International Journal of Security and Its Applications
Vol.9, No.1 (2015), pp.249-258
http://dx.doi.org/10.14257/ijsia.2015.9.1.24

Privacy Protection for LBS in Mobile Environments: Progresses, Issues and Challenges

Julong Pan [1], Zhengwei Zuo [1], Zhanyi Xu [1] and Qun Jin [1, 2]

[1]*College of Information Engineering, China Jiliang University, Hangzhou, China*
[2]*Department of Human Informatics and Cognitive Sciences, Waseda University, Tokorozawa, Japan*
pjl@cjlu.edu.cn; zzwstriver@163.com

Abstract

In recent years, mainly driven by the availability of modern mobile devices with cheap integrated position sensors, location based services (LBS) have become more and more popular. Prominent examples are nearest friends finding in mobile social networks and the points of interest finders such as the nearest gas stations, hospitals, or places of interests etc. However, the users' privacy information such as location is threatened, when users enjoy the convenience and effectiveness provided by such location services. Therefore, how to secure users' privacy information must be taken into consideration. Many different approaches have been proposed to protect users' identity, location and so on. This paper reviews and analyzes existing privacy protection research works from an integrated perspective, gives the LBS category from two views, and discusses the challenges of securing privacy information. Lastly, our suggestions for future research works are presented.

Keywords: *LBS; privacy protection; security; smart phone*

1. Introduction

The convergence of technologies such as *Geographic Information System* (GIS), networks and mobile devices give birth to *location based services* (LBS), In these services, requests which contain users' locations are sent to *Third Service Provider* (TSP), and TSP then responds with some needed information for the users. Common examples are finding the nearest *Places of Interests* (POI) such as gas stations, hospitals, *etc*. In this process, users' privacy like locations is open and threatened. Privacy is generally the information that you don't want others to know. There was news which reported tracing other people with GPS before. With the popularity of LBS, users' privacy information has aroused much concern [1, 2]. Although the security of users' privacy is vigorously advocated by related entities, a seemingly bigger problem is that users themselves show few interests on this.

In his survey, Colbert [3] asked some college students whether they wanted to share their locations with others, the result was that they were so willing that Colbert even doubted the survey's authenticity. In fact, users' privacy information has become a tradable commodity: users give their location information in exchange for free personalized services. Profits gained by such information reversely guarantee the operation of such services. Therefore, from this perspective, it seems that there is no need to put much intervention on users' location information, because both users and service providers seem to have no willingness to protect privacy. But in reality, most users don't understand this process. By this mean, it's necessary to strengthen users' consciousness of privacy data flow from the point of security, thus fairness can be ensured. There have been many representative works about protecting users' privacy so far, such as these mentioned in [4-7]. But one problem is that they only considered just one aspect. For example, [5] just discussed privacy protection techniques, but neglected related policy and law issues, and they all assumed TSP is the potential threat entity. In fact, in LBS, users must first get their own locations which are provided by *Location Provider* (LP). If LPs are not trusted, all security methods will not work correctly.

ISSN: 1738-9976 IJSIA

This paper summarizes and discusses the state-of-art security mechanisms of LBS. Firstly, the architecture, objects and category of LBS are introduced. Since more and more web-based location providers are involved, they will be discussed separately. And two new perspectives of analyzing the privacy protection in LBS are discussed. Lastly, challenges in LBS are described.

2. Architecture, Objects and Category of LBS

The application architecture is showed in Figure 1. In this model, we assume LP is trusted. Users first request their locations from LP. When done, service request together with their locations will be sent to the LBS provider. Then the LBS providers will response with some kind of service after receiving these requests. In this process, we can apparently see the data flow of user's location, which is LP, user, LBS provider and user. Actually, the data flow of user's information is much more complicated in practice. For example, the LBS providers may sell users' information to other individuals or organizations for their profits. But, users' privacy security largely depends on the LBS providers, especially different kinds of LBS providers are rife in recent years; this is also the reason why most research works on privacy protection in LBS focused on the LBS providers.

As mentioned above, we assume LP is trusted. Therefore, we can treat the LBS provider as the main protecting object in LBS, not the LP in Figure 1, neither the users themselves. If this is not the case, LP will be the potential threat on users' location information, because it can gather a large amount of users' location data. Of course, conventional location sources such as GPS or mobile phone operators are also supposed to be trusted in most cases, but with the progress of technology, more and more web-based third part LPs have emerged. We have reason to doubt their reliability.

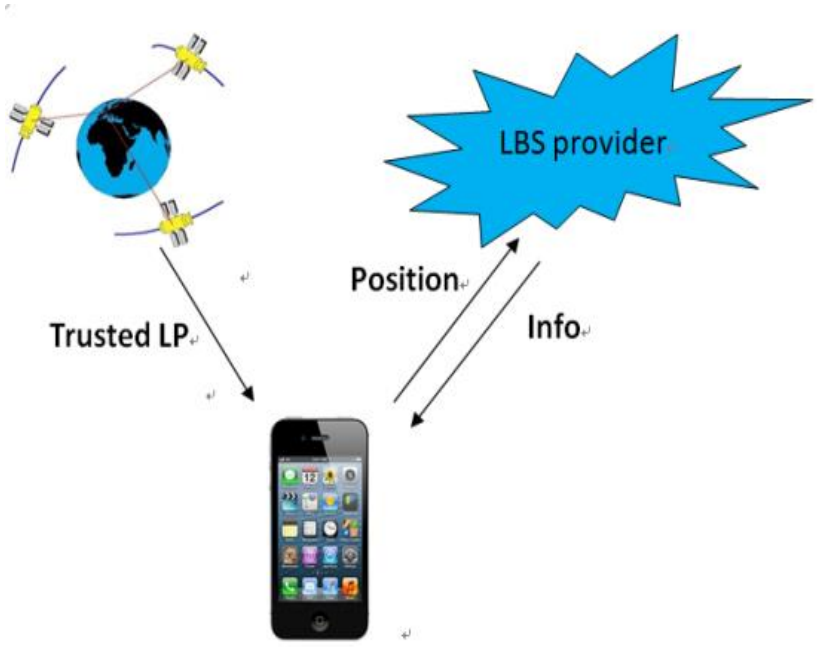

Figure 1. Conventional LBS Application Model

The protecting goals in LBS include three parts: user's *identity* (ID), spatial information and time-related information. Imagine that you are using an attracting navigation system which is based on your own locations; if you don't want the system know your name, you may choose to be anonymous. But as described in [8], the user ID can still be leaked out if his/her locations are revealed. An attacker can infer that some two places are home and work place if someone visits these two places regularly every day. If the attacker has some other background knowledge, the user's identity will most probably be figured out easily. Therefore,

users' location information must be protected and used in a safe way. In some cases, only when locations combine with time will it make sense. Imagine when you are driving on the highway, you just want to share your locations with your friends, but not your driving speed. In this case, the user's time information is equally important. One possible method is to delay the user's location update.

We can divide the LBS into two classes: one is conventional satellite-based LBS and another is web-based LBS [9]. The first one depends on positioning systems like American GPS, Chinese Beidou, Russian GLONASS and EU's Galileo. Generally, all these four satellite-based location providers are supposed to be trusted. On the other hand, the safety cannot be guaranteed in the web-based LBS.

3. How to Protect Web-based LBS Privacy

Not only LBS providers but also untrusted *Third Location Provider* (TLP) can threaten users' privacy information, especially when TLPs emerge so fast nowadays. Figure 2 depicts the application model in this case. The main IT leaders such as Google, Microsoft, Apple and Skyhook Wireless have all started their own location services, which are widely spread in all kinds of smart phones and network applications. Google Latitude, Foursquare and Skyhook Wireless are typical TLPs which are called the web-based LBS [9]. Users can access services by a special browser which has positioning function in the web-based LBS. That is, users' locations are not given by conventional location sources such as GSP, but calculated and sent to users by TLPs, and such a data process gives us reasons to doubt their reliability. The positioning principle of TLP is: the user's smart phone senses the locations of nearest communication equipment such as Wi-Fi Access Point (AP) or GSM/3G/4G base stations which are fixed and maintained in TLP's databases, and sends them to TLP. Then TLP can calculate users' location by such information. Wide positioning range is the advantage of this technology; whenever you are indoor or outdoor, in rural or urban, it works well, and its accuracy level limits to tens of meters, which can satisfy most LBS applications.

Figure 2. LBS Application Model Containing Untrusted TLP

Based on what we have described, you may think why the positioning function is not implemented within smart phones directly, that is, let users themselves be the location providers, so that we needn't to worry about threat from LPs. In fact, this is the solution called "Place Lab" [10] which was implemented by Intel. It's the precursor of business mix positioning system. The defect is that implementing all these LBS functions within a phone is costly. Furthermore, compared with TLP, the service quality is even worse. It's uneconomical and difficult to popularize.

The main idea to protect user privacy in the web-based LBS is to reduce or prevent the harm from TLP. Techniques to protect privacy in LBS include two parts: computational-based

and policy-based. Computational-based technology means to use some algorithms to deal with users' raw data to get processed data which are used in LBS safely. Policy-based technology relies on law or policy (standard protocol) which is made by third regulatory authority and must be forced to comply with. We suppose TLPs are not trusted, and they can infer users' locations by context such as Wi-Fi AP sent by users. If we can avoid communicating with TLP, the problem is solved, just as Intel's "Place Lab" has done. But, this idea has its own limitation and cannot be used broadly. What we can do is to minimize the communication times and information traffic with TLP so as to protect users' privacy to the greatest extent.

The amount of information user send to TLP is more than what is really needed to positioning [11]. Every time when users make a request to TLP, they will send the same context such as Wi-Fi AP, thus TLP can infer lots of related information about the users. If TLP finds a place where the user regularly stays at night, it can infer that this place is most likely to be the user's home. If the user's device can recognize and remember the visited places automatically, this problem can be partly solved. One possible solution is to buffer these sensitive places in mobile devices. When the user visits these places again in web-based LBS, the user's device can refer to local databases; if requested information is available, no more context should be sent to TLP. Damiani, *et al.* proposed a method to recognize the visited sensitive places, and combined every sensitive place with a radio fingerprint which is notated by Wi-Fi AP [12]. Every time a request is made, the user's device can recognize whether it is located in sensitive places by matching the fingerprint with Wi-Fi AP. Thus only when the user first visits some place will his/her context be sent to TLP. The second time and so on, the user can get needed information from local databases directly, so that the user can communicate with TLP as few as possible, and privacy threat can be reduced.

It's important to note that reducing the times communicating with TLP cannot avoid leakage of the user's location completely, because when the user visits some sensitive place the first time and makes a request to the web-based LBS, his/her context still need to be sent to TLP.

4. From the View of Policy and Law

A privacy policy specifies the privacy practices of an organization, mainly what kind of personal information is collected, for what purpose, and how the information will be used. Policy-based technology relies mainly on standard protocols which are made by third standard associations such as IETF Geopriv [13], and these protocol must comply with some law such as European standard and must be enforced to protect users' privacy. The simplest protocol is described by natural language and enforced manually. Apparently, this way is too costly and is easily neglected by users. The early research on standard protocol such as the famous P3P [14] focused on the compulsory automation. P3P allows websites to release their own standards by XML files. Although P3P is supported by most main browsers, it's not used very often because of its bad usability. IETF Geopriv regulates the way users' locations transfer on the Internet, and regulations to comply with are specified by users themselves, not some other organizations.

From the viewpoint of law, privacy protection standards state that users' privacy information can only be used under some conditions. In summary, there are three aspects: transparency, legitimate purpose and proportionality. Transparency means users have the right to be informed and access all related data when their privacy data is being dealt with. Legitimate means users' privacy data can only be used for lawful purposes which are clearly declared. Proportionality means that only those data which are needed to complete a user's request can be dealt, and not more than what is needed. IETF Geopriv allows users to specify when their locations can be dealt with and to whom they want to share their locations, such specification is implemented and enforced by machine language and privacy protection rules. Machine language means it's implemented by a kind of computer language, not natural language; and enforcement means these protection rules are enforced and supervised by a system which is on behalf of users.

From the above, policy-based mechanism is flexible and support personality in LBS privacy protection, the users can specify the protecting way and level themselves. But it's important to note that this mechanism can only provide a deterrent for privacy abuse. As for those who have decided to ignore the law, it will not work. That's why only policy-based mechanism is not enough to protect users' privacy and computational-based mechanism is essential. What is more, the precondition of policy-based mechanism is that LP is trusted, if this is not the case, such a mechanism will also fail to protect users' privacy. Therefore, computational-based technology is of the same importance.

5. From the View of Techniques

Privacy protection techniques use some algorithms to deal with users' privacy raw data and communicate with TSP using processed data. This is the most studied field in LBS at present, and lots of progresses have been made, such as mentioned in [4-7]. But all these techniques are designed just for particular scenario and conditions, and their classification criteria are different. [1] proposed three protection architectures, which are used as classification criterion; Three aspects: identity, location and language are described in [15]. In this paper, we summarize existing research achievements in principle and discuss from three aspects: fake data method, k-anonymity, obfuscating method. They have some intersections, for example, k-anonymity contains both fake data and obfuscating ideology.

5.1. Fake Data Methods

Fake data methods protect users' privacy by sending fake data instead of real data to service providers, such as using pseudonym to protect the user's ID. By sending false locations which are called dummies, users' real location can be protected [16]. This method can be easily implemented, because users themselves can make dummies. The degree of privacy protection depends on the distance between false and real locations. Shankar, et al [17] proposed a new method to make dummies, and they supposed that users maintain a historical database by which dummies were made.

5.2. K- anonymity

K-anonymity [18] was proposed by Latanya Sweeney from Carnegie Mellon university, firstly used in privacy protection about data publication in relational database systems. K-anonymity requires that one piece of data cannot be distinguished from at least other k-1 pieces of data. Marco Gruteser, *et al.* [19] firstly applied this idea to location privacy protection and proposed location k-anonymity. That is, when one user's location cannot be distinguished from other k-1 users', it meets k-anonymity. Now location k-anonymity has been widely extended to protect users' privacy, and the most famous are strong k-anonymity [20], l-diversity [21], t-closeness [22], p-sensitivity [23], and historical k-anonymity [24].

K-anonymity can not only protect users' location but the query privacy. Sometimes when making a query to a TSP, users are willing to protect not only their locations but the queries they made. When a user requests his/her nearest tumor hospital, he/she must be reluctant to let out this information. Some researchers call this kind of problem the query privacy and discuss it separately. Snapshot single query privacy can be solved by k-anonymity, but when a user makes continuous queries in a period of time, conventional k-anonymity is powerless. For example, there are six users {A, B, C, D, E, F} in a system, an attacker finds a continuous query but he/she doesn't know who makes it and what the query is. At three different time t_i, t_{i+1}, t_{i+2} ,user A gets three anonymous sets {A,B,D}, {A,B,F}, {A,C,E}. If the attacker intersects them, he/she will know user A makes the query and what A has requested. To solve this problem, [25] firstly proposed a method, which used the initially formed anonymous set as the final result in the period of validity, but this method could cause location leakage and poor service quality. The solution proposed in [26] supposed that users' locations are not uniformly distributed in the anonymous area, and the concept of entropy which is used in information theory was brought in to define the degree of privacy protection. But entropy

doesn't take the difference of locations into consideration, thus can lead to the condition that all users are at one point, and users' location privacy is leaked out.

5.3. Obfuscation Method

Obfuscating method protects users' definite location by reducing the location precision, mainly includes spatial obfuscation and spatio-temporal obfuscation. Spatial obfuscation extends users' definite location to a larger area which contains users' location, such as a circle. Spatio-temporal obfuscation adds time obfuscating on the basis of spatial obfuscation.

5.3.1. Spatial Obfuscation: One example is using a circle area instead of a point to represent users' location. Since the user can be randomly located in any point in the area, attacker only knows the user is in this area but the accurate location is obscure. This method is proposed by Ardagna [27]. Cheng [28] proposed another obfuscation method. In the paper, Cheng used an obfuscation Graph instead of geometrical shapes to obfuscate users' locations. The advantage of spatial obfuscation is simple implementation; users themselves can specify obfuscation area. But if the users' definite location is replaced with an area, the quality of service will decrease. Therefore, how to make a balance between privacy protection and quality of service is worth studying.

5.3.2. Spatio-temporal Obfuscation: Spatio-temporal obfuscation reduces the precision of not only location but the time-related information so as to satisfy the predefined k-anonymity standard. Gruteser [19] applied this method to protect users' trajectory. Similar idea was proposed and applied in spatio-temporal cloaking by Ghinita [29], the author even took background map knowledge into consideration. Other similar methods are trajectory clustering [30] and so on.

There are many other privacy protection techniques proposed, but this paper only chooses some typical methods to compare and discuss, and all techniques mentioned in this paper are summarized in Table 1.

Table 1. Commonly used Privacy Techniques in LBS

Goals and techniques / methods	Goals			Techniques			
	ID	Location	Time	Fake data	k-anony	Spatio Obf.	Spatio-temporal Obf.
Historical k-anony [24]	√	√	×	×	√	×	×
l-diversity [21]	√	√	×	×	√	×	×
p-sensitivity [23]	√	√	×	×	√	×	×
k-anonymity [20]	√	√	×	×	√	×	×
t-closeness [22]	√	√	×	×	√	×	×
Spatio-temp [27]	×	√	√	×	×	√	√
Traj. Cluster [28]	×	√	√	×	×	√	√
Geo-Obf.[25]	×	√	×	×	×	√	×
Obf.-Graph [26]	×	√	×	×	×	√	×
Dummy [16][17]	×	√	×	√	×	×	×

(" √ " means supported, " × " means not supported)

6. Conclusion

With the development of powerful smart phones, LBS will become more and more popular. When enjoying these convenient services, users need to provide their privacy information such as location which is likely abused. Many kinds of existing privacy preserving technologies are reviewed in this paper. We first introduced two classes of LBS: satellite-based LBS and web-based LBS which is rife nowadays, and pointed out the privacy threats that users encounter and possible solutions. Then, we continued to discuss security mechanisms from two aspects: policy-based and computational-based techniques. Policy-based mechanism can support more flexible protection, but the precondition is that LP is trusted. Otherwise, it will not work correctly. By this mean, computational-based technology is essential, which can guarantee the validity of privacy protection. They should become mutually complementary. As for the users, they should make a tradeoff between the benefit and privacy threat brought by LBS.

Privacy protection is an important research challenge. Although progresses have been made in recent years, lots of problems are still unresolved and more solutions are needed to put forward. Following are some possible research challenges that we can foresee in near future.

(1) Evaluation criteria about privacy protection

Progresses of technology in any fields largely rely on the advance of evaluation criteria in that field, so as to privacy protection in LBS, that is, how to evaluate the user privacy and quality of service. As for some specific techniques, how to quantify properly is of much importance. For example, since spatial obfuscation mentioned above uses an area to replace users' location, it's obvious that bigger obfuscation area means higher degree protection but lower QoS. How to figure out a proper area is worth studying. Actually, finding optimal solution in obfuscation methods is an NP problem which is challenging.

(2) New universal framework of privacy protection

Internet is global, so that users' data is most likely to flow to different regions or countries, but exiting security framework such as European standard can work only in limited regions. Therefore, a new universal framework must be proposed by some international organization, or at least let users have the right to choose when meeting with several different frameworks.

(3) New attack model in context-aware condition

Most recent privacy protection techniques are designed just for specific context, but when attackers gained new knowledge of users, lots of new problems will appear. How to find new attack model and propose corresponding solutions is also a research challenge in the future.

Acknowledgements

This work is partly supported by the Research Project No. Y201329915, the Department of Education of Zhejiang Province, China.

References

[1] M. F. Mokbel, "Privacy in Location-based Services: State-of-the-art and Research Directions", 2007 International Conference on Mobile Data Management, (**2007**), pp. 228-228.
[2] D. Pedreschi, F. Bonchi, F. Turini, V. S. Verykios, M. Atzori, B. Malin, B. Moelans and Y. Saygin, "Privacy protection: regulations and technologies, opportunities and threats", Mobility, Data Mining and Privacy: Springer, (**2008**), pp.101-119.
[3] M. Colbert, "A diary study of rendezvousing: implications for position-aware computing and communications for the general public", Proceedings of the 2001 International ACM SIGGROUP Conference on Supporting Group Work, (**2001**), pp. 15-23.
[4] C.-Y. Chow and M. F. Mokbel, "Trajectory privacy in location-based services and data publication", ACM SIGKDD Explorations Newsletter, vol. 13, no. 1, (**2011**), pp.19-29.
[5] J. Krumm, "A survey of computational location privacy", Personal and Ubiquitous Computing, vol. 13, no. 6, (**2009**), pp. 391-399.
[6] A. Solanas, J. Domingo-Ferrer and A. Martínez-Ballesté, "Location privacy in location-based services: Beyond TTP-based schemes", Proceedings of the 1st International Workshop on Privacy in Location-Based Applications (PILBA), (**2008**), pp. 12-23.
[7] T. Wang and L. Liu, "From data privacy to location privacy", Machine Learning in Cyber Trust: Springer, (**2009**), pp. 217-246.
[8] P. Golle and K. Partridge, "On the anonymity of home/work location pairs", Pervasive Computing: Springer, (**2009**), pp. 390-397.
[9] M. L. Damiani, "Third party geolocation services in LBS: privacy requirements and research issues", Transactions on data privacy, vol. 4, no. 2, (**2011**), pp.55-72.
[10] A. Lamarca, Y. Chawathe, S. Consolvo, J. Hightower, I. Smith , J. Scott, T. Sohn, J. Howard, J. Hughes and F. Potter, "Place lab: Device positioning using radio beacons in the wild", Pervasive computing: Springer, (**2005**), pp. 116-133.
[11] M. L. Damianim and M. Galbiati, "Handling user-defined private contexts for location privacy in LBS", Proceedings of the 20th International Conference on Advances in Geographic Information Systems, (**2012**), pp. 574-577.
[12] M. L. Damiani and C. Cuijpers, "Privacy challenges in third-party location services", Mobile Data Management (MDM), 2013 IEEE 14th International Conference on, (**2013**), pp. 63-66.
[13] R. Barnes, M. Lepinski, A. Cooper, J. Morris, H. Tschofenig and H. Schulzrinne," An Architecture for Location and Location Privacy in Internet Applications, BCP 160, RFC 6280, (**2011**) July.
[14] L. F. Cranor, "P3P," Making Privacy Policies More Useful, IEEE Security & Privacy, New York, (**2003**), pp. 50-55.
[15] M. L. Damiani, "Privacy enhancing techniques for the protection of mobility patterns in LBS: research issues and trends", European Data Protection: Coming of Age: Springer, (**2013**), pp. 223-239.
[16] H. Kido, Y. Yanagisawa and T. Satoh, "An anonymous communication technique using dummies for location-based services", Pervasive Services, 2005, ICPS'05, Proceedings, International Conference on, (**2005**), pp. 88-97.
[17] P. Shankar, V. Ganapathy and L. Iftode, "Privately querying location-based services with sybilquery", Proceedings of the 11th international conference on Ubiquitous computing, (**2009**), pp. 31-40.
[18] L. Sweeney, "k-anonymity: A model for protecting privacy [J]", International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, vol. 10, no. 05, (**2002**), pp. 557-570.
[19] M. Gruteser and D. Grunwald, "Anonymous usage of location-based services through spatial and temporal cloaking", Proceedings of the 1st international conference on Mobile systems, applications and services, (**2003**), pp. 31-42.
[20] C. Zhang and Y. Huang, "Cloaking locations for anonymous location based services: a hybrid approach [J]", GeoInformatica, (**2009**), vol. 13, no. 2, pp.159-182.
[21] B. Bamba, L. Liu, P. Pesti and T. Wang, "Supporting anonymous location queries in mobile environments with privacygrid", Proceedings of the 17th international conference on World Wide Web, (**2008**), pp. 237-246.
[22] N. Li, T. Li and S. Venkatasubramanian, "t-Closeness: Privacy Beyond k-Anonymity and l-Diversity", ICDE, (**2007**), pp. 106-115.

[23] A. Solanas, F. Sebé and J. Domingo-Ferrer, "Micro-aggregation-based heuristics for p-sensitive k-anonymity: one step beyond", Proceedings of the 2008 international workshop on Privacy and anonymity in information society, (**2008**), pp. 61-69.
[24] S. Mascetti, C. Bettini, X. S. Wang, D. Freni and S. Jajodia, "Providenthider: An algorithm to preserve historical k-anonymity in lbs", Mobile Data Management: Systems, Services and Middleware, 2009. MDM'09. Tenth International Conference on, (**2009**), pp. 172-181.
[25] C.-Y. Chow and M. F. Mokbel, "Enabling private continuous queries for revealed user locations", Advances in Spatial and Temporal Databases: Springer, **(2007**), pp. 258-275.
[26] T. Xu and Y. Cai, "Location anonymity in continuous location-based services", Proceedings of the 15th annual ACM international symposium on Advances in geographic information systems, (**2007**), pp. 39.
[27] C. A. Ardagna, M. Cremonini, E. Damiani, S. D. C. Di Vimercati and P. Samarati, "Location privacy protection through obfuscation-based techniques", Data and Applications Security XXI: Springer, (**2007**), pp. 47-60.
[28] R. Cheng, Y. Zhang, E. Bertino and S. Prabhakar, "Preserving user location privacy in mobile data management infrastructures", Privacy Enhancing Technologies, (**2006**), pp. 393-412.
[29] G. Ghinita, M. L. Damiani, C. Silvestri and E. Bertino, "Preventing velocity-based linkage attacks in location-aware applications", Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, (**2009**), pp. 246-255.
[30] J.-G. Lee, J. Han and K.-Y. Whang, "Trajectory clustering: a partition-and-group framework", Proceedings of the 2007 ACM SIGMOD international conference on Management of data, (**2007**), pp. 593-604.

Authors

Julong Pan, he received his Ph.D. from Zhejiang University, China in 2011, and now he is a professor in College of Information Engineering, China Jiliang University, Hangzhou, China. His main research interests are wireless sensor networks security, mobile computing and wireless access technology.

Zhengwei Zuo, he is a master degree candidate in China Jiliang University. His research interest is wireless sensor networks security issues.