Electronic Commerce Research and Applications 18 (2016) 58–70

Contents lists available at ScienceDirect

Electronic Commerce Research and Applications

journal homepage: www.elsevier.com/locate/ecra

User privacy protection for a mobile commerce alliance

Chunhui Piao [a,b,*], Xiaoyan Li [a], Xiao Pan [b], Changyou Zhang [c]

[a] *School of Information Science and Technology, Shijiazhuang Tiedao University, Shijiazhuang, Hebei, China*
[b] *School of Economics and Management, Shijiazhuang Tiedao University, Shijiazhuang, Hebei, China*
[c] *Lab. of Parallel Software and Comp. Sci., Inst. of Software, Chinese Acad. of Sci., Beijing, China*

ARTICLE INFO

Article history:
Received 15 March 2015
Received in revised form 26 March 2016
Accepted 26 March 2016
Available online 28 March 2016

Keywords:
Anonymity model
Location-based services
Mobile commerce
Personalized privacy profile
Privacy-preserving algorithm
Privacy requirements
Service framework
User privacy protection

ABSTRACT

The risk of privacy disclosure in mobile commerce has received increasing attention worldwide. Although many papers related to information privacy and privacy-preserving technologies exist, few are based on a particular mobile commerce model to study the anonymity models and privacy-preserving algorithms. A privacy-preserving service framework for the mobile commerce alliance providing location-based services is established. According to the defined personalized privacy profile of the mobile user, a (K, L, P)-anonymity model is formally described. Based on the model, a new privacy-preserving algorithm for *exchanging and merging processes for generating anonymity sets* (EMAGAS) is proposed, which features the construction of minimal initial K-anonymity sets, an exchanging process and a merging process. The processes of exchanging and merging are formally described. EMAGAS can be used to protect the location, identifier and other sensitive information of the mobile user on a road network. The availability of EMAGAS is illustrated by an example. Finally, based on a real road network and generated privacy profiles of mobile users, the feasibility and advantages of EMAGAS are experimentally validated.

1. Introduction

Mobile commerce refers to the e-commerce activities conducted using mobile handheld devices such as cellular telephones and personal digital assistants (PDAs) through mobile Internet. Compared with conventional electronic commerce, m-commerce has some new features including mobility, instantaneity, personalization and convenience. *Location-based services* (LBS), a general class of information services accessible to mobile users, uses information about the geographical locations of mobile devices based on mobile communication technologies such as global positioning system (GPS), wireless local area networks (WLAN) and cellular networks. In recent years, with the pervasive application of new information and communication technologies, m-commerce has been developing rapidly. One of the most widely used location-based m-commerce applications is mobile advertising (Tähtinen and Salo, 2004). New types of m-commerce applications providing LBS have become popular. Meanwhile, various types of m-commerce alliances beneficial to share resources are emerging.

To use LBS, mobile users usually are required to send their query requests and accurate locations to the service providers. The service providers may collect, process, and store the users' locations on an unprecedented scale, and location privacy-related issues have naturally attracted increasing attention (Terrovitis, 2011). According to a survey conducted by Microsoft in 2011, the main reason why people were unwilling to adopt LBS was the concern of personal privacy. Many events related to the disclosure of mobile users' privacy information have been reported by public media. In practice, the untrustworthy service providers may collect mobile users' privacy information from the service requests, then disclose or misuse the privacy information. In the research on privacy protection related to m-commerce providing LBS, three types of information must be protected: location, identifier and sensitive information (Wu et al., 2014). Location information can reveal sensitive information about the mobile users, such as health problems, commercial practices.

The effective protection of sensitive information should ensure that the adversary has low confidence to link sensitive information with a specific user, such as the user may be ill, some type of sensitive services may be needed by the user. Anonymous communication, data conversion, k-anonymity and cryptography-based techniques are the commonly used privacy preserving technologies in the research on protecting privacy information.

Although many papers related to information privacy and privacy-preserving technologies exist, few are based on a particular m-commerce model to study the anonymity models and

[*] Corresponding author at: School of Information Science and Technology, Shijiazhuang Tiedao University, Shijiazhuang, Hebei, China.

E-mail address: pchls2011@126.com (C. Piao).

http://dx.doi.org/10.1016/j.elerap.2016.03.005
1567-4223/© 2016 Published by Elsevier B.V.

privacy-preserving algorithms. This study attempts to answer three research questions: What is the applicable privacy-preserving service framework for a specific m-commerce alliance? How can the personalized privacy requirements of the mobile user in the context of m-commerce be formally defined? Based on the defined anonymity model, can a new privacy-preserving algorithm be established?

In Section 2, after reviewing the concepts and works related to information privacy and privacy concerns in m-commerce, the commonly used privacy preserving technologies in mobile environments are discussed. A privacy-preserving service framework for the m-commerce alliance providing LBS is established in Section 3. According to the defined personalized privacy profile of the mobile user, a (K, L, P)-anonymity model is described in Section 4. Based on the anonymity model, a new privacy-preserving algorithm for *exchanging and merging processes for generating anonymity sets* (*EMAGAS*) is proposed. The processes of exchanging users and merging users are discussed in detail and described formally. In Section 5, the availability of *EMAGAS* is illustrated by an example. In Section 6, based on a real road network and generated privacy profiles of its mobile users, the feasibility and advantages of *EMAGAS* are experimentally validated. Conclusions are presented last.

2. Related work

2.1. Personal information privacy

The concept of privacy is widely relevant in many fields. The word *privacy* has different meanings in different disciplines such as Psychology, Law, Sociology, Economics, Management and Informatics. Warren and Brandeis (1890) published the article “The Right to Privacy” in the 1890 *Harvard Law Review*, which defined the privacy of the individual as a right to be let alone. It is widely regarded as the first publication in the United States to advocate aright to privacy. As one of the basic human rights, privacy can be divided into three dimensions: decision privacy, information privacy and residence privacy. Among these, information privacy represents the central dimension (Shen, 2013).

There are many different definitions of personal information privacy. For example, Mason (1986) defined information privacy as the right to control, collect and use personal information. Culnan (1995) considered it as the person’s ability to control her own information to be accessed by others. And Belanger and Crossler (2011) defined information privacy as the desire of individuals to control or have some influence over data about themselves.

Currently, many developed countries and independent territories have adopted comprehensive information privacy or data protection laws to prohibit the disclosure or misuse of information held on private individuals (Greenleaf, 2014). These laws are based on Fair Information Practice, first developed in the United States in the 1970s by the Department for Health, Education and Welfare. Early in 1990, Straub and Collins (1990) reviewed laws relevant to information privacy and highlighted the implications for information managers. However, the developing countries lag behind the developed countries in the laws, regulations and directives related to the protection of information systems. For example, China has not enacted its own data protection law regulating personal data (Liu et al., 2015).

The pervasive application of information and communication technologies have raised concerns about privacy information. Information privacy concerns occur when consumers find it difficult to control their own private information from misuse or unauthorized distribution. Culnan and Armstrong (1999) argued that consumers have two types of privacy concerns. First, they are concerned over unauthorized access to personal data because of security breaches or the lack of internal controls. Second, consumers are concerned about the risk of secondary use—the reuse of their personal data for unrelated purposes without their consent, including data sharing with third parties that were not part of the transaction. Privacy concerns exist wherever uniquely identifiable data relating to an individual or a group of individuals is collected and stored, in digital form or otherwise.

The Internet has introduced new concerns about privacy in the age where computers can permanently store records of everything. Internet privacy is the ability to determine what information one reveals or withholds about oneself over the Internet, who has access to such information, and for what purposes one’s information may or may not be used. For example, Web users may be concerned that the websites they visit collect, store, and possibly share personally identifiable information about them—such as their IP addresses, online social networks. Similarly, Internet email users generally consider their emails to be private and hence would be concerned if their emails were being accessed, read, stored or forwarded by third parties without their consent. In e-commerce, although the interactivity of the Internet provides more business opportunities for the vendors, Internet privacy issues are more serious for consumers (Jiang et al., 2010; Lee et al., 2010).

There are many studies on the issue of information privacy concerns. Some of them are based on the level of analysis: individual, group, organizational and societal (Smith et al., 2011). Li (2014) proposed a multi-level model of individual information privacy beliefs. This model consists of three levels of privacy beliefs: disposition to privacy, representing a person’s fundamental beliefs; online privacy concern, representing a person’s overall perception of privacy risks in the online environment; and website privacy concern, representing a person’s perception of privacy risks on a particular website. Kim and Lee (2009) studied the impact of consumers’ information privacy concerns on firms’ collection and use of consumer information for Web-based personalization. The result shows that a firm of inferior ability in customer information utilization is more affected by privacy concerns than a firm of superior ability in terms of choosing to collect and use consumer information for personalization.

Aiming at online consumers’ concerns regarding information privacy, a major problem hampering the growth of e-commerce, Malhotra et al. (2004) proposed a theoretical framework on the dimensionality of Internet users’ information privacy concerns (IUIPC). Based on the understanding that legitimate concerns regarding privacy and trust are important issues to both individuals and organizations in e-commerce, Liu et al. (2004) proposed and tested a theoretical model that considers an individual’s perceptions of privacy and how they relate to his or her behavioral intention to conduct an online transaction. In the emerging social commerce, shoppers’ information disclosure intention may be driven by the fairness of information exchange, privacy benefits and privacy apathy (Sharma and Crossler, 2014).

2.2. Privacy concerns in m-commerce

Privacy concerns in m-commerce applications mainly involve the mobile user’s basic personal information, mobile device information, mobile payment information, location information, interest information, transaction information, social relationship information and other privacy-related information (Piao et al., 2013). The information from the local sensors such as the GPS receiver, camera, microphone, and accelerometer that can be used by third-party applications is viewed as privacy-sensitive data on the smartphone user (Enck et al., 2014). Data privacy and security are major concerns for m-commerce (Hamad et al., 2009). The concern for the disclosure of location-related privacy information is a key factor affecting the adoption of m-commerce. Consumers' demographic differences had varying degrees of impact on concerns for information privacy in m-commerce (Zhang et al., 2013). Based on the consideration that information privacy is one of the major factors influencing the use of m-commerce, a research model to identify factors influencing behavioral intention to provide private information in m-commerce is established (Kim, 2015).

Location privacy is a special type of information privacy (Bereford and Stajano, 2003). It can be defined as the ability to prevent unauthorized parties from learning one's current or past location (Hong and Landay, 2004). In mobile contexts, user identity privacy refers to preventing queries of the user's identity identifier from disclosure when the attacker obtains and analyzes the user's other identity-related information. For example, location information can be used as a pseudo-identity in the query. In mobile applications providing LBS functionality, sensitive data may be the information of general concern, such as health information (Guo et al., 2016) or financial data that can be transmitted as part of a service request. It may also be spatio-temporal information on mobile users collected by a service provider, for example, information on the locations of users at a specific time, movement patterns, or personal points of interest (frequent visits to specific shops, clubs, or institutions) (Bettini et al., 2005).

Users with privacy concerns worry about private information shared in m-commerce being properly collected, stored and used, and believe that disclosure of their location information may incur serious risks to them (Zhou, 2011). However, to use LBS, mobile users must send their location information over the Internet to the service providers, which may lead to two types of privacy issues. One is that the attacker may gain access to the original location data by positioning location transmission equipment or eavesdropping data transmission channels and obtain location-related personal privacy information using the data collected from multiple sources. The other is that the service provider may obtain the users' personal sensitive information by analyzing query request information and accurate location data received, and even resell or release the privacy information for profit. Thus, service providers should take measures to reduce users' privacy concerns to support adoption of m-commerce.

2.3. Privacy-preserving technologies in the mobile environment

2.3.1. Privacy preserving technologies based on road networks

Research on location privacy-preserving technologies can be classified into free-space-based (Kido et al., 2005; Gruteser and Grunwald, 2003; Solanus et al., 2008) and constrained-space-based (Xue et al., 2011; Wang and Liu, 2009). For free-space-based technologies, it is assumed that the users can move freely in the Euclidean space without constraints. However, in real mobile environments, the mobile users usually move in constrained networks, such as road networks and railway networks, regardless of what types of LBS they request. Applying the location anonymity methods designed for Euclidean space to road network environments may result in privacy leakage (Chow et al., 2011). The basic ideas of location anonymity methods in both spaces can be classified into dummies (Kido et al., 2005), spatio-temporal cloaking (Gruteser and Grunwald, 2003; Mokbel et al., 2006; Bamba and Liu, 2008; Chow and Mokbel, 2007; Pan and Meng, 2013) and encryption (Robello-Monedero et al., 2008; Solanus and Martinez-Balleste, 2007, 2008; Ghinita et al., 2008; Khoshgozaran and Shahabi, 2007). Encryption-based anonymization methods have the highest security, but their efficiency is low; dummy-related methods are simple and efficient, but their degree of privacy preservation is low; and spatio-temporal cloaking methods can be used to balance privacy and efficiency.

In the last few years, with the aim of preserving the location privacy of mobile users on road networks, various solutions have been proposed. Most existing works focus on protecting the users' location and identity identifier information based on the location *K*-anonymity model. The essential feature of the *k*-anonymity model is *K*-sharing; namely, the location or identifier of a specific user cannot be distinguished from at least *k*-1 users who also appear on the road network.

Mouratidis and Yiu (2010) first applied the location *K*-anonymity model to a road network. The road network was modeled as an undirected graph. Assume that users *A*, *B* and *C* in Fig. 1 constitute a cloaking set, which follows the location 3-anonymity model. The anonymity server sends the cloaking set to the platform providing LBS. Thus, it is difficult for the attacker to distinguish the three users in the set. Wang and Liu (2009) presented a study on the problem of protecting location privacy of the mobile users under a network-constrained mobility model and proposed a general location anonymity model, XStar, for privacy-aware mobile services over road networks. Chow et al. (2011) proposed a location anonymization algorithm specifically for road network environments, in which a user location was cloaked into a set of connected road segments of a minimum total length *L*, including at least *K* users and considering the proximity of the road segments in the geographical contexts.

Xue et al. (2011) discussed the location privacy-preserving issue in road networks including one-way roads and proposed a location privacy preserving approach using cloaking cycle and forest. Li et al. (2008) presented a hierarchical index structure for location anonymization on road networks, which was used to obtain cloaked segment sets very similar to the user's privacy requirements. Huo et al. (2013) explored a *K*-anonymity trajectory privacy-preserving method. Lin et al. (2012) proposed a road network-bichromatic reverse nearest neighbor query algorithm, which processed the query request on the premise of protecting the location privacy of the user. Pan et al. (2012) presented a location cloaking algorithm, called ICliqueCloak, which considered the effect of continuous location updates in the process of location cloaking. All of the research works mentioned above aimed at protecting mobile users' locations and identity identifiers, without considering the sensitivity of information disclosure.

There has been less research on sensitive information protection in m-commerce. One reported work is TaintDroid, an extension to the Android mobile phone platform that tracks the flow of privacy-sensitive data through third-party applications (Enck et al., 2014). Pan et al. (2014) were the first researchers to study the protection technology of users' sensitive information on a road network. They proposed a (*K*, *L*, *P*)-anonymity model and a P^3RN cloaking algorithm to cope with sensitive homogeneity attacks that were query-semantic-aware over road networks. This approach supports personalized privacy requirements and ensures the location information security of the users. Starting from where the main steps of the approach may be changed, we will present a new privacy-preserving algorithm.

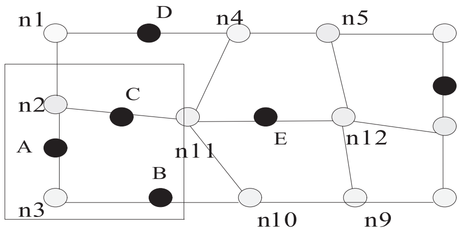

Fig. 1. Example road network RN_1.

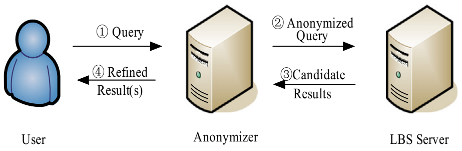

Fig. 2. Centralized privacy-preserving system architecture.

2.3.2. Centralized privacy-preserving system

The majority of location privacy preservation studies have been based on the premise that LBS providers cannot be completely trusted. In the communication scenario between the user and LBS provider, different techniques have been proposed to provide protection to the user against location disclosure by hiding the exact location of the user from untrustworthy entities, against identity disclosure by hiding the identity of the user from adversaries that may already know other identifying information, and against disclosure of sensitive information by preventing the adversaries from inferring that the user visited a certain place or adopted a sensitive service (Terrovitis, 2011).

In terms of the system architecture for privacy preservation in mobile applications, existing approaches can be classified into three categories: non-cooperative, centralized and decentralized. The most widely used architecture is the centralized system architecture as shown in Fig. 2, involving mobile users, the location anonymizer and LBS providers. The location anonymizer, a trusted third-party, is located between the mobile user and the service provider (Pan et al., 2012). The anonymizer performs the process of location anonymization and query result refinement with complete knowledge of the users' locations and query requests, such as the category of the query and the sensitivity of the query. Our work also applies the centralized architecture.

The general procedure for query processing based on the centralized architecture is as follows:

- *Query submission*. The mobile user sends an LBS query request containing the user's exact location and sensitive information to the anonymizer.
- *Anonymization*. Upon receiving a query, the anonymizer performs an anonymization algorithm, and each of the cloaking sets that is generated meets the requirements of the specified anonymous model. The anonymizer then forwards the anonymized query request to the LBS server.
- *Query execution*. When it receives the query request, the LBS server executes the query and returns the candidate results to the anonymizer.
- *Results refinement*. When receiving the candidate results, the refinement engine finds the best result or ranks the candidate results and sends the refinement result to the user.

3. Privacy-preserving service framework for m-commerce alliance

3.1. The privacy-preserving service framework of MCA

With the rapid development of m-commerce, various types of m-commerce alliances have emerged in recent years. In contrast to the *intentionally-developed business network* (IDBN) (Salo et al., 2008), which focuses on B2B marketing, the *mobile commerce alliance* (MCA) that we explore provides B2B2C services. The m-commerce alliance aims to provide trusted, reliable and value-added IT infrastructure services, promote resource sharing, and facilitate win–win relationships among the players involved. In this article, we view MCA as a trustworthy information service platform collaborating with many mobile information service providers, each of which owns its own specific vendor resources.

The privacy-preserving service framework of MCA is shown in Fig. 3. It has four categories of entities: the MCA platform, information service providers, vendors with physical stores, and mobile users (Piao et al., 2013). Each type of entity has its own rights and obligations. The MCA platform acts as a trusted third-party providing an anonymization service and comprehensive location-based information services, including vendor recommendation according to the mobile user' query request and current location. The information service providers protect the privacy of the vendors through the implementation of fair information practices and standards governing the collection and use of personal information. The mobile users have the right to specify their privacy requirements.

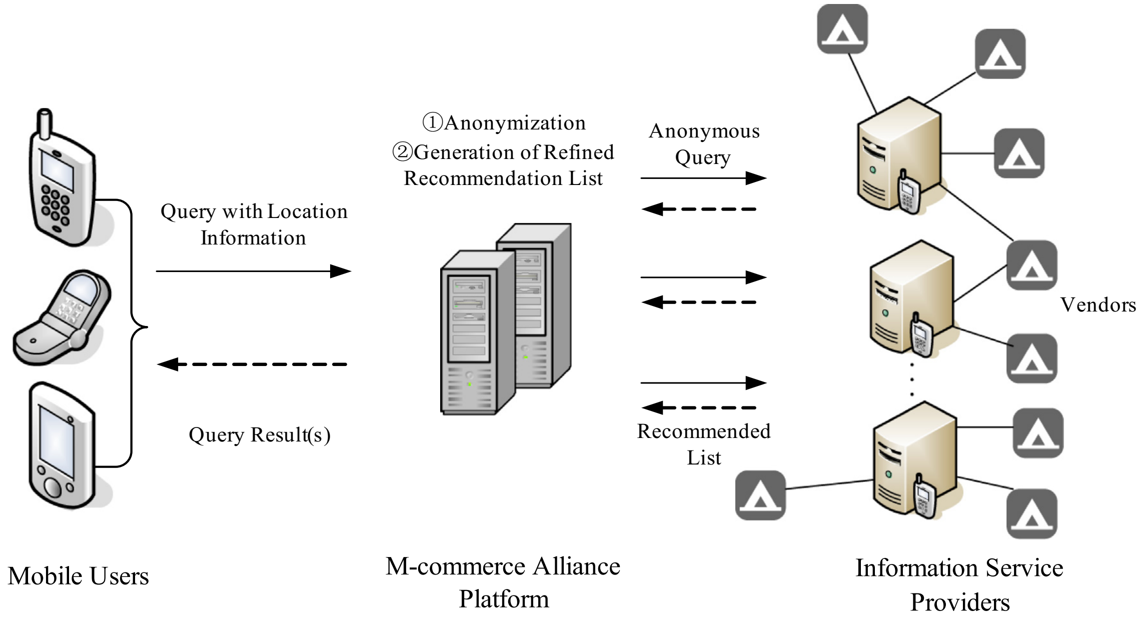
①Anonymization
②Generation of Refined Recommendation List
Query with Location Information
Query Result(s)
Anonymous Query
Recommended List
Vendors
Mobile Users
M-commerce Alliance Platform
Information Service Providers

Fig. 3. The privacy-preserving service framework of MCA.

3.2. Service model for vendor recommendation of MCA based on privacy-preservation

The service model is as follows.

- Mobile users register their basic information on the MCA platform and enable it to access their location information when needed.
- When a query request is received, the MCA platform performs an anonymization algorithm based on the location of the querying user to generate a cloaking set for the user that satisfies the requirements of a specified anonymous model.
- The anonymous query request and the cloaking set are sent to the information service providers of the MCA.
- Each of the information service providers executes the query using its particular algorithm and returns the query results generated to the MCA platform, such as the recommended vendors' names, home page links and other needed information.
- Based on the accurate location of the query user and the recommended lists of the vendors from multiple information service providers, the MCA platform ranks the vendors using a ranking algorithm and sends the optimal recommendation list back to the query user.
- If the user adopts the recommendation—visits the home page of a recommended vendor—it can be viewed as an effective recommendation, or feedback. To facilitate the improvement of vendor recommendation quality, the collected feedback can be used as one of the ranking indicators in the recommendation algorithm.

Under the model of vendor recommendation described above, the MCA platform receives accurate location information for the query user and generates the anonymous query, whereas each of the information service providers must perform its own recommendation algorithm and generate its vender recommendation list based on the cloaking set. We next turn to the anonymity model established for the platform and present a privacy-preserving algorithm based on exchanging and merging processes.

4. Personalized privacy-preserving algorithm on road network

4.1. Basic definitions

Information sensitivity refers to the control of access to information that might result in loss of security if disclosed to or shared with others. Different mobile users may assign different levels of sensitivity to the same type of information based on different privacy dispositions. The information privacy requirements of mobile users vary depending on their information sensitivity and personal disposition.

Referring to the work of Pan et al. (2014), all query requests submitted to the anonymizer can be classified into different sensitivity categories, and each category of requests are labeled with a value in the range of [0, 1] to represent the corresponding information sensitivity. The larger the value labeled, the more sensitive the category of requests. For example, a query Q_1, to find the nearest hospital, can be labeled with sensitivity level 0.75; another query Q_2, to find the nearest primary school, can be labeled with sensitivity level 0. Thus, Q_1 is more sensitive than Q_2.

We first formally define the personalized privacy profile for the mobile user.

Definition 1 (*Personalized privacy profile*). Given that the mobile user can specify a personalized privacy profile, it can be formally described as four-tuple (k, qsr, sd, p).

In this definition, the *anonymity requirement* (k) represents the minimum anonymity level that the mobile user can accept, which means that at least k mobile users are included in the anonymity set. The larger the value of k, the higher the anonymity level of the mobile user's identifier. The *query sensitivity requirement* (qsr) represents the mobile user's maximum tolerable query sensitivity. If the sensitivity of a query Q is larger than qsr, then Q is a sensitive query for the user; otherwise, Q is a non-sensitive query for the user. The larger the value of qsr, the higher the tolerability of the sensitive queries. The *segment diversity requirement* (sd) represents the minimum number of different road segments in the cloaked region. The larger the value of sd, the higher the level of location anonymity of the mobile user. And the *set sensitivity requirement* (p) implies the mobile user's maximum tolerance ratio for sensitive queries involved in the anonymity set. The larger the value of p, the stricter the set sensitivity.

For example, a mobile user u sets the privacy profile as (4, 0.25, 3, 0.5). This means that u requires that the anonymity set contains at least 4 different mobile users and at least 3 different road segments are included in the cloaked region. If a query's sensitivity is larger than 0.25, u regards this query as a sensitive query and requires that the ratio of sensitive queries involved in the anonymity set not be larger than 0.5.

Definition 2 (*Anonymity set*). This is composed of the mobile users in the cloaked region. For a user's anonymity set AS, $|AS|$ denotes the number of mobile users in AS.

For all $u \in AS$, u's anonymity requirement is $u.k$, its query sensitivity requirement is $u.qsr$, and its set sensitivity requirement is $u.p$. The sensitivity of the query sent by u is $u.qs$. Further, the maximum anonymity requirement of the users in AS is denoted as Max $u.k$ for $u \in AS$, the number of road segments in the cloaked region is denoted as $AS.Count\_S$, and the number of sensitive queries for u is $AS.Count\_SQu$.

Definition 3 (*(K, L, P)-anonymity model*). If the anonymity set AS satisfies three conditions, then we say that it the (K, L, P)-anonymity model. The conditions are: (1) location K-anonymity, $|AS| \geqslant K$, where $K = Max\ u.k$ for all $u \in AS$; (2) road segment diversity, $AS.Count\_S \geqslant L$, where $AS.Count\_S$ denotes the number of road segments in the cloaked region, $L = Max\ u.sd$ for all $u \in AS$; and (3) sensitive information P-anonymity, such that for all $u \in AS$, $u.p\ AS.Count\_SQu/|AS|$.

By applying the (K, L, P)-anonymity model, the privacy information of the users in mobile contexts can be comprehensively protected. The first condition implies that the user in AS has the K-sharing property, which can prevent disclosure of a query user's identifier and location. The second condition guarantees that the diversity of the published road segments and exact location of the query user are hidden from attackers. Finally, the third condition contributes to protecting sensitive information.

For example, let $AS_1 = \{u_1, u_2, u_3\}$ and $AS_2 = \{u_4, u_5, u_6, u_7\}$. The mobile users' personalized privacy profiles are shown in Table 1a, and the queries and labeled sensitivity values are shown in Table 1b. According to the (K, L, P)-anonymity model, the values of all variables are listed in Table 1c. The values of the unsafe user u_6 in AS_2 are highlighted in green. It can be shown that AS_1 satisfies the (K, L, P)-anonymity model, whereas AS_2 does not.

4.2. The anonymization algorithm, EMAGAS

To protect the privacy information of mobile users, the MCA platform must group all query requests sent by mobile users to the platform into different anonymity sets satisfying the

Table 1a
Personalized privacy profile.

Mobile user	*K*	*qsr*	*sd*	*p*
u_1	2	1	2	0.4
u_2	3	0.6	2	0.5
u_3	3	0.4	2	0.8
u_4	3	0.6	2	0.6
u_5	2	0.4	2	0.8
u_6	2	0.25	2	0.5
u_7	3	0.5	2	1

Table 1b
Query and query sensitivity.

Anonymity set	Mobile user	Query	Query sensitivity
AS_1	u_1	Q_1	0.25
	u_2	Q_2	0.5
	u_3	Q_3	1
AS_2	u_4	Q_4	0.5
	u_5	Q_5	1
	u_6	Q_6	0.5
	u_7	Q_7	0

(*K*, *L*, *P*)-anonymity model. We propose an algorithm for by exchanging and merging processes for generating anonymity sets (*EMAGAS*). It features of *EMAGAS* include: (1) constructing the minimum initial anonymity sets to satisfy *k*-anonymity; (2) for the initial anonymity sets that do not satisfy the (*K*, *L*, *P*)-anonymity model, exchanging the users in adjacent anonymity sets, and then merging adjacent anonymity sets; and (3) for the anonymity sets that do not satisfy the (*K*, *L*, *P*)-anonymity model after the exchanging and merging processes have finished, adding dummies to make them satisfy the (*K*, *L*, *P*)-anonymity model. The flowchart of the *EMAGAS* algorithm is shown in Fig. 4.

The steps in Fig. 4 are as follows:

(1) Traverse the road network using the *depth-first search* (DFS) algorithm. In the traversal, the road segments and the mobile users are numbered in sequence.
(2) Select a mobile user on a road segment as the first component of a new anonymity set who has not been included in any constructed anonymity set. Find and add adjacent users who have not been included in any constructed anonymity set to the new anonymity set in sequence until the number of users in the set is equal to the maximum anonymity requirement of the users and the condition of the road segment diversity is satisfied.
(3) Repeat Step 2 until every mobile user is included in one of the constructed anonymity sets.

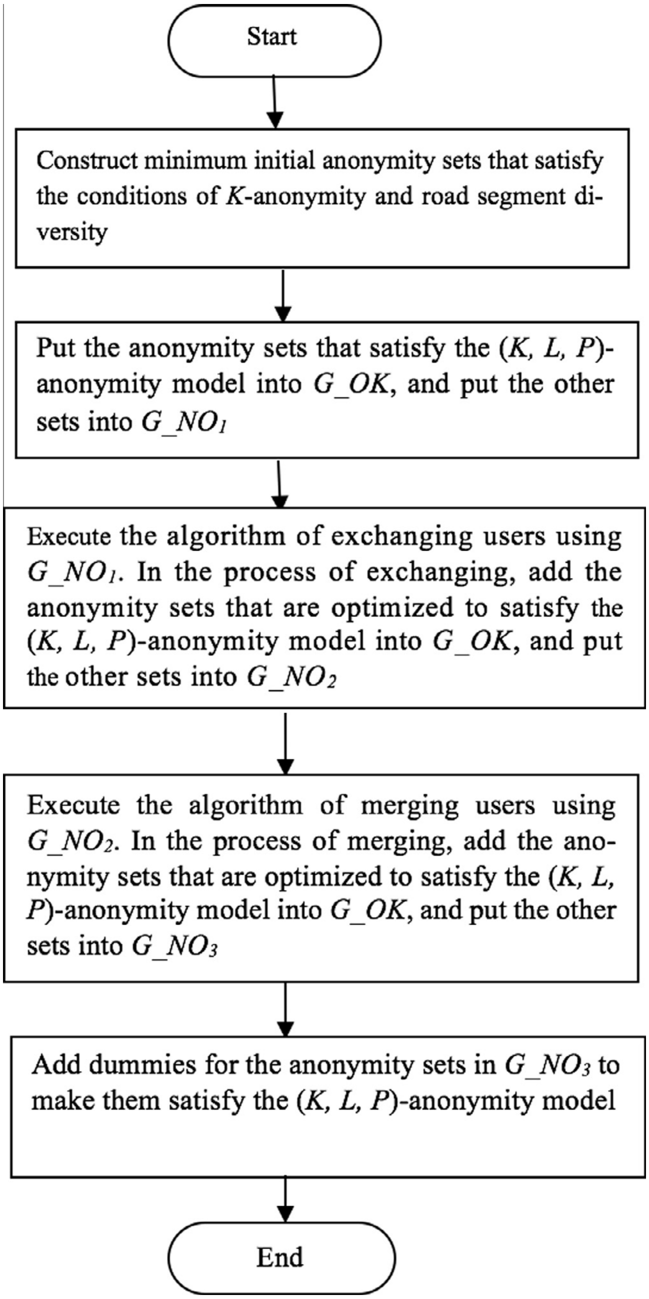
Start
Construct minimum initial anonymity sets that satisfy the conditions of *K*-anonymity and road segment diversity
Put the anonymity sets that satisfy the (*K, L, P*)-anonymity model into *G_OK*, and put the other sets into *G_NO₁*
Execute the algorithm of exchanging users using *G_NO₁*. In the process of exchanging, add the anonymity sets that are optimized to satisfy the (*K, L, P*)-anonymity model into *G_OK*, and put the other sets into *G_NO₂*
Execute the algorithm of merging users using *G_NO₂*. In the process of merging, add the anonymity sets that are optimized to satisfy the (*K, L, P*)-anonymity model into *G_OK*, and put the other sets into *G_NO₃*
Add dummies for the anonymity sets in *G_NO₃* to make them satisfy the (*K, L, P*)-anonymity model
End

Fig. 4. The flowchart of algorithm *EMAGAS*.

A road network example is shown in Fig. 5(a), and the result of a DFS traversal of the network is shown in Fig. 5(b). The anonymity requirements of users {u_1, u_2, u_3, u_4, u_5, u_6, u_7} are {2, 3, 4, 4, 3, 2, 2}, and the road segment requirements are {2, 2, 2, 2, 2, 2, 2}. Using the method described above, two initial anonymity sets {u_1, u_2, u_3, u_4} and {u_5, u_6, u_7} can be constructed.

Table 1c
Values of the variables.

Anonym. set	Mobile user	*|AS|*	*Max*(*u.k*)	*AS.Count_S*	*Max* (*u.sd*)	*AS.Count_SQu*	*u.p*	*AS.Count_SQu*/*|AS|*
AS_1	u_1	3	3	3	2	0	0.4	0
	u_2					1	0.5	1/3
	u_3					2	0.8	2/3
AS_2	u_4	4	3	4	2	1	0.6	1/4
	u_5					3	0.8	3/4
	u_6					3	0.5	3/4
	u_7					1	1	1/4

4.2.1. Algorithm for exchanging users

The initial anonymity sets that do not satisfy the (*K*, *L*, *P*)-anonymity model are processed using the algorithm for exchanging users. The processing steps are described below.

(1) Sort in ascending order the initial anonymity sets in G_NO_1 by the number of unsafe users that do not satisfy the (*K*, *L*, *P*)-anonymity model.
(2) Obtain an unprocessed anonymity set AS_1 from G_NO_1. Extract the anonymity sets that are unprocessed and adjacent to AS_1 in G_NO_1 into *G_Temp*, and sort these sets by the adjacency degrees between them and AS_1 in descending order.
(3) Obtain an unprocessed unsafe user u_1 from AS_1.
(4) Obtain an anonymity set AS_2 from *G_Temp* in turn.
(5) Obtain an unprocessed unsafe user u_2 from AS_2. If these four conditions hold: (a) $u_2.k \leqslant |AS_1|$; (b) $u_2.sd \leqslant AS_1.Count\_S$, and the value of $AS_1.Count\_S$ would not become smaller than before if u_1 and u_2 were exchanged; (c) $u_2.qs \leqslant u_1.qs$; and (d) $u_2.p \geqslant i/K$, where i represents the number of sensitive queries for u_2 sent by the users in AS_1 if u_1 and u_2 were exchanged, then u_1 and u_2 are exchanged.
(6) If AS_1 fits the (*K*, *L*, *P*)-anonymity model, move AS_1 into *G_OK*;
(7) Repeat steps (3) through (6) to optimize the anonymity sets by exchanging unsafe users in different anonymity sets.
(8) Put the remaining anonymity sets in G_NO_1 into G_NO_2.

In the process of exchanging users, the four conditions in Step 5 ensure that unsafe user u_2 in AS_2 becomes a safe user in AS_1, and the existing safe users in AS_1 are still safe after the exchanging operation. That is, the safe users in AS_1 will not be impacted negatively by exchanging u_1 and u_2. If AS_1 contains multiple unsafe users, sort the unsafe users by the sensitivity of the queries sent by them in descending order, then obtain the currently unprocessed unsafe user in sequence.

Algorithm 1 for exchanging users is shown next

```
Input: G_NO1, G_OK, privacy profiles, query sensitivities
Output: G_NO2, G_OK
1: Sort the sets in G_NO1 by the number of unsafe users
2: Tag all sets in G_NO1 as unprocessed
3: If there exists an unprocessed set in G_NO1 then
4:    fetch an unprocessed set AS1 in G_NO1
5:    tag AS1 as processed
6:    sort AS'1 adjacent unprocessed sets in G_temp
7:    for each user u1 in AS1
8:       if G_temp ≠ {} then
9:          fetch an anonymity set AS2 from G_temp
10:           choose unsafe users from AS2 into UU
11:           if UU ≠ {} then
12:              fetch a user u2 from UU
13:              m = AS1.Count_S //if u1 and u2 were exchanged
14:              i = AS.Count_SQu2///if u1 and u2 were exchanged
15:              if u2.k ⩽ |AS1| and u2.sd ⩽ AS1.Count_S
16:                    and AS1.Count_S ⩽ m and u2.p ⩾ i/K
17:                 then exchange u1 and u2
18:                    tag u2 as a safe user
19:                    if AS1 fits (K, L, P)-anonymity model
20:                       then move AS1 into G_OK
21:                          go to Step 3
25:              go to Step 11
26:           else go to Step 8
27:    end of for statement
28: go to Step 3
29: Put the remaining sets in G_NO1 into G_NO2
```

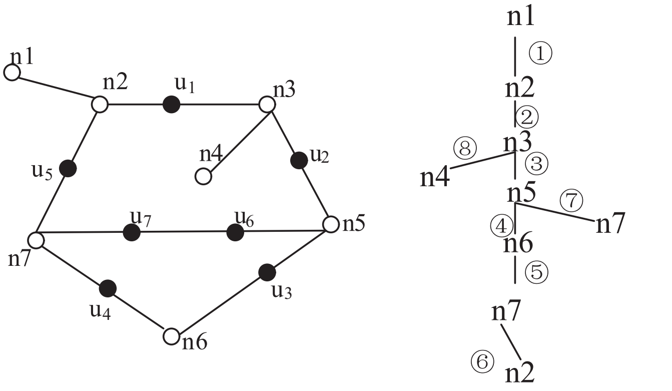

Fig. 5. Result of DFS traversal of a road network.

4.2.2. Algorithm of merging users

Through the process of exchanging users, some initial anonymity sets unfit for the (*K*, *L*, *P*)-anonymity model may become fit for the model. The remaining sets in *G_NO*2 can be further optimized by merging the users of different anonymity sets. First, we give the definitions of a *p*-critical user and *p*-dissatisfied user.

Definition 4 (*p-critical user*). Given a user *u* in anonymity set *AS*, if merging a user *u*′ (*u*′.qs ⩾ *u.qsr*) into *AS* will result in a negative effect on user *u* — that is, making the set sensitivity requirement of u is not satisfied — then u is a p-critical user in *AS*.

Definition 5 (*p-dissatisfied user*). If the set sensitivity requirement of a user in anonymity set AS is dissatisfied, the user is p-dissatisfied

For example, let $AS = \{u_1, u_2, u_3\}$. If $AS.Count\_SQu_1 = 1$ and $u_1.p = 0.4$, then $AS.Count\_SQu1/|AS| = 1/3$. This means that the set sensitivity requirement of u_1 is satisfied. Conversely, if $AS.Count\_SQu_2 = 2$ and $u_2.p = 0.6$, then $AS.Count\_SQu_2/|AS| = 2/3$, which means that the set sensitivity requirement of u_2 is dissatisfied, so u_2 is a *p*-dissatisfied user. Once a user *u*' (*u*'.qs ⩾ $u_1.qsr$) is merged into *AS*, then $AS.Count\_SQu_1/|AS| = 2/4$, which means that the set sensitivity requirement of u_1 becomes dissatisfied. Thus, u_1 is a *p*-critical user in *AS*.

The processing steps of merging users are described next below:

(1) Sort the sets in G_NO_2 by the number of unsafe users in ascending order.
(2) Tag all sets in G_NO_2 as unprocessed.
(3) Obtain an unprocessed anonymity set $AS_1$from G_NO_2. Extract the anonymity sets that are unprocessed and adjacent to AS_1 into *G_Temp*, and sort these sets by the adjacency degrees between them and AS_1 in descending order.
(4) Obtain an anonymity set AS_2 from *G_Temp* in turn.
(5) Extract the unsafe users in AS_2 into *UU*. The sensitivities of the queries sent by the selected users are less than the minimum query sensitivity requirement of the *p*-critical users and *p*-dissatisfied users in AS_1.
(6) Fetch a user *u* from *UU*, and merge it into AS_1.
(7) If AS_1 fits the (*K*, *L*, *P*)-anonymity model, move AS_1 into *G_OK*.
(8) Repeat Steps (3) through (7) to optimize the anonymity sets by merging unsafe users.
(9) Put the remaining anonymity sets in G_NO_2 into G_NO_3.

Algorithm 2 for merging users is described below now

```
Input: $G\_NO_2$, $G\_OK$, privacy profiles, query sensitivities
Output: $G\_NO_3$, $G\_OK$
1: Sort the sets in $G\_NO_2$ by the number of unsafe users
2: Tag all sets in $G\_NO_2$ as unprocessed
3: If there exists an unprocessed set in $G\_NO_2$ then
4:   fetch an unprocessed set $AS_1$ in $G\_NO_2$
5:   tag $AS_1$ in $G\_NO_2$ as processed
6:   sort $AS_1$'s adjacent unprocessed sets in G_temp
7:   if G_temp ≠ {} then
8:     fetch an anonymity set $AS_2$ from G_temp
9:     $m_1$ = minimum u.qsr//u∈ set of p-critical users
10:     $m_2$ = minimum u.qsr//u∈ set of p-dissatisfied users
11:     min_qsr = min($min_1$, $min_2$)
12:     for each user u in $AS_2$
13:       if u.qs < min_qsr
14:         then put u into UU
15:     if UU ≠ {} then
16:       fetch a user u from UU
17:       merge u into $AS_1$
18:       if $AS_1$ fits (K, L, P)-anonymity model
19:         then move $AS_1$ into $G\_OK$
20:           go to Step 3
21:         else go to Step 15
22:   go to Step 7
23: Put the remaining sets in $G\_NO_2$ into $G\_NO_3$
```

5. Applying *EMAGAS*: an illustration

To illustrate the ability of the proposed algorithm, *EMAGAS*, let us assume that Fig. 6 is a road network example. Using the method that we presented, the initial anonymity sets are generated. The personalized privacy profiles, and the query sensitivities of 18 users on the road network, and the initial anonymity sets generated are shown in Table 2.

According to Definition 3 described, it can be shown that AS_1, AS_3, AS_6 and AS_7 do not satisfy the (*K*, *L*, *P*)-anonymity model. So they are put into the set G_NO_1. Anonymity sets AS_1, AS_3, AS_6, and AS_7 are processed using the algorithm for exchanging users. User u_2 in AS_1 and user u_{15} in AS_6 are exchanged, so AS_1 satisfies the (*K*, *L*, *P*)-anonymity model. Other pairs of unsafe users in different anonymity sets cannot be exchanged. The remaining anonymity sets that do not satisfy the (*K*, *L*, *P*)-anonymity model, as shown in Table 3, are put into G_NO_2.

By applying the algorithm for merging users, user u_{16} in AS_6 is merged into AS_7, and AS_7 then satisfies the (*K*, *L*, *P*)-anonymity model. The remaining anonymity sets that do not satisfy the (*K*, *L*, *P*)-anonymity model, as shown in Table 4, are put into G_NO_3. Finally, we can make the anonymity sets in G_NO_3 satisfy the (*K*, *L*, *P*)-anonymity model by inserting dummies into them.

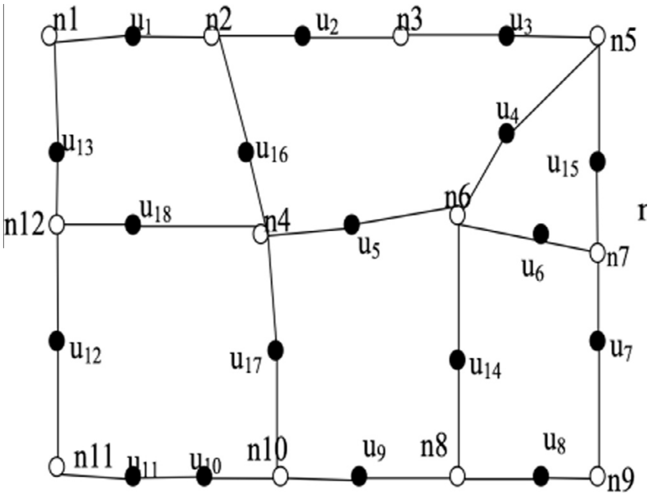

Fig. 6. Example road network RN_2.

6. Experimental analysis of *Emagas*

6.1. Experimental dataset and parameter settings

To validate the effectiveness of the proposed *EMAGAS* algorithm, we use a real dataset from a California road network with 21,048 intersections and 21,693 road segments (Li et al., 2005). 32,400 simulated users are generated on the network, and their personalized privacy profiles are created also based on the specified rules. The mobile users can dynamically set the parameter values of their personalized privacy profiles. The settings of the query sensitivities and the parameters used to define the personalized privacy profiles are described in Table 5.

The query sensitivities are categorized into five levels labeled {0, 0.25, 0.5, 0.75, 1}. All users are completely insensitive to the queries with sensitivity level "0" and are most sensitive to the queries with sensitivity level "1". The anonymity requirement of each user is randomly set to an integer within [2, *kmax*], where *kmax* represents the maximum anonymity requirement. With increasing *kmax*, the average number of users and average number of road segments in the anonymity sets generated usually increase, which means that the privacy protection requirements of the users become more demanding. In our experiments, *kmax* is set to an integer within [2, 30].

The road segment diversity requirement of each user is randomly set to an integer within [2, 10]. The query sensitivity requirement of each user is randomly determined as one of the five sensitivity levels. We experimentally analyzed and compared the performance of the *EMAGAS* algorithm based on the metrics for information entropy, dummy ratio, query cost and anonymization time. As shown in Figs. 7a–7d, we chose [0.6, 1] as the range of values for the users' set sensitivity requirements.

Table 2
Initial anonymity sets.

Anonymity sets	Users	*k*	*sd*	*qsr*	*p*	*qs*
AS_1	u_1	2	2	0.5	0.6	0.5
	u_2	2	2	0.4	0.6	1
AS_2	u_3	3	2	0.4	0.7	0
	u_4	3	2	1	0.4	0.25
	u_5	3	2	0.6	0.5	0.5
AS_3	u_6	2	2	0.4	0.8	1
	u_7	2	2	0.25	0.4	0.5
AS_4	u_8	3	2	0.5	1	0
	u_9	3	2	0.5	0.7	0.5
	u_{10}	2	2	0.6	0.5	0.75
AS_5	u_{11}	3	2	0.4	0.7	0.25
	u_{12}	2	2	0.7	0.4	0.5
	u_{13}	2	2	0.5	0.5	0
AS_6	u_{14}	3	2	0.3	0.7	1
	u_{15}	2	2	0.8	0.3	0.25
	u_{16}	3	2	1	0.3	0.25
AS_7	u_{17}	3	2	0.5	0.7	1
	u_{18}	2	2	0.6	0.8	0.75

Table 3
Anonymity sets in G_NO_2.

Anonymity sets	Users	*k*	*sd*	*Qsr*	*p*	*qs*
AS_3	u_6	2	2	0.4	0.8	1
	u_7	2	2	0.25	0.4	0.5
AS_6	u_{14}	3	2	0.3	0.7	1
	u_2	2	2	0.4	0.6	1
	u_{16}	3	2	1	0.3	0.25
AS_7	u_{17}	3	2	0.5	0.7	1
	u_{18}	2	2	0.6	0.8	0.75

Table 4
Anonymity sets in G_NO_3.

Anonymity sets	Users	*k*	*sd*	*qsr*	*p*	*qs*
AS_3	u_6	2	2	0.4	0.8	1
	u_7	2	2	0.25	0.4	0.5
AS_6	u_{14}	3	2	0.3	0.7	1
	u_2	2	2	0.4	0.6	1

Table 5
Experimental parameter settings.

Parameter	Setting
Number of intersections	21,048
Number of road segments	21,693
Number of users	32,400
K	[2, *maxk*]
Sd	[2, 10]
P	[0.6, 1]
qsr	{0.25, 0.5, 0.75, 1}

6.2. Experimental analysis of the performance of the EMAGAS algorithm

We compared the performance of the *EMAGAS* algorithm with P^3RN (Pan et al., 2014) via experiments based on a real-world road network and generated privacy profiles for the previously discussed users using four metrics: *average information entropy*, *average dummy ratio*, *query cost* and *average anonymization time*. Among the existing location privacy protection works based on road networks, P^3RN seems to be the first algorithm that addresses the query sensitivity problem. We implemented both algorithms are implemented in JAVA. We also ran them on a desktop PC with a dual-core AMD 2.2 GHz processor and 4 GB of main memory.

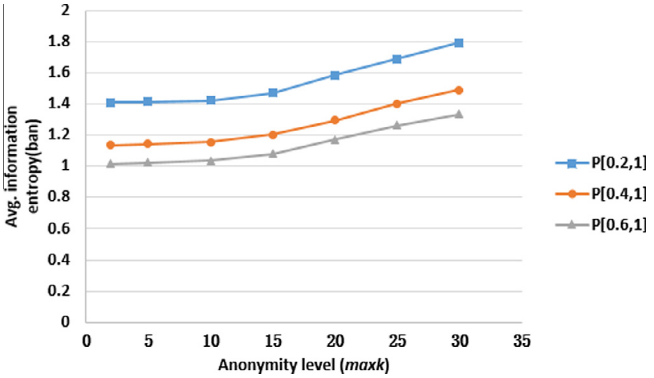

Fig. 7a. Average information entropy.

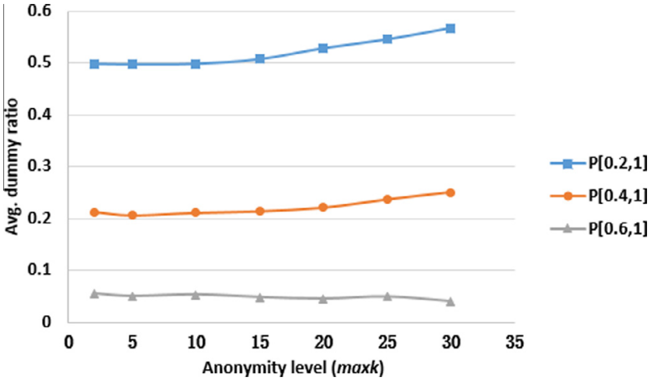

Fig. 7b. Average dummy ratio.

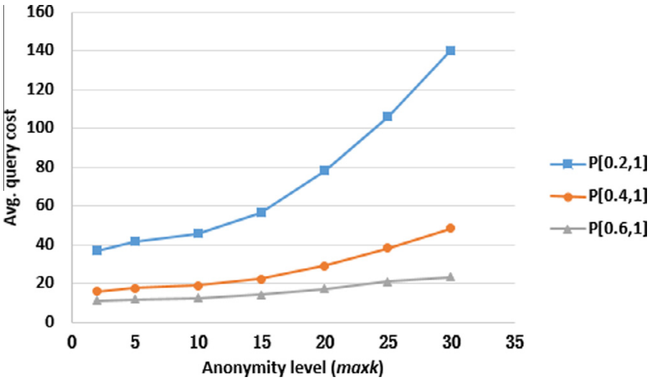

Fig. 7c. Average query cost.

6.2.1. Information entropy

Information entropy indicates the protection strength of the privacy-preserving algorithm. Tests of information entropy are used to determine the effect of the cloaking region scale and user density in the region on the cloaking algorithm (Wang and Liu, 2009). Larger information entropy implies that the attacker is less certain of the mobile user's specific location. The *average information entropy* can be calculated via $h = -K\log(1/L,10)$, where L is the number of road segments in the cloaking region, and K the number of the users in the anonymity set, and then $\overline{H} = \frac{-\sum_{i=1}^{n} K_i \log\left(\frac{1}{L_i},10\right)}{user_{num}}$, where n is the number of anonymity sets, and $user_{num}$ is the total number of users.

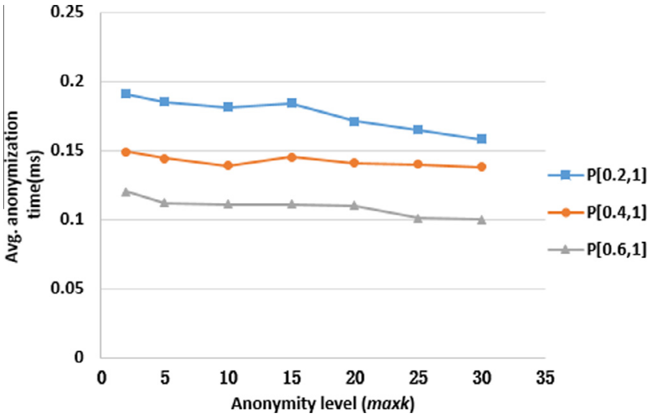

Fig. 7d. Average anonymization time.

The test results for average information entropy are shown in Fig. 8a. For both algorithms, information entropy increases with increasing *maxk* owing to more users and road segments included in the anonymity sets. The protection strength of *EMAGAS* is higher than that of P^3RN when $maxk \geqslant 15$; when $maxk > 15$, the corresponding information entropy of both algorithms tends to be the same.

6.2.2. Dummy ratio

The success rate of the anonymity sets that are generated to satisfy a specific anonymity model can be increased by adding dummies to the anonymity sets that do not satisfy the model (Kido et al., 2005; Pan et al., 2014). The *dummy ratio* refers to the percentage of added dummies in the anonymity sets. With the *EMAGAS* algorithm, 100% success in the generation of anonymity sets that satisfy the (*K*, *L*, *P*)-anonymity model can be achieved by adding dummies after the exchanging and merging processes are completed. The average dummy ratio is calculated using $D = \frac{d_{num}}{d_{num}+user_{num}}$, where d_{num} denotes the number of dummies, and $user_{num}$ denotes the total number of users.

The experimental results are shown in Fig. 8b. The performance of *EMAGAS* is worse than P^3RN when both algorithms are compared using the dummy ratio, but the gaps become smaller with increasing values of *maxk*. This is because the reason for adding dummies into an anonymity set for P^3RN is that the *P* condition is not satisfied, whereas for *EMAGAS* may be that any one of the conditions (*K*, *L*, *P*) is not satisfied, and thus more dummies must be added. However, compared with P^3RN, *EMAGAS* may generate smaller anonymity sets and smaller cloaking regions, which contribute to providing more accurate and useful recommendations for the LBS provider based on the anonymous query information.

6.2.3. Query cost

The cost of processing a location-based query consists of the query execution cost of the LBS platform and the communication cost of transferring the query result back to the mobile user (Wang and Liu, 2009). We consider only the query execution cost in this study. Given the *set of the road segments SS* corresponding to anonymity set *AS*, *VS* is the *set of endpoints of the road segments* in *SS*. According to Chow et al. (2011) for all $v \in VS$. If there is a road segment for which one endpoint is v and another endpoint is not in *VS*, v is referred to as an *open endpoint*. Let the *set of open endpoints* in *VS* be *OVS*. Then the average query cost is calculated as follows: (1) Calculate the query cost of each anonymity set via $QCost(AS) = AS.Count\_S + |OVS|$, where $AS.Count\_S$ denotes the number of road segments in *SS*, and $|OVS|$ represents the number of the open endpoints in the set *OVS*. Then calculate the average query cost with $\overline{QCost} = \frac{\sum_{i=1}^{n}(AS_i.Count\_S+|OVS_i|)}{n}$, where n is the number of anonymity sets.

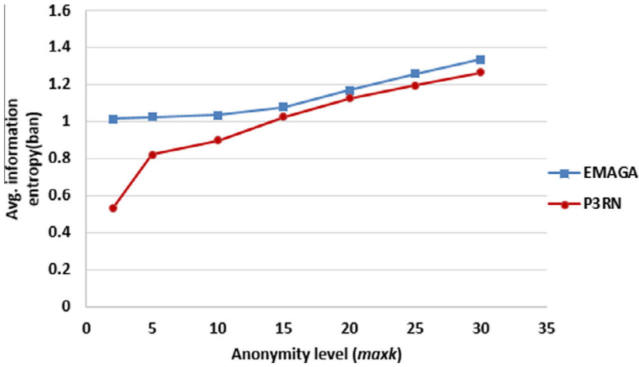

Fig. 8a. Average information entropy.

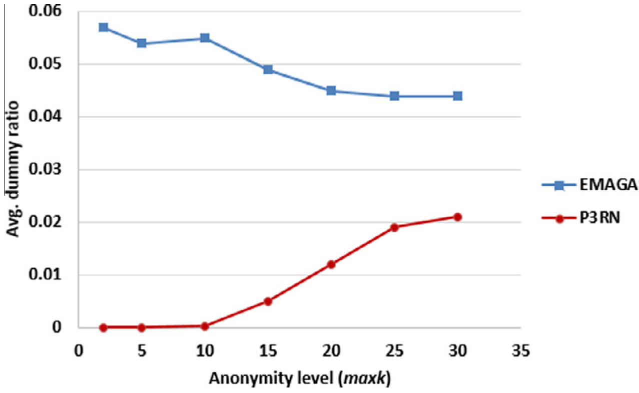

Fig. 8b. Average dummy ratio.

The experimental results for the query cost are shown in Fig. 8c. The average query cost of both algorithms increases with increasing *maxk*, and the average query cost of *EMAGAS* is lower than that of P^3RN when $maxk \geqslant 10$. Thus, the performance of *EMAGAS* is better than that of P^3RN when both algorithms are compared using query cost. The reason lies in the improvements in the process of generating anonymity sets in the *EMAGAS* algorithm. Using P^3RN, the size of each initial anonymity set is *maxk*, while the size of the anonymity set resulting from the merging operation may become 2 × *maxk*. Using the *EMAGAS* algorithm, the sizes of the initial anonymity sets are dynamically determined to be no more than *maxk*. The merging operation involves bringing one user from an adjacent anonymity set into the specified anonymity set. Analysis of the experimental results suggested that the average size of the anonymity sets generated using *EMAGAS* is less than that of the anonymity sets generated using P^3RN.This is the main reason why the query cost of *EMAGAS* is lower than P^3RN.

6.2.4. Anonymization time

The *anonymization time* indicates the efficiency of the algorithm. The shorter the time, the higher the efficiency of the algorithm. The average anonymization time is given by $\bar{T} = T/user_{num}$, where $user_{num}$ represents the total number of users on the road network, and *T* represents the runtime of the anonymization algorithm.

Fig. 8d presents the anonymization time of both anonymization algorithms. When *maxk* is equal to or greater than 10, the anonymization time using P^3RN is approximately 0.8 ms, whereas the anonymization time using *EMAGAS* becomes less than 0.1 ms.

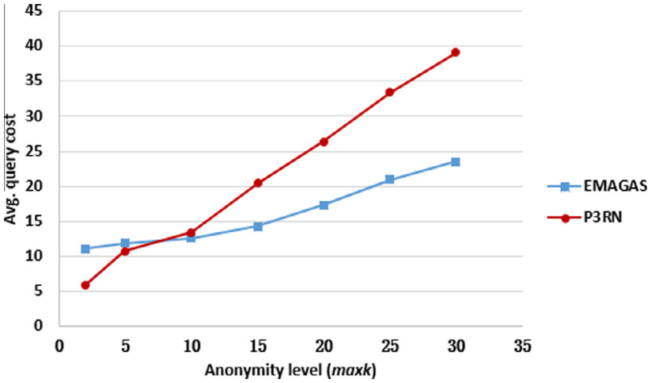

Fig. 8c. Average query cost.

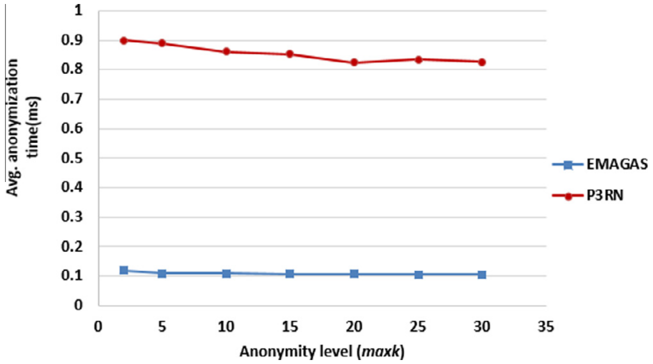

Fig. 8d. Average anonymization time.

Thus, the efficiency of *EMAGAS* is much higher than that of P^3RN. The main reason lies in the different processing methods in the main steps of both algorithms.

7. Conclusions

7.1. Discussion of the results

In this article, a privacy-preserving service framework for an m-commerce alliance (MCA) providing LBS was established, which enables the integration of the service resources of multiple information service providers and contributes to achieving a win–win for all participants in the alliance. The privacy information of the mobile users can be prevented from collection and misuse by the information service providers and vendors, and the users can receive comprehensive information services.

One of the keys to implementing the proposed privacy-preserving service framework for MCA is the anonymization process based on an appropriate anonymity model. The personalized privacy profile is formally defined as four-tuple (k, qsr, sd, p) and can be used to describe the mobile user's privacy requirements, including anonymity, query sensitivity, segment diversity and set sensitivity requirements. According to the defined personalized privacy profile, a (K, L, P)-anonymity model is introduced and formally described.

Based on the (K, L, P)-anonymity model, a new privacy-preserving algorithm named *EMAGAS* was proposed, which features the construction of minimal initial K-anonymity sets, an exchanging process and a merging process. The processes of exchanging and merging are discussed step by step in detail and formally described. The proposed *EMAGAS* algorithm is used to protect the mobile user's location and identifier but also contributes to the protection of sensitive information. To demonstrate the ability of *EMAGAS*, we illustrated the results of each major step through an example application. Using a real road network and generated privacy profiles of the mobile users, the feasibility of *EMAGAS* was validated by experimentally analyzing and comparing the performance of *EMAGAS* and P^3RN using the metrics of information entropy, dummy ratio, query cost, and anonymization time. The experimental results demonstrated that *EMAGAS* has advantages in the performance of information entropy, query cost and anonymization time to different extents.

7.2. Contributions and implications

There are two main contributions. First, from the business perspective, a privacy-preserving service framework for the mobile commerce alliance providing LBS was established, and the (K, L, P)-anonymity model was used to describe the mobile user's personalized privacy profile is presented. Second, from the technical perspective, a new privacy-preserving algorithm named *EMAGAS* is proposed, and the feasibility and performance advantages of *EMAGAS* were validated experimentally.

In practice, however, there are many LBS applications, and different information service providers usually have different service resources. A single information service provider is unable to provide mobile users with comprehensive and high-quality LBS recommendations. The proposed service framework for MCA not only contributes to protecting the privacy of mobile users but also facilitates sharing of the service resources. Moreover, given the wide use of smart mobile phones and the popularity of various mobile applications, the mobile user's privacy information may be disclosed at any location and time. Location fixing and mobile phone number binding are commonly used features in various m-commerce applications, which may result in the disclosure of privacy information of mobile users, such as location, identity and other sensitive information. Privacy disclosure risk has become one of the constraints affecting the further adoption of mobile services. Thus, it is significant to explore a solution for protecting the privacy of mobile users theoretically and technologically.

7.3. Limitations and future research directions

Our work has some limitations. First, an application architecture of the presented privacy-preserving service framework for the MCA must be designed, and the performance of the proposed algorithm must be further tested using not only a real road network but also a real m-commerce application scenario. Second, problems regarding how each of the information service providers generates the list of query results based on the anonymous query request and how the MCA platform refines the results received from multiple information service providers must be addressed.

With the emergence of the big data era, m-commerce provides unprecedented opportunities for service providers to implement precision marketing in real-time. Meanwhile, new challenges related to privacy issues in m-commerce have been identified. On the one hand, the problem of information overload and the risk of privacy information disclosure are increasingly acute (Feng et al., 2014). Recommendation systems are effective solutions to solve the problem of information overload. Providing high-quality service recommendations under the condition of protecting the privacy information of mobile users is a challenging problem (Zhao et al., 2014).

There have been many published works on personalized recommendation methods (Moradi et al., 2015; Meng et al., 2013; Huang et al., 2010; Piao et al., 2009; Sarwar et al., 2001). As far as we know though, few reported works related to m-commerce have proposed a holistic solution for service recommendation based on privacy awareness. On the other hand, sensitive information protection based on location big data is also a challenging problem. Location big data not only involve location privacy information of the users but also implies sensitive information such as the personal habits, health conditions and social positions of users. The attackers may extract, integrate, analyze and find sensitive information on mobile users via the accumulation and association of big data from multiple sources. Solutions to prevent attackers from obtaining sensitive information on mobile users via the connections between the location data and service-related data must be studied in depth.

Acknowledgments

This research was funded by the Natural Science Foundation of Hebei Province of China (No. F2015210106), and was partially supported by Grants from the National Natural Science Foundation of China (Nos. 61303017, 61379048).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.elerap.2016.03.005.

References

Bamba, B., Liu, L., 2008. Supporting anonymous location queries in mobile environments with privacy grid. In: Proceedings of the 17th International Conference on the World Wide Web. Beijing, China, pp. 237–246.

Belanger, F., Crossler, R.E., 2011. Privacy in the digital age: a review of information privacy research in information systems. MIS Q. 35 (4), 1017–1042.

Bereford, A.R., Stajano, F., 2003. Location privacy in pervasive computing. IEEE Pervasive Comput. 2 (1), 46–55.

Bettini, C., Wang, X.S., Jajodia, S., 2005. Protecting privacy against location-based personal identification. In Proceedings of the VLDB Workshop on Secure Data Management. ACM Press, New York, NY, pp. 185–199.

Chow, C.Y., Mokbel, M.F., 2007. Enabling private continuous queries for revealed user locations. In: Advances in Spatial and Temporal Databases. Lecture Notes in Computer Science, vol. 4605. Springer, Berlin, Germany.

Chow, C.Y., Mokbel, M.F., Liu, X., 2011. Query-aware location anonymization for road networks. Geoinformatica 15 (3), 571–607.

Culnan, M.J., 1995. Consumer awareness of name removal procedures: implications for direct marketing. J. Direct Marketing 9 (2), 10–19.

Culnan, M.J., Armstrong, P.K., 1999. Consumer privacy concerns, procedural fairness, and impersonal trust: an empirical investigation. Organiz. Sci. 10 (1), 104–115.

Enck, W., Gilbert, P., Han, S., Tendulkar, V., Cox, L.P., Jung, J., McDaniel, P., Sheth, A., 2014. TaintDroid: an information-flow tracking system for realtime privacy monitoring on smartphones. ACM Trans. Comput. Syst. 32 (2), 99–106.

Feng, D., Zhang, M., Hai, L., 2014. Big data security and privacy protection. Chin. J. Comput. 37 (1), 246–258.

Ghinita, G., Kalnis, P., Khoshgozaran, A., Shahabi, C., Tan, K.L., 2008. Private queries in location based services: anonymizers are not necessary. In Proceedings of the 28th ACM International Conference on Management of Data, Vancouver, Canada. ACM Press, New York, NY, pp. 121–132.

Greenleaf, G. 2014. Global data privacy laws: 89 countries, and accelerating. Privacy Laws & Business International Report 115, Special Supplement, Social Science Electronic Publishing.

Gruteser, M., Grunwald, D., 2003. Anonymous usage of location-based services through spatial and temporal cloaking. In: Proceedings of the International Conference on Mobile Systems, Applications, and Services, San Francisco, CA. ACM Press, New York, NY, pp. 163–168.

Guo, X., Zhang, X., Sun, Y., 2016. The privacy-personalization paradox in m-health services acceptance of different age groups. Electron. Commerce Res. Appl. 16, 55–65.

Hamad, F., Smalov, L., James, A., 2009. Energy-aware security in m-commerce and the Internet of Things. IETE Tech. Rev. 26 (5), 357–362.

Hong, J.I., Landay, J.A., 2004. An architecture for privacy-sensitive ubiquitous computing. In: Boston, M.A. (Ed.), Proceedings of the International Conference on Mobile Systems, Applications, and Services. IEEE Computer Science Press, Los Alamitos, CA, pp. 177–189.

Huang, C., Yin, J., Wang, J., Wang, H., 2010. Uncertain neighbors' collaborative filtering recommendation algorithm. Chin. J. Comput. 33 (8), 1369–1377.

Huo, Z., Meng, X., Huang, Y., 2013. PrivateCheckIn: trajectory privacy-preserving for check-in services in MSNS. Chin. J. Comput. 36 (4), 716–726.

Jiang, X., Zhong, Q., Ji, S., 2010. Research in online privacy: concept, progress and trend. Inf. Sci. 28 (2), 305–310.

Khoshgozaran, A., Shahabi, C., 2007. Blind evaluation of nearest neighbor queries using space transformation to preserve location privacy. In: Proceedings of 10th International Symposium of Advances in Spatial and Temporal Databases, Boston, MA, July 16–18, pp. 239–257.

Kido, H., Yanagisawa, Y., Satoh, T., 2005. An anonymous communication technique using dummies for location-based services. In: Proceedings of the IEEE International Conference on Pervasive Services, Santorini, Greece. IEEE Computer Society Press, Los Alamitos, CA, pp. 88–97.

Kim, J., 2015. Impact of concerns for information privacy on behavioral intention of providing privacy information in the context of m-commerce. J. Internet Electron. Commerce Res., 115–127

Kim, E., Lee, B., 2009. E-service quality competition through personalization under consumer privacy concerns. Electron. Commerce Res. Appl. 8 (4), 182–190.

Lee, D.J., Ahn, J.H., Bang, Y., 2010. Managing consumer privacy concerns in personalization: a strategic analysis of privacy protection. MIS Q. 35 (2), 423–444.

Li, Y., 2014. A multi-level model of individual information privacy beliefs. Electron. Commerce Res. Appl. 13 (1), 32–44.

Li, F., Cheng, D., Hadjieleftheriou, M., Kollios, G., Teng, S.H., 2005. On trip planning queries inspatial databases. In: Proceedings of the 9th International Symposium on Advances in Spatial and Temporal Databases, Angra dos Reis, Brazil, August 22–24. Springer, Berlin, Germany.

Li, P.Y., Peng, W.C., Wang, T.W., Ku, W.S., Xu, J., Hamilton Jr., J.A., 2008. A cloaking algorithm based on spatial networks for location privacy. In: Proceedings of IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing. IEEE Computer Science Press, Los Alamitos, CA.

Lin, X., Zhou, L., Chen, P., Gu, J., 2012. Privacy preserving reverse nearest-neighbor queries processing on road network. In: Lecture Notes in Computer Science. Springer, Berlin, Germany, pp. 19–28.

Liu, C., Marchewka, J.T., Lu, J., 2004. Beyond concern: a privacy-trust-behavioral intention model of electronic commerce. Inf. Manage. 42 (2), 127–142.

Liu, Y., Zhang, T., Jin, X., Cheng, X., 2015. Personal privacy protection in the era of big data. J. Comput. Res. Dev. 52 (1), 229–248.

Malhotra, N.K., Kim, S.S., Agarwal, J., 2004. Internet users' information privacy concerns (IUIPC): the construct, the scale, and a causal model. Inf. Syst. Res. 15 (4), 336–355.

Mason, R.O., 1986. Four ethical issues of the information age. MIS Q. 10 (1), 5–12.

Meng, X.W., Hu, X., Wang, L.C., Zhang, Y.J., 2013. Mobile recommender systems and their applications. J. Software 24 (1), 91–108.

Mokbel, M.F., Chow, C., Aref, W.G., 2006. The new Casper: query processing for location services without compromising privacy. In: Proceedings of the International Conference on Very Large Data Bases, Seoul, Korea, September. ACM Press, New York, NY, pp. 763–774.

Moradi, P., Ahmadian, S., Akhlaghian, F., 2015. An effective trust-based recommendation method using a novel graph clustering algorithm. Phys. A 436, 462–481.

Mouratidis, K., Yiu, M.L., 2010. Anonymous query processing in road networks. IEEE Trans. Knowl. Data Eng. 22 (1), 2–15.

Pan, X., Meng, X., 2013. Location preserving without exact locations in mobile services. Front. Comput. Sci. 7 (3), 317–340.

Pan, X., Xu, J., Meng, X., 2012. Protecting location privacy against location-dependent attacks in mobile services. IEEE Trans. Knowl. Data Eng. 24 (8), 1506–1519.

Pan, X., Wu, L., Piao, C., 2014. P^{3}RN: personalized privacy protection using query semantics over road networks. In: Proceedings of the 15th International Conference Web Age Information Management, Macau, China, June. Springer, Berlin, Germany, pp. 16–18.

Piao, C., Zhao, J., Zheng, L., 2009. Research on entropy-based collaborative filtering algorithm and personalized recommendation in e-commerce. SOCA 3 (2), 147–157.

Piao, C., Dong, C., Cui, L., 2013. A novel scheme on service recommendation for mobile users based on location privacy protection. In: Coventry, U.K. (Ed.), Proceedings of the IEEE 10th International Conference on e-Business Engineering, September 11–13. IEEE Computer Society Press, Washington, DC, pp. 300–305.

Robello-Monedero, D., Forne, J., Solanus, A., Martinez-Balleste, A., 2008. Private location-based information retrieval through user collaboration. Comput. Commun. 31 (6), 762–774.

Salo, J., Sinisalo, J., Karjaluoto, H., 2008. Intentionally developed business network for mobile marketing: a case study from Finland. J. Business Ind. Marketing 23 (7), 497–506.

Sarwar, B., Karypis, G., Konstan, J., Riedl, J., 2001. Item-based collaborative filtering recommendation algorithms. In: Proceedings of the Tenth International World Wide Web Conference. ACM Press, New York, NY, pp. 285–295.

Sharma, S., Crossler, R.E., 2014. Disclosing too much? Situational factors affecting information disclosure in social commerce environment. Electron. Commerce Res. Appl. 13 (5), 305–319.

Shen, Q., 2013. Internet information privacy concerns and privacy protection behavior of college students in Shanghai. J. Int. Commun. 35 (2), 120–129.

Smith, H.J., Dinev, T., Xu, H., 2011. Information privacy research: an interdisciplinary review. MIS Q. 35 (4), 989–1016.

Solanus, A., Martinez-Balleste, A., 2007. Privacy protection in location-based services through a public-key privacy homomorphism. In: Proceedings of the 4th European Conference on Public Key Infrastructure: Theory and Practice. Springer, Berlin, Germany.

Solanus, A., Martinez-Balleste, A., 2008. A TTP-free protocol for location privacy in location-based services. Comput. Commun. 31, 1181–1191.

Solanus, A., Domingo-Ferrer, J., Martinez-Balleste, A., 2008. Location privacy in location-based services: beyond TTP-based schemes. In: Proceedings of the 1st International Workshop on Privacy in Location-Based Applications, Malaga, Spain, October 9, CEUR-WS.org, vol. 397, pp. 12–23.

Straub Jr., D.W., Collins, R.W., 1990. Key information liability issues facing managers: software piracy, proprietary databases, and individual rights to privacy. MIS Q. 14 (2), 143–156.

Tähtinen, J., Salo, J., 2004. Special features of mobile advertising and their utilization. In: Proceedings of the 33rd European Marketing Academy Conference, Murcia, Spain.

Terrovitis, M., 2011. Privacy preservation in the dissemination of location data. In: Proceedings of the 17th ACM International Conference on Knowledge Discovery and Data Mining. ACM Press, New York, NY, pp. 6–18.

Wang, T., Liu, L., 2009. Privacy-aware mobile services over road networks. In: Proceedings of the 35th International Conference on Very Large Data Bases, Lyon, France. ACM Press, New York, NY, pp. 1042–1053.

Warren, S., Brandeis, L., 1890. The right to privacy. Harvard Law Rev. 4 (5), 193–220.

Wu, L., Pan, X., Piao, C., Li, Z., 2014. Personalized privacy preservation against sensitivity homogeneity attack in location-based services. J. Comput. Appl. 34 (8), 2356–2360.

Xue, J., Liu, X., Yang, X., 2011. A location privacy preserving approach on road network. Chin. J. Comput. 5 (34), 866–878.

Zhang, R.D., Chen, J.Q., Lee, C.J., 2013. Mobile commerce and consumer privacy concerns. J. Comput. Inf. Syst. 53 (4), 31–38.

Zhao, Y., Ye, J., Henderson, T., 2014. Privacy-aware location privacy preference recommendations. In: Proceedings of the 11th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services, London, UK, December 2–5. ACM Press, New York, NY, pp. 120–129.

Zhou, T., 2011. The impact of privacy concern on user adoption of location-based services. Ind. Manage. Data Syst. 111, 212–226.