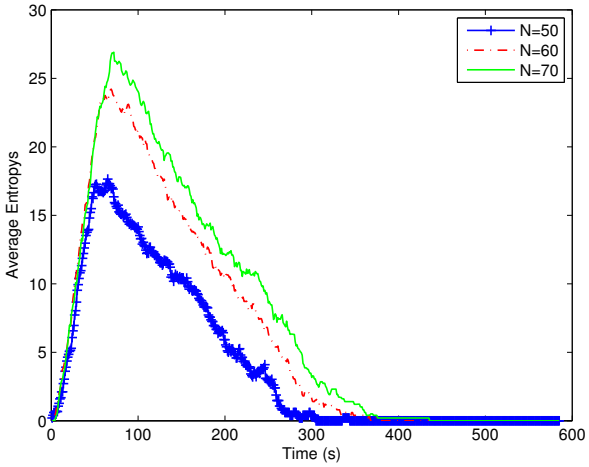

Fig. 10. The Entropy

As for our future work, we will investigate the location privacy issue under the context of our proposed sheme.

References

[1] G. Yan, S. Olariu, and M. Weigle, "Providing vanet security through active position detection," *Computer Communications*, vol. 31, no. 12, pp. 2883–2897, 2008.

[2] C. Li, M. Hwang, and Y. Chu, "A secure and efficient communication scheme with authenticated key establishment and privacy preserving for vehicular ad hoc networks," *Computer Communications*, vol. 31, no. 12, pp. 2803–2814, 2008.

[3] G. Samara, W. A. Al-Salihy, and R. Sures, "Security analysis of vehicular ad hoc networks (vanet)," *2010 Second International Conference on Network Applications, Protocols and Services*, pp. 55–60, 2010.

[4] S. Zhang, J. Tao, and Y. Yuan, "Anonymous authentication-oriented vehicular privacy protection technology research in vanet," in *Electrical and Control Engineering (ICECE), 2011 International Conference on*. IEEE, 2011, pp. 4365–4368.

[5] D. Boneh and H. Shacham, "Group signatures with verifier-local revocation," in *Proceedings of the 11th ACM conference on Computer and communications security*. ACM, 2004, pp. 168–177.

[6] J. Guo, J. Baugh, and S. Wang, "A group signature based secure and privacy-preserving vehicular communication framework," in *2007 Mobile Networking for Vehicular Environments*. IEEE, 2007, pp. 103–108.

[7] X. Lin, X. Sun, P. Ho, and X. Shen, "Gsis: a secure and privacy-preserving protocol for vehicular communications," *Vehicular Technology, IEEE Transactions on*, vol. 56, no. 6, pp. 3442–3456, 2007.

[8] G. Calandriello, P. Papadimitratos, J. Hubaux, and A. Lioy, "Efficient and robust pseudonymous authentication in vanet," in *Proceedings of the fourth ACM international workshop on Vehicular ad hoc networks*. ACM, 2007, pp. 19–28.

[9] M. Raya and J. Hubaux, "Securing vehicular ad hoc networks," *Journal of Computer Security*, vol. 15, no. 1, pp. 39–68, 2007.

[10] B. Bellur, "Certificate assignment strategies for a pki-based security architecture in a vehicular network," in *Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008. IEEE*. IEEE, 2008, pp. 1–6.

[11] R. Lu, X. Lin, H. Zhu, P. Ho, and X. Shen, "Ecpp: Efficient conditional privacy preservation protocol for secure vehicular communications," in *INFOCOM 2008. The 27th Conference on Computer Communications. IEEE*. IEEE, 2008, pp. 1229–1237.

[12] A. Wasef, Y. Jiang, and X. Shen, "Dcs: An efficient distributed-certificate-service scheme for vehicular networks," *Vehicular Technology, IEEE Transactions on*, vol. 59, no. 2, pp. 533–549, 2010.

[13] Y. Jiang, M. Shi, X. Shen, and C. Lin, "Bat: a robust signature scheme for vehicular networks using binary authentication tree," *Wireless Communications, IEEE Transactions on*, vol. 8, no. 4, pp. 1974–1983, 2009.

[14] L. Huang, K. Matsuura, H. Yamane, and K. Sezaki, "Enhancing wireless location privacy using silent period," in *Wireless Communications and Networking Conference, 2005 IEEE*, vol. 2. IEEE, 2005, pp. 1187–1192.

[15] L. Huang, H. Yamane, K. Matsuura, and K. Sezaki, "Silent cascade: Enhancing location privacy without communication qos degradation," *Security in Pervasive Computing*, pp. 165–180, 2006.

[16] C. Zhang, R. Liu, P. Ho, and A. Chen, "A location privacy preserving authentication scheme in vehicular networks," in *Wireless Communications and Networking Conference, 2008. WCNC 2008. IEEE*. IEEE, 2008, pp. 2543–2548.

[17] L. Yao, X. Kong, and Q. Fan, "A privacy-preserving authentication scheme using biometrics for pervasive computing environments," *Journal of Electronics*, vol. 27, pp. 68–78, 2010.

[18] J. Anil K, N. Karthik, N. Abhishek *et al.*, "Biometric template security," *EURASIP Journal on Advances in Signal Processing*, vol. 2008, 2008.

[19] Y. Sun, R. Lu, X. Lin, X. Shen, and J. Su, "An efficient pseudonymous authentication scheme with strong privacy preservation for vehicular communications," *Vehicular Technology, IEEE Transactions on*, vol. 59, no. 7, pp. 3589–3603, 2010.

[20] K. El Defrawy, J. Solis, and G. Tsudik, "Leveraging social contacts for message confidentiality in delay tolerant networks," in *Computer Software and Applications Conference, 2009. COMPSAC'09. 33rd Annual IEEE International*, vol. 1. IEEE, 2009, pp. 271–279.

[21] A. Kate, G. Zaverucha, and U. Hengartner, "Anonymity and security in delay tolerant networks," in *Security and Privacy in Communications Networks and the Workshops, 2007. SecureComm 2007. Third International Conference on*. IEEE, 2007, pp. 504–513.

[22] T. Kwon, "Authentication and key agreement via memorable password," in *ISOC Network and Distributed System Security Symposium*, vol. 20. Citeseer, 2001, pp. 31–33.

[23] K. Ren and W. Lou, "Privacy-enhanced, attack-resilient access control in pervasive computing environments with optional context authentication capability," *Mobile Networks and Applications*, vol. 12, no. 1, pp. 79–92, 2007.

[24] B. Gedik and L. Liu, "Protecting location privacy with personalized k-anonymity: Architecture and algorithms," *Mobile Computing, IEEE Transactions on*, vol. 7, no. 1, pp. 1–18, 2008.

[25] Q. He, D. Wu, and P. Khosla, "The quest for personal control over mobile location privacy," *Communications Magazine, IEEE*, vol. 42, no. 5, pp. 130–136, 2004.

[26] P. Kamat, Y. Zhang, W. Trappe, and C. Ozturk, "Enhancing source-location privacy in sensor network routing," in *Distributed Computing Systems, 2005. ICDCS 2005. Proceedings. 25th IEEE International Conference on*. IEEE, 2005, pp. 599–608.

[27] "http://sumo.sourceforge.net/."

[28] "http://www.openstreetmap.org/."