Mobile-Aware Anonymous Peer Selecting Algorithm for Enhancing Privacy and Connectivity in Location-Based Service

CONFERENCE PAPER · DECEMBER 2010

DOI: 10.1109/ICEBE.2010.32 · Source: IEEE Xplore

4 AUTHORS:

Jian Xu
Hangzhou Dianzi University
18 PUBLICATIONS **34** CITATIONS
SEE PROFILE

Z.P. Jin
Deggendorf University of Applied Sciences
3 PUBLICATIONS **7** CITATIONS
SEE PROFILE

Ming Xu
Hangzhou Dianzi University
42 PUBLICATIONS **89** CITATIONS
SEE PROFILE

Ning Zheng
Hangzhou Dianzi University
40 PUBLICATIONS **81** CITATIONS
SEE PROFILE

Mobile-aware Anonymous Peer Selecting Algorithm for Enhancing Privacy and Connectivity in Location-based Service

Jian Xu, Z.P. Jin, Ming Xu, Ning Zheng
Computer Science Department of
Hangzhou Dianzi Technological University
HangZhou, China, 310085
{jian.xu, jin_z_p mxu, nzheng}@hdu.edu.cn

Abstract—**Privacy preserving in location-based service (LBS) has been an important issue in Vehicular Ad hoc Network (VANET). Traditional k-anonymity algorithm dealing with location privacy protecting problem does not consider vehicle's mobility and inner collaborator. In this paper, we propose a new algorithm which uses dynamic and mobile-aware anonymous peer selecting algorithm to improve the anonymity and connectivity in a mobile environment. We also give out detailed analysis and show experimentally that better privacy and connectivity can be achieved with this method.**

Keywords-Mobile-aware; privacy; location-based service;

I. INTRODUCTION

Location-Based Service (LBS) [1] is an important application in Vehicular Ad hoc Network (VANET). It provides various services to mobile user based on its location. For examples, "Finding the nearest gas station" and "Finding all of the restaurants on my route to office". The results of the two LBS queries depend on the location of user. Emerging position-detection (e.g. GPS, RFID) and wireless short-range communication techniques (e.g. Bluetooth, IEEE 802.11p and DSRC) help vehicles to position accurately and improve communication capacity. When a mobile user launches a LBS application, such as traffic information querying, street-level routing and location-based adverting, the LBS server will ask for the accurate location of user. The more accurate location offer, the better quality service get.

However, it brings a crucial security issue in privacy preserving [2] [3]. Location information of a vehicle is private. Untrustworthy LBS server may expose user's location to adversary. Then adversary can infer user's habit and interests by knowing the places user visits and the frequency of each place. To solve the problem, a peer-to-peer k-anonymity algorithm [6] is proposed. As an effective solution, the algorithm composes a *k* peers (vehicles) group though a spatial cloaking algorithm and randomly picks a leader peer acting as a communication agent to LBS server. The initial peer only exposes its location to agent peer. Then the server can't distinguish the initiator from other k-1 peers. But the algorithm has a flaw that it builds on a trust to agent peer. It will be unsafe if the agent is captured by the adversary. Meanwhile, the algorithm ignores the mobility of peers which brings a high disconnection rate in mobile network. In this paper, we propose a mobile-aware anonymous peer selecting algorithm to enhance the privacy and connectivity. Contributions of this paper are follows:

1) We design an anonymous architecture to protect the privacy against untrustworthy agent.
2) We propose an algorithm considering with the fast mobility of vehicles on the highway and reduce the communication disconnect rate.
3) We also provide experimental evidence to prove the superiority of our approach in anonymity, connectivity and scalability.

The rest of the paper is organized as follows: Section 2 surveys the related work. In section 3, we expand a universal system model. We describe the architecture and give a theoretic analysis in section 4. Section 5 gives experimental evidences and discusses about the results. Finally, we conclude the paper in section 6.

II. RELATED WORK

Location privacy preserving is to achieve the unlinkability [12] [13] [14] between identity information and location information. There are two typical approaches: identity protection and location protection.

Identity protection approach protects the identity from eavesdropped but allow exposing location information. Anonymity [4] [5] is a simple solution but loses the system authentication. Gruteser and Grunwald propose k-anonimity model [4] through spatial and temporal cloaking for VANET. It composes a k peers group and keeps the initiator indistinguishable from other k-1 peers. But the group composing algorithm is complex and the member peers are dynamic. ChiYin Chow et al improved the model in Peer-to-Peer architecture [6]. It builds on a strong trust to agent peer. Pseudonymity [7] is a special anonymity approach using a pseudonym instead of user's identity. It gives a consideration to anonymity and authentication but is not reliable under data mining technology. Beresford and Stajano [8] give another model——Mix-zone. It defines two zones: application zone and mix zone. Users keep silent in mix zone. The communication only occurs in the application zone. The Mix-zone is effective for privacy preserving and authentication but it is hard to be deployed in large scale system.

Location protection approach has a different focus from identity protection approach. It pays more attention to location data and allows revealing real identity. Kido H et al propose a dummy algorithm [9]. It offers a lot redundant location data to the LBS server which reduces the usability of the location data collected from users. Obfuscation [10] is a similar algorithm. It exposes a regional location instead of the accurate location. The k-anonymity approach [6] is also a location protection algorithm. It uses a regional location which cloaks all of the k peers instead of accurate location. All of those approaches have good authentication but reduce the quality of service because the server unable to collect accurate location.

III. P2P K-ANONIMITY MODEL

A. *Networking Model*

Figure 1 illustrates a typical VANET model that consists of three major domains: content provider domain, network provider domain and user domain. Content provider domain is comprised by several commercial content provider units (CPUs). They provide services to users. Network provider domain provides internet access interfaces though some roadside units (RSUs). In the user domain, peers with on board units (OBU) constitute a vehicular ad hoc network. There are two type communications: vehicle to vehicle (V2V) and vehicle to infrastructure (V2I). If the vehicle is in the transmission coverage of the RSU, it can access to Internet directly via a V2I communication. Otherwise, it can build a path with single or multi hops vehicle V2V and a V2I communication to access the Internet.

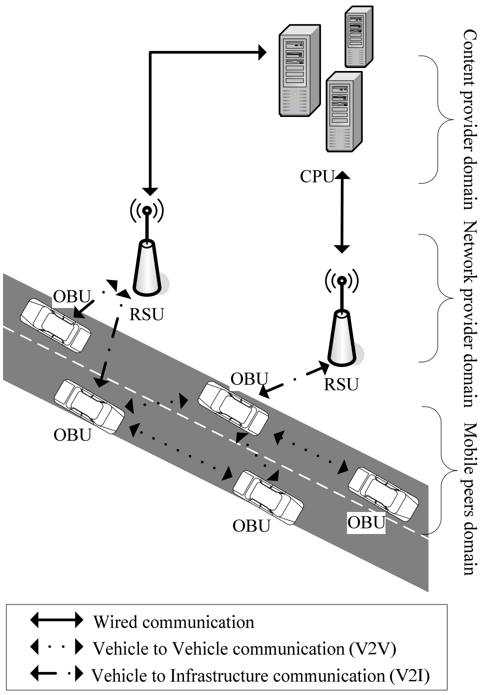


Figure 1. The Vehicular Ad hoc Network Architecture

The more detailed of VANET is depicted in Figure 2. An intelligent vehicle is equips four main units: On-board processor is in charge of controlling and coordinating other devices. Positioning device gets exact location data from the Global Positioning System (GPS). Sensor and database collect both internal information (e.g. velocity, direction, acceleration) and external information (e.g. temperature, road condition). Communication interface with several network cards offers communication capacity with roadside infrastructure and neighbor peers. There is also a third part (Registration Authority) to authenticate or authorize new peers.

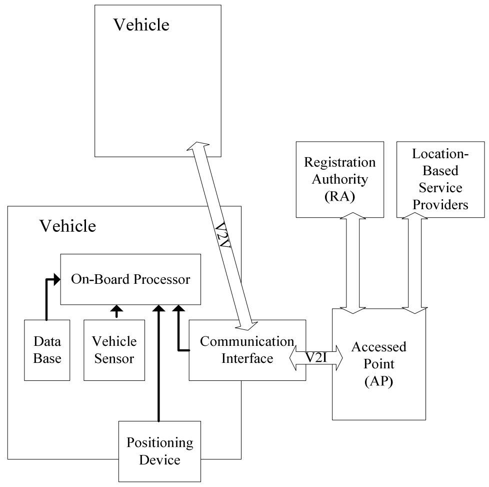


Figure 2. The Detail of VANET Architecture

B. *P2P k-anonymity model*

The P2P k-anonymity model is proposed by ChiYin Chow et al. The algorithm constitutes the initial peer and other k-1 mobile peers to a group to achieve the user's privacy profile with two parameters: k and A_{min}. k indicates the indistinguishable degree for the group (contain k different peers). So the initiator can not be distinguished from other k-1 peers and it will be k-anonymity. A_{min} means the minimum resolution of the cloaked spatial region. The privacy profile can be customized by user at any time.

The P2P k-anonymity algorithm is to find a regional cloaking that covers k peers. It has several steps: 1) Select a central peer who will act as a agent for the group. 2) The central peer will discover other k-1 different peers via single-hop or multi-hop to compose the group. 3) Find a cloaked region covering all locations that every peer may arrive. 3) Adjust the cloaked region to A_{min} parameters. Once the cloaked region is less than A_{min}, the region will be expanded. Figure 3 shows a running example of the algorithm.

A query from initiator to LBS server includes some necessary information which composes a tuple. We assume that its normal format is:

$$T \rightarrow \{SID, DID, L, E[query]\}$$

SID, DID denote source peer's ID and destination peer's ID respectively. L means the spatial cloaked location. E[query] is the semantic query content which is encrypted.

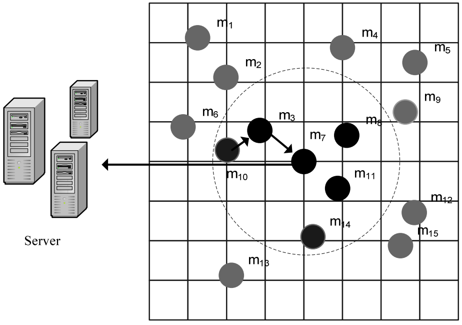


Figure 3. An Running Examlpe of P2P Architecture

From the example in Figure 3, the request initiator m_{10} collects other 5 peers (represent as solid black circles) for the user's privacy profile k=6 (we don't discuss A_{min} here). Firstly, it selects m_7 as its agent (central peer). And then the central peer discovers other 5 peers through one hop (m_3, m_8, m_{11}) searching or two hops searching (m_{14} and m_{10}). After the group is created, the initiator will send a query to the agent first:

$$T_{initiator} \rightarrow \{m_{10}, m_7, L, E[query]\}$$

The agent forward the tuple to server:

$$T_{agent} \rightarrow \{m_7, Server, L, E[query]\}$$

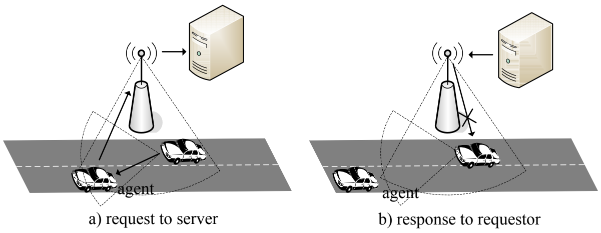
As a result, the server receives a query from m_7 and the initiator is unknown from the server.

C. Anonymity and Connectivity

Without a doubt, the k-anonymity model achieves the user's privacy profile effectively. It preserves the unlinkability between initiator's ID and its exact location to LBS server. But it is not flawless.

One of the problems is that the P2P k-anonymity algorithm builds on a strong inner trust and just considers the anonymity to server. But sometimes a collaborator may be in the group and capture the agent. We can summarize three types of attackers: 1) **Local eavesdropper.** The attacker invades the user's computer and can observe all messages the user sends. 2) **Collaborator eavesdropper.** The collaborator is on the path from initiator to agent (include agent), the sensitive information may be exposed to the attacker because of the plaintext of SID and DID. 3) **Server eavesdropper.** Server eavesdropper can observe all the messages that the server receives. Local eavesdropper is a global attacker and can't be defeated in this model. Once the user's computer is invaded by attacker, we provide no algorithmic protections against it even the message is encrypted. Generally, only the query content is encrypted while SID and DID are plaintext for routing. So a collaborator eavesdropper can know the indeed sender and receiver if the routed path gets through it. But it can't decrypt the message and know the real content. The attacker at the server side can decrypt the encrypted message and record the direct sender. But if both the server and the agent are captured by attackers, the linkability between location and initiator may be revealed to attackers. The privacy of the user may be exposed.

Another problem is the variable network topology (because the fast moving of vehicles) which brings a high disconnect rate and brings down the system's availability. Figure 4 shows an example of disconnection. Initiator chooses an agent randomly to send its query to server. Sometimes two peers are at the ultimate transmission range (Figure 4a). But when the server responses to the requester, the agent has moved out of both the infrastructure and initiator's transmission range (Figure 4b). It makes a disconnection and the high speed makes the disconnection rate high. The high speed moving of vehicles exacerbate the situation.



a) request to server

b) response to requestor

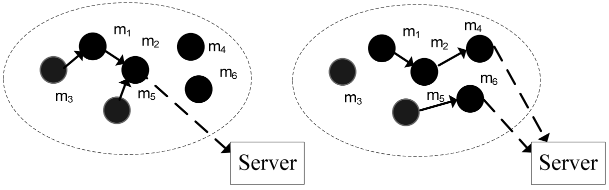
Figure 4. An example of disconnection

IV. System Model

In this section, we will expand our anonymous architecture and mobile-aware peer selecting algorithm.

A. Anonymous Architecture

Anonymous architecture is non-central and self-organization architecture. In the architecture, all peers are anonymous peers. *Anonymous peer* is only capable to know from whom a query is received and determine to whom the query will be forwarded in one hop. But it has no knowledge of the initiator and the final receiver. Figure 5 shows the difference between P2P architecture and anonymous architecture. The P2P architecture is central architecture. Both of the two architectures create a communication path from initiator to LBS server. But the SID and DID are expose to every peer on the path in P2P architecture while



a) P2P architecture

b) Anonymous architecture

Figure 5. The difference of P2P architecture and anonymous architecture

the peers in the anonymous architecture only knows its last peer and next peer.

In anonymous architecture, we also need to achieve a spatial cloaking area that is more than A_{min} as well as covers k indistinguishable anonymous peers. Both of the two requirements can be customized by user. Algorithm 1 gives the pseudo code for the anonymous architecture. The architecture usually consists of three phases:

Phase 1: Peers Discovering. The initiator will broadcast a hello message to neighbors to discover new peers. Those new peers are stored in a set T. $k'=|T|$ denotes the size of the set.

Phase 2: Mobile-aware selecting. If the discovered peers are not up to k, we need to expand the peer set. We define a function $f(d_i,\theta_i,\tau_i,\upsilon_i)$ to measure the disconnect probability of two moving peer (see the detail in next subsection). Then we can select a stronger connection peer as the next hop of the path.

Phase 3: Adjustment. Phase 1 and Phase will be repeated until discovers k peers. And then we compute an area covering of all discovered peers. It should be expanded if less than A_{min}(in line 7).

Once both of the requirement k and A_{min} are achieved, the initiator could send the query request to server through the communication path. The last peer will act as agent that communicates with server.

```
Algorithm 1: Mobile-Aware Cloaking
1:Function Mobile-awareCloaking(k, $A_{min}$)
2:Originally, the number of discovered peers k'=0
while k'<k-1 do
// Phase 1: Peers Discovering
3:Discovering new peers to a set T.
4:k'= k'+$|T|$
// Phase 2: Mobile-aware selecting
5:for all peers in T do
    $V_i= f(d_i,\theta_i,\tau_i,\upsilon_i)$
  end for
// Phase 3: Adjustment
6: if k'<k-1 then
    next=max($V_1$,$V_2$…)
  end if
end while
7: Expland covering area of discovered peer to $A_{min}$
   if necessary
```

B. Mobile-aware peer selecting algorithm

The range of those wireless short distance transmission is limited, such as IEEE802.11p, DSRC (about 500m). The vehicle's rapid movement brings a high disconnect rate in the vehicular ad hoc network. Once two connected vehicles move out of the transmission range, the connection will be broken. The topology of the network is changing at any time.

The probability of two moving peers disconnect is related to the relative movement between them. For example, two vehicles which are moving toward each other will have lower probability of disconnection than two which are not. We define S_t to measure the capability of two vehicles keeping connected at time *t*. The value of S_t is calculated as $S_t=\frac{r}{D}$. r means the transmission range while D denotes the distance between two peers. It means that two near vehicles (with a high value of S_t) have higher probability to keep connection than two far away vehicles. If the value S_t <1 (means out of the transmission range), the connection will be broken.

Figure 6 shows a pervasive model of two moving vehicles. In a period τ (from t to t'), two vehicles at m_1 and m_2 move to m_1' and m_2' with the speed of v_1 and v_2 respectively. The distance changes from d to d'. υ is the vector sum of $v_1$and v_2. But now we give it a new division: υsinθ and υcosθ (θ is the angle from υ and d). υsinθ is a component of υ changes the direction of d while υcosθ changes the value of d. The probability of the two peers will keep connected at t' is:

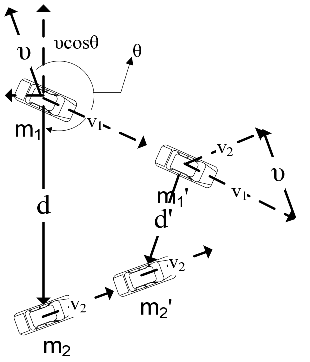


Figure 6. The mobility analysis between two vehicles

$$P(S_{t'}\ge 1)=P(\frac{r}{d+|\upsilon|\tau\cos\theta}\ge 1)=P(|\upsilon|\le\frac{r-d}{\tau\cos\theta})$$

We give a presumption that the car velocity $|\upsilon|$ (the quantitative value of υ) is from a normal distribution $N(\mu,\sigma)$ with a mean μ and variance σ . The probability can be calculated as follows:

$$P(S_{t'}\ge 1)=P(|\upsilon|\le\frac{r-d}{\tau\cos\theta})=\int_0^{\frac{r-d}{\tau\cos\theta}}\frac{1}{\sqrt{2\pi\sigma}}e^{-\frac{(\upsilon-\mu)^2}{2\sigma^2}}$$

From the equation above, we know that the probability is decided by a function with four parameters d, υ, τ and θ:

$$P(S_{t'}\ge 1)=f(d,\theta,\tau,\upsilon)$$

The value of *f* is used to measure the communication capability of two cars. We calculate *f* for all of the neighbor peers and choose an optimal one as next hop. This algorithm is used in last sub section. Comparing to the random peer selecting algorithm, the mobile-aware peer selecting algorithm enhances the connectivity of VANET remarkably considering the mobility for vehicles.

V. EXPERIMENTAL RESULT

In this section, we evaluate the anonymity, connectivity and scalability comparing with P2P architecture.

All of the experiments run on a simulator developed by Thomas Brinkhoff [11]. The mobile peers are generated by the simulator based on a small town's road network. They are randomly distributed in a spatial space of 5000×5000 square meters. Every mobile peer has a presupposed transmission range of 500 meters and randomly determines the speed υ from a normal distribution $N(\mu,\sigma)$ with a mean μ and variance σ. The simulator runs following the system time. At every system time, a record with id, location, velocity and direction is created for every peer.

A. Anonymity

We assume that we totally have n peers with c collaborators. For the k-anonymity, the initiator routed a l long (l intermediate peers) path to the server. In P2P architecture, once a peer on the path is a collaborator, the initiator will be exposed to attacker. The probability of keeping initiator anonymous P_1 can be calculated as follows:

$$P_1 = \frac{\binom{n-c}{l}}{\binom{n}{l}} = \frac{(n-c)!(n-l)!}{n!(n-c-l)!}$$

In anonymous architecture, the initiator is exposed only if the initiator's next peer is occupied by collaborator. The probability P_2 is:

$$P_2 = \frac{\binom{n-c}{1}}{\binom{n}{1}} = \frac{n-c}{n}$$

From the two equations, if $c<n-1$, the probability $P_2>P_1$. Anonymous architecture has a higher probability to keep initiator anonymous.

Figure 7 shows the comparison of disclosure rate between P2P architecture and anonymity architecture. In the experiment, we generate 500 mobile peers with 2% collaborators on the simulator. As is depicted in Figure 8, the disclosure rate of typical architecture grows along with the length of routed path linearly. But it has almost no influence to the anonymous architecture. Anonymous architecture is better than P2P architecture in anonymity. It is because that in the P2P architecture any collaborators on the routed path can know the initiator ID. But in the anonymous architecture only the collaborator captures the initiator's next peer can know the initiator ID. We can get similar result with respect to increasing percentage of collaborators (Figure 8).

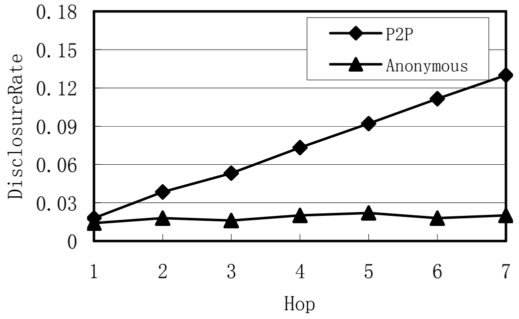


Figure 7. The disclosure rate with respect to path length

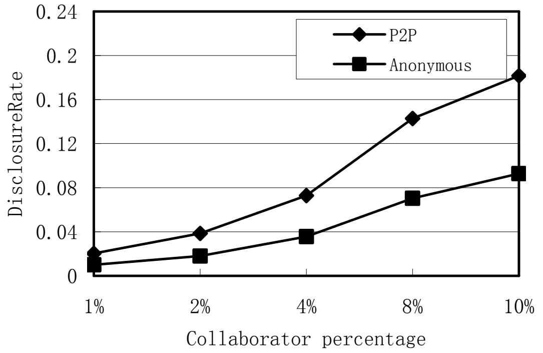


Figure 8. The disclosure rate with respect to collaborator percentage

B. Connectivity

Variable network topology causes a connectivity issue. In the P2P architecture, peer chooses a random mode to create the connection of two moving vehicles while we propose a mobile-aware mode in anonymous architecture.

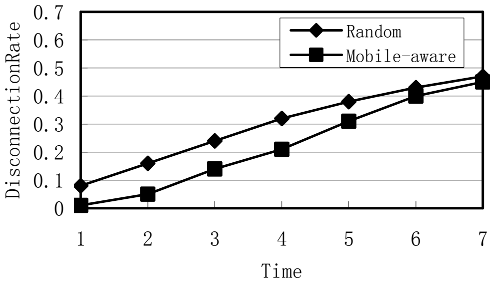


Figure 9. The disconnection rate with respect to time

Figure 9 is a statistical result of the disconnect rate for 1000 simulative peers. All connections are created at system time t=0 at both random mode and mobility-aware mode. The result indicates that the mobility-aware mode gets a remarkably prior for the connectivity especially at the beginning of the connection. The disconnect rate grows with respect to time, because more vehicles move out of original transmission range. This linear growth of disconnect rate is one of the evidence for the sharp variation of network topology.

C. Scalability

Scalability is also an important parameter in VANET. Figure 10 gives the scalability of the comparison about two modes. We increase the simulative peer from 500 to 2000 in the same 5000×5000 spatial space. The connectivity mildly gets better and both of two modes have a good scalability. It is because that a vehicle in a crowded environment has a higher probability to create a connection than in a sparse environment. The result also shows that the mobility-aware mode has lower disconnect rate. It proves that the mobile-aware mode gets better connectivity than random mode.

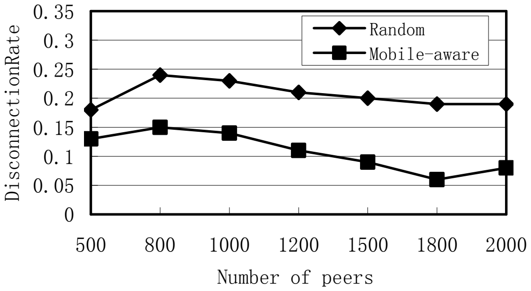


Figure 10. The disconnect rate with respect to number of peers

VI. Conclusion

This paper introduces an enhanced architecture comparing with P2P k-anonymity spatital cloaking algorithm. The architecture includes two respects: anonymous architecture and mobile-aware peer selecting algorithm. The anonymous architecture improves the anonymity via anonymous peers. It is against both of collaborator eavesdropper and server eavesdropper. The mobile-aware algorithm considers the mobility of vehicles and selects a relatively stable peer as its next hop which improves the system connectivity remarkably. We also give analysis and experimental evidence for those two improvements. The results support our proposal solution.

Acknowledgment

The whole work is supported by Science and Technology Program of Zhejiang Province (NO:2008C21075). Thanks very much.

References

[1] Shu Wang, Jungwon Min and Byung K. Yi, “Location Based Services for Mobiles: Technologies and Standards”, IEEE International Conference on Communication (ICC) 2008, Beijing, China.

[2] M. E. Zarki, S. Mehrotra, G. Tsudik, and N. Venkatasubramanian, “ Security issues in a future vehicular network,” in European Wireless,2002.

[3] J.-P. Hubaux, S. Capkun, and J. Luo, “The security and privacy of smart vehicles,” IEEE Security & Privacy, vol. 2, no. 3, pp. 49–55, 2004.

[4] Gruteser M and Grunwald D, “Anonymous Usage of Location-based Services Through Spatial and Temporal Cloaking”, Processing of the International Conference on Mobile Systems, Applications, and Services. MobiSys 2003 :163-168.

[5] Gedik B and Liu L, “A Customizable k-Anonymity Model for Protecting Location Privacy”, ICDCS, 2005.

[6] ChiYin Chow, Mohamed F. Mokbel and Xuan Liu, “A Peer-to-Peer Spatial Cloaking Algorithm for Anonymous Location based Services”, ACMGIS’06, November 1011, 2006, Arlington, Virginia, USA.

[7] Pfitamann A and Kohntopp M, “Anonymity, unobservability, and pseudonymity－a proposal for terminology”, Designing Privacy Enhancing Technologies. Volume 2009 of Lecture Notes in Computer Science. Springer ,2001 :1-9.

[8] Beresford A R, Stajano F, “Location privacy in pervasive computing”. IEEE Pervasive Computing, 2003: 46-55.

[9] Kido H, Yanagisawa Y and Satoh T, “An Anonymous Communication Technique Using Dummies for Location-based Service”, Processing of IEEE International Conference on Pervasive Services. 2005: 88-97.

[10] Duckham M and Kulil L, “A formal model of obfuscation and negotiation for location privacy”, Pervasive, 2005

[11] T. Brinkhoff, “Generating Network-Based Moving Objects”, In Proc. 12th International Conference on Scientific and Statistical Database Management, Berlin, Germany, 2000, pp. 253–255.

[12] John Krumm, “A survey of computational location privacy”, Springer-Verlag Pers Ubiquit Comput (2009) 13:391–399

[13] HaoJin Zhu, “Security in service-oriented vehicular networks”, IEEE Wireless Communications August 2009.

[14] Lars Kulik, “Privacy for Real-time Location-based Services”, SIGSPATIAL Special 2009, Volume 1.