Journal of Theoretical and Applied Information Technology
31st December 2017. Vol.95. No 24

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A SURVEY ON PRIVACY OF LOCATION-BASED SERVICES: CLASSIFICATION, INFERENCE ATTACKS, AND CHALLENGES

[1]MOHAMAD SHADY ALRAHHAL, [2]MAHER KHEMAKHEM, [3]KAMAL JAMBI

[1]King Abdulaziz University (KAU), Department of Computer Science, Saudi Arabia

E-mail: [1]shady.rahal1986@gmail.com, [2]maherkhemekhem@yahoo.com, [3]kjambi@kau.edu.sa

ABSTRACT

In recent years, Location-Based Services (LBS) have become very popular, especially in the light of enhancements that are daily performed on both mobile devices and wireless networks. The popularity of LBS is derived from its valuable benefits, where they enable the users to search for nearest Points of Interest (POI), share ideas and comments, and enjoy playing games, making our life easier and more enjoyable. However, LBS have some risks associated with it. The privacy issue is considered one of the most important risks in this field since the users are forced to build their queries based on their real geographic locations. This paper studies the different privacy protection approaches through a survey, where a new classification is proposed based on the amount of collaboration between LBS users and LBS server. The protection goals (identity ID, Location Information LI, and Temporal Information TI) that any LBS user aims to protect are defined and measured. Based on the provided protection goals, the most advanced inference attacks (Location Homogeneity Attack LHA, Map Matching Attack MMA, Query Sampling Attack QSA, and Semantic Location Attack SLA) are analyzed and evaluated. As for challenges in LBS privacy protection field, an eight research questions and open problems are explored. In addition, we present some rules-based recommendations, which can help the LBS users to select the most optimal way to achieve a higher privacy protection level.

Keywords: *Inference Attacks, Research Questions, Privacy Protection, Protection Goals, Privacy Metrics, Rule.*

1. INTRODUCTION

Recently, there has been a rapid development in the world of mobile technology and Internet Networking resulting a variety of new mobile devices and social networks as well as the development of emerging Internet of Things (IoT) services [1, 2, 3, 4, 5]. Most of these developments rely on location-based services (LBS) or LBS applications. IoT devices, smartphones, as well as LBS all have built on Global Positioning System (GPS) with a powerful computation capability. Users can easily get the benefits of LBS applications through downloading them from various sites such as the Apple Store or Google Play Store. With the help of these applications, users can send their queries together with their identities, locations, interests, and other information (e.g., time, query range) to the LBS server. In return, they enjoy the benefits provided by LBS such as searching for the Points of Interests (POI) like the nearest shopping mall, supermarket, restaurant [6], or even ask help in emergency situations [7]. Moreover, integrating LBS applications with wireless communication technologies have enabled the creation of location-based social networking services, such as Foursquare, Twinkle, and GeoLife [8]. This integration bridges the gap between the physical world and the digital online social networking services, opening the door to new challenges. To accurately identify these challenges, we need to have a look at the concept of the location-based social networks and how to use LBS-enabled applications, which form the skeleton of such networks.

An online social networking service can be defined as a participatory digital representation of real-world social networks. The social networking services reveal the social connection networks of the user and also enhance the growth by allowing

Journal of Theoretical and Applied Information Technology
31st December 2017. Vol.95. No 24

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

them to share and communicate with ideas, activities, events, news, and interests in a much easier fashion. The addition of spatial aspect in a location-based social networking service strengthens the connection between the social networking services and the real-world social networks. According to this added spatial aspect, three graphs, namely, user-location graph, user-user graph, and location-location graph can be explored. Figure 1, illustrates the three graphs mentioned above besides to the user correlation and the location correlation.

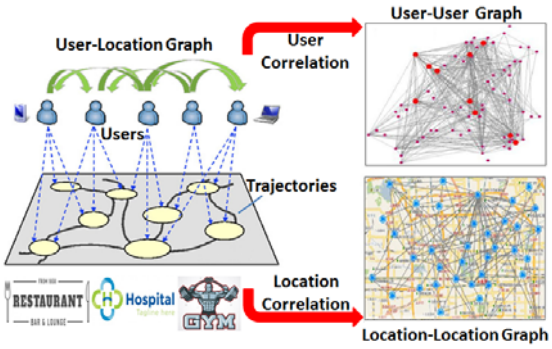


Figure 1:The Concept of Location-Based Social Networks [9].

- ***User-Location Graph.*** In this graph (shown on the left of Figure 1), the location is tightly-coupled with the user, which in turn reveals the travel histories of the user.
- ***User-User Graph. In this graph*** (shown on the top-right of Figure 1), relations of the user are represented such as the friendship relations. In addition, this graph reveals the correlation of the users based on their locations, where two users may be connected if they have visited the same location or similar types of places.
- ***Location-Location Graph.*** In this graph (shown on the bottom-right of Figure 1), the correlations of the users' locations are illustrated, where the physical distances between the locations or the similarities between the locations in terms of their functionality/category (health centers, sports centers, or religious centers for examples) are revealed.

According to the three previous graphs, a lot of sensitive data about the users could be derived from the travel histories, the friendship relations, and the similarities among the locations. This, in turn, reflects a personal information about the users and their daily lives, which could be exploited and misused by the attackers. As a result, the privacy of the users is threatened.

Since LBS form the skeleton of the location-based social networking services and to accurately state the problem, we need to have a look at the LBS-enabled applications usage. Figure 2 illustrates the classical scenario of using LBS-enabled applications.

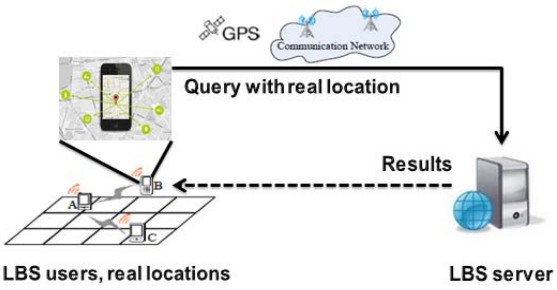


Figure 2: The Classical Scenario of LBS-enabled Applications.

According to Figure 2, LBS user builds his/her queries based on his/her real geographical location and then sends them to the LBS server, asking for the benefits provided by the LBS-enabled applications. After processing the queries on the LBS server side, the results are returned to the LBS user. The units of the query sent to the LBS server are summarized in Table 1.

Table 1: The General Form of The LBS Query.

Symbol	<X, Y>	POI	R	ID
Description	The coordinates of the exact location of the LBS user.	The queried interest.	The queried range.	The identity of the LBS user.

The simple and traditional scenario illustrated above includes risks, even as users are carried away by the advantages of LBS. The reason behind these risks is that the locations of the LBS user may be tracked by an attacker, constructing a malicious profile about the user. This malicious profile is full of sensitive personal information that could be exploited later in our real life for burglary, blackmail, or mugging for examples. Moreover, in light of existing advanced methods that could be used to track users, such as [10, 11], gathering private information has become more serious. Furthermore and according to the data units included in Table 1, the attacker can apply query analyzing-based attacks on the sent queries, obtaining more auxiliary information to be added to the previous malicious profile. Beyond that, these two ways (i.e., tracking the locations and analyzing

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

the queries of the LBS user) could be used by the LBS server itself or its maintainer (i.e., acting as a malicious party or an attacker), which in turn aggravates the risks since all information related to the LBS users are stored and accessible by the LBS server. In addition, the two mentioned ways lead to two different types of privacy which are location privacy and query privacy. So, if a full privacy protection is needed in the LBS field, the two types of privacy must be ensured. Besides the need of protecting location privacy and query privacy, many issues are related to privacy protection in LBS field. Inference attacks can be applied to weaken the privacy protection methods. Inference attacks have a high negative impact on the protection methods. However, there is no study, presented previously to the best of our knowledge, that explores the negative of inference attacks on the LBS privacy protection approaches. Moreover, in the light of existing various privacy metrics, it is worth to differentiate between the standards privacy metrics and those specialized for a certain privacy protection approaches. Furthermore, highlighting the challenges in the LBS privacy protection field enables the researchers to focus on the presenting the corresponding solution to such challenges. However, the challenges differ from one privacy protection technique to another. Classifying the privacy protection approaches in a standard categories and highlighting the advantages, disadvantages, and challenges or open problems related to each category enables the researchers to accurately determine their objectives in their works.

In this paper, we introduce a survey related to the different approaches proposed previously to protect the privacy of the LBS users. The contributions of our work are as follows:

- We introduce a novel classification for the proposed privacy protection in LBS field. Our classification depends on the amount of collaboration between the two major components involved in any LBS privacy protection system (i.e., LBS users and LBS server), highlighting both the power points and the drawbacks of each category.
- We analyze, evaluate, and rank the most advanced inference attacks depending on the negative impact on the LBS privacy protection approaches.
- We explore the various privacy metrics used to quantify the privacy in LBS privacy protection field.
- We provide some rules-based recommendations that can help the LBS users to select the best protection method to ensure their privacy.
- We highlight the challenges in LBS privacy protection area, which help the researchers to focus on in the future works.

The rest of the paper is structured as follows: Section 2 introduces our new classification, where a common system model that matches most approaches described in the literature is presented. In section 3, we present the different protection goals from the LBS user's point of view supported by examples. Section 4 explores the concepts of the different metrics used to measure the privacy. The most advanced attacks that LBS privacy protection approaches are suffering from are presented in section 5. The challenges and recommendations are provided in section 6 and 7 respectively. Finally, we conclude the paper and give some future works in section 8.

2. CLASSIFICATION OF LBS PRIVACY PROTECTION APPROACHES

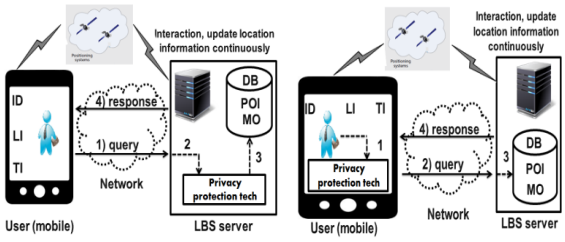
Many efforts were made to classify the proposed privacy protection approaches, and these classifications were taken from different points of view according to their objectives [79], topologies of network [80], or structure features [81]. The previous classifications groups the privacy protection approaches in a two main categories which are server-based approaches and user-based approaches. However, our classification differs from the previous ones in two aspects which are i) providing a common model which determines where the privacy protection method is installed and ii) a third category is added which depends on the amount of collaboration between the LBS server and the LBS user. These two aspects an excellent and positive effect on the researches to understand the art of LBS privacy protection. In addition, having a third category grants the researches an additional flexibility in exploring the LBS privacy protection approaches for comparison purpose.

In this section, we introduce a common system architecture that any LBS privacy protection system should have, determining where the privacy protection method is installed. Then, we present our new classification, discussing the provided privacy protection approaches in detail.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

2.1 System Model

The system model provided in this paper consists of two main components. They are the mobile device of LBS user/users and LBS server, where the mobile device of the user is equipped with an integrated position sensor to determine the current user position. Mobile devices send their position information to the LBS server, which stores and manages mobile device positions on behalf of the user. According to the two main components, the protection mechanism is installed on the mobile device of LBS user or on the LBS server as shown in Figure 3. It should be mentioned that the robust privacy protection mechanism aims at protecting three goals which are: The identity of the LBS user (ID), the location information of the LBS user (LI), and the temporal information of the LBS user (TI). The next section is dedicated to discussing these goals in details.



(a)server-based model. *(b)user-based model.*

Figure 3: The Common Model of LBS Privacy Protection Systems.

Regarding server-based system model, the LBS server is considered a trusted server. So, the privacy protection technique is installed on the LBS server, and the process goes through four steps as follows: 1) The LBS user issues a non-protected query; 2) the units of the received query, mentioned in Table 1, from the inputs of the protection technique where it is executed on the LBS server side; 3) to manipulate the received query, the required data is obtained from the database stored on the LBS server side, where the information about the POIs and the moving objects (MO) (i.e., LBS users' locations and their motion trajectories) is stored and continuously updated in the database and 4) after the query manipulation, the results are returned to the LBS user.

As for the user-based system model, the LBS server is considered untrusted server. So, the privacy protection technique is installed on the mobile device of the LBS user. For the process, it goes through the same previous four steps, except that the execution of the protection technique is performed on the mobile device of the LBS user. So, the issued query will be protected (i.e., to ensure privacy protection) before sending to the LBS server.

Another system model is derived from the common system model above, where a trusted third party (TTP) is used between the LBS user and the LBS server as shown in Figure 4.

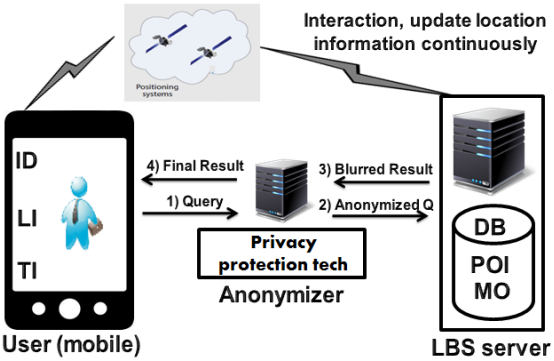


Figure 4: A TTP-based System Model of LBS Privacy Protection.

In a TTP-based system model, the LBS server is considered untrusted server and the privacy protection technique is installed and performed on the TTP side. This TTP is represented either by a special trusted hardware-based CPU, connected to the LBS server or by a trusted anonymizer. The CPU preserves the privacy using private information retrieval (PIR) such as [12, 13]. As for the trusted anonymizer, it blurs the real position of the user before sending the query to the LBS server such as [14, 15, 16, 17].

2.2 Classification of LBS Privacy Protection Approaches

In our classification, we introduce the principle of the amount of collaboration between the LBS user/users and the LBS server as the main base to distinguish among the privacy protection approaches. Since the TTP mentioned in Figure 4 is tightly-coupled with the LBS server, the majority of the load will be on the LBS server side. It, in turn, forms the first category of the classification (i.e., the most load on the LBS server side or server-based approaches). In contrast, the user-based approaches state that the majority of the load is on the mobile device of the LBS user, which forms the second category. However, there is a chance to collaborate among the LBS users themselves to preserve their privacy, which forms the third category. Figure 5 gives a general look at the three categories, where each category has its own techniques.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

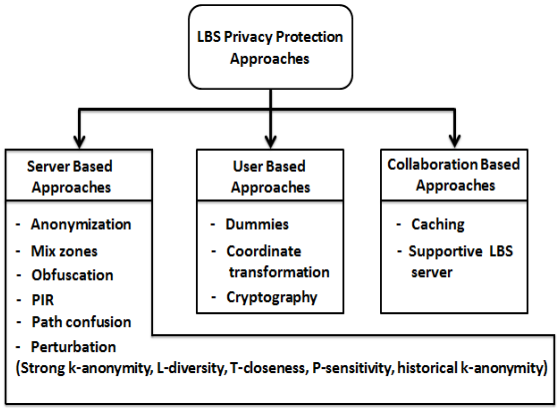


Figure 5: Classification of LBS Privacy Protection Approaches.

Before discussing the details of each category, it is worth to mention that most of LBS privacy protection approaches share the same general concept, which is applying K-anonymity concept in different ways. Originally, K-anonymity concept was provided by Gruteser and Grunwald [18]. The key idea is that the user deliberately sends a perturbed area rather than his/her accurate location so that it is hard to determine the real location of the LBS user among k-1 other users' locations.

2.2.1 First group

(Most of the load on the server side): Means that the LBS server is the main responsible for the execution of protection method, while the mission of the LBS user is only sending a query. So, the shining assumption in this category is that the LBS server is trusted the server. Various techniques are used in this category as shown in Figure 5.

In their work [19], Gedlik et al. present a personalized K-anonymity approach, where LBS server acts as an ***anonymizer.*** This approach adopts with conditions provided by the user (i.e., to protect the privacy), where a spatial-temporal mask is applied on the position of the user, providing k-anonymity level of tolerance that the user wishes. Depending on the same idea, [20] suggested personalization according to the user profile which contains the conditions of privacy protection. One of the most popular techniques used in this group was proposed by Beresford et al. in [21] called ***mix zones.*** The users located in an area will be grouped into many spatial regions. This region protects the real positions for the users through hiding them within such regions. Then, these regions will be mixed together, where no location updates inside a mixing zone occur during the motion of the objects. The work [22] improved the mix zones approach through adding pseudonym concept. Therefore, another condition is satisfied, which is the user must utilize another pseudonym when leaving one mix zone to another. Another development was performed on mix zones, where the authors of [23] proposed the MobiMix approach. The essence of the development idea is to the make mix zones approach more robust against the attackers. To end this, the authors took into consideration various context information that can be exploited to derive detailed trajectories such as geometrical and temporal constraints.

Spatial ***obfuscation*** techniques try to protect the privacy by minimizing the accuracy of the location information sent by the LBS user to the LBS server. A classic spatial obfuscation approach is provided in [24], where a user sends a circular area instead of the accurate user position. Depending on the same idea, the work [25] presented a new approach. The difference was: instead of using geometric obfuscation shapes (i.e., circles), the authors use obfuscation graphs to apply the concept of position obfuscation to road networks. The obfuscation technique is developed in [26] to present robustness against semantic location attack, where the location of the user cannot be mapped with a high probability to certain critical locations such as a hospital. Therefore, a map-aware obfuscation approach was proposed, where the key idea is expanding the obfuscation area adaptively in a way that the probability of the user for being in a certain semantic location is below a given threshold. ***Cloaking region*** is a protection method inspired from the obfuscation technique. The key idea is to cloak the real location of the user in spatial and temporal domains. To protect the privacy, the authors of the work [27] played on the resolution of the cloaking region through modifying the spatial-temporal dimensions, satisfying certain conditions to achieve a high k-anonymity level. Using the cloaking region method, the authors of [28] manipulated the problem of applying a constant level of privacy protection (i.e., k = constant to achieve k-anonymity concept), where this constant level may be out of the user feeling and do not be needed. So, they allowed the user to express the privacy level he wishes so that the user can minimize the resolution of the cloaking region in the regions that the user feels relax and maximize it in other regions. A hierarchical grouping algorithm integrated with cloaking region was proposed in [29]. To ensure the privacy protection of the users, the hierarchical grouping algorithm groups the users in different sets, then the cloaking region method is applied to the orders of the users (i.e., their queries when asking for POIs), and finally the hierarchical

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

grouping algorithm collects the orders in each group, sending them together to the LBS server. This makes the attacker confused about determining the real locations of the users. In [30], the LBS server acts as a location hider to camouflage the actual position of the user. The basic idea is to exploit the landmarks located in the area the user resides, hiding the real position of the user in a one landmark such as a university or sports city. In the case of no landmark is located, the LBS server creates an imaginative landmark depending on the information stored previously about the successful tries. Similar to [30], [31] exploited the geographic context of the area where the user is located to build landmarks. The difference was that [31] deals with moving objects, avoiding creating imaginative landmarks and taking into account that the motion of the objects can be exploited to find effective landmarks.

Private Information Retrieval (PIR) protocol was proposed in [32] to retrieve POIs queried by the user. The strategy followed by the authors is that instead of determining his real position, the user defines an index through the LBS provider. Depending on processing this index, the LBS provider executes the PIR protocol to extract the corresponding POI with an encryption stage. Another PIR-based approach was developed in [33], where a combination between the concept of ϵ-differential privacy and PIR is performed to ensure getting the same amount of the information that represents the query response. The key idea that the authors used is to rely on the statistics of the queries to retrieve a similar heap of information for each query, where it could be employed to weaken the ability of the attacker who tries to obtain private information.

Path confusion method was introduced in [34]. It targeted the confusion of the attacker and minimized his ability about gathering historical information related to the trajectories of the users during their motion. This approach used a mixing algorithm, to be executed at the server side, where paths intersections are exploited in the region where at least two users are meeting. Meyerowitz et al. [35] used the same idea mentioned in [33] with an enhancement using both cache clock and position prediction.

Perturbation techniques rely on the basic concept of k-anonymity, where this concept is extended by various approaches to increase privacy protection, such as strong k-anonymity, L-diversity, T-closeness, P-sensitivity, and historical k-anonymity. Under reciprocity term, the authors of [36], [37], and [38] guarantee strong k-anonymity by ensuring that the calculated cluster of k users remains the same over several queries. The key idea was to use adaptive nearest neighborhood cloaking to achieve this term. The idea of location L-diversity was presented by Bamba et al. [39]. L-diversity approach aims at making the location of the user indistinguishable from a set of L different physical locations such as hospitals, clinics, churches, sports clubs, etc. To achieve this, the approach guarantees that the position of the k-cluster members is not just different, but is also located distant enough from each other. The concept of l-diversity developed by Li et al. [40], where they proposed T-closeness approach. Here, the parameter T represents the distance between an attribute's distribution within the selected cluster of k users and the same attribute's distribution over the total set of user, where this distance should be higher than a certain threshold T. The guarantees of k-anonymity was enhanced through proposing P-sensitivity concept [41]. The key idea of the P-sensitivity concept is to ensure that within a k-cluster, each group of highlighted key attributes has at least p distinct values for each highlighted attribute within the same group. Another approach extended the k-anonymity concept taking into consideration the temporal dimension of the user's location information. Mascetti et al. [42] presented an approach called historical k-anonymity for MOs. Similarly to strong k-anonymity clustering, historical information of multiple users is divided into blocks, where each block includes locations of at least k users. CliqueCloak [43] is a personalized k-anonymity model in which users can control and adjust their minimum level of anonymity, and the maximum temporal and spatial resolutions they can tolerate. This approach modeled the anonymization constraints as a constraint graph and thus transformed the problem of finding cloaking boxes into that of finding cliques that satisfy certain conditions in the constraint graph.

Regarding the advantages and disadvantages of this category, many issues are taken into account to be checked, which are: 1) Technical issues; 2) quality attributes that should be satisfied in any LBS privacy protection system; 3) mission of the LBS server. Table 2 summarizes the mentioned issues, where ○ symbol means taking an advantage and ● means suffering from a disadvantage.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 2: Advantages and Disadvantages of Server-based Privacy Protection Approaches.

Issue \ Term		Advantages	Disadvantages
Technical issues	Have control		•
	High storage	○	
	High computation	○	
	Short life battery	○	
Quality attributes	Network overhead		•
	Scalability	○	
	Availability	○	
	Reliability	○	
	Portability		•
	Performance	○	
	Maintenance	○	
LBS server	Malicious party		•

The main drawback of server-based approaches is that the LBS server (or its maintainer) can act as a malicious party (i.e., an attacker). The reason behind this is that all information related to the LBS users, their trajectories motion, and the details of POIs they prefer are stored in the LBS server and can be accessible by an attacker. This fundamental drawback formed the major motivation of the researchers to turn to the second group.

2.2.2 Second group

(Most of the load on the user's mobile side): Means that the mobile device is the main responsible for protection method execution, where the LBS user mission is to set some parameters as inputs and related to the protection method. The shining assumption in this category is that the LBS server is a untrusted server.

In the work [44], Yanagisawa et al. provided ***dummies*** idea to protect the privacy of the LBS user. The key idea was that the user creates many of false positions (dummies), building instances of the current query using both the dummies and the true position of the user, and then sending all of the copies to the LBS server asking for the same POI. Randomizing the real position among dummies ensures the privacy protection, where the LBS server cannot recognize the real position among dummies. Similarly, [45] uses dummies to protect the privacy protection of the LBS users. It depends on selecting the dummy using normalized distance to confuse the attacker and limit his/her ability to track or infer some sensitive information about the query issuer (i.e., the LBS user). Another approach used dummies idea was presented in [46] called DUMMY-Q, but the idea is applied to the query itself rather than the location. Therefore, dummy queries of different attributes from the same location are generated to hide the real query. To make the generated dummied stronger, two aspects are taken into consideration which is 1) The query context; 2) the motion model. Hara et al. [47] developed a dummy-based approach, manipulating dummies generation from our real life. So, they took into account the physical constraints of the real world. The feature that distinguishes this work was that the trajectories of the generated dummies cross the trajectories of the actual movement of the LBS user.

Gutscher et al. propose the idea of ***coordinate transformation*** [48], where the users apply some geometric operations, such as shifting or rotating, over their locations before sending them to the LBS server. In order to retrieve the original locations, inverses transformation function is used. Similar to [48], the work [49] proposed a solution that allows the user to protect his/her real position using mathematical operations. These mathematical operations include enlarging radius, shifting center, increasing the radius, or applying double obfuscation (i.e., mixing shifting center with any of remainders).

Cryptographic privacy approaches utilize encryption to protect the locations of the users. Mascetti et al. [50] proposed an approach to notify users when friends (also called buddies) are within their proximity without revealing the current location of the user to the LBS server. To achieve this, the authors assume that each user shares a secret with each of his buddies and use symmetric encryption techniques. Another approach was provided in [51], manipulating the problem of dealing with untrusted LBS server. The authors based on the distributed management of position information using the concept of secret sharing. The key idea of this approach is to partition the location information of the user into shares. Then, the shares are distributed among a set of untrusted LBS servers. In order to recover positions, the user needs the shares from multiple LBS servers.

For the advantages and disadvantages of the user-based category, Table 3 summarizes the same issues taken into consideration and related to the previous category.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 3: Advantages and Disadvantages of User-based Privacy Protection Approaches.

Issue \ Term		Advantages	Disadvantages
Technical issues	Have control	○	
	High storage		●
	High computation		●
	Short life battery		●
Quality attributes	Network overhead	○	
	Scalability		●
	Availability		●
	Reliability		●
	Portability		
	Performance		●
	Maintenance		●
LBS server	Malicious party	○	

The short life battery drawback, which in turn leads to a poor availability and reliability quality attributes formed the fundamental motivation behind turning the researchers to the third group. That is because this drawback results in a disconnecting problem, which limits the ability of the LBS users from tacking the benefits of LBS-enabled applications.

2.2.3 Third group

(Load balancing between the LBS users and LBS server): Means that the LBS user controls and ensures his/her privacy protection with a help provided by the LBS server. Similar to the previous group, the LBS server is considered untrusted or malicious party.

A ***cache-based*** approach was proposed in [52]. The key idea is exploiting the collaboration among the LBS users to avoid dealing with the LBS server. The privacy protection is achieved by answering the queries within the mobile crowd, where the queries responses are stored in the cache of each mobile device at each user. Similar to [52], the work [53] used the cache integrated with dummies idea, where Enhanced Dummy Selection Algorithm (enhanced-DSA) is proposed. The basic idea of the enhanced-DSA algorithm is to generate dummy locations, and the answers of the dummies are stored in the cache to answer future queries. The idea enhanced-DSA algorithm was developed by Niu et al. [54], proposing Caching-aware Dummy Selection Algorithm (CaDSA). Here, CaDSA algorithm generates dummies using normalized distance, and the cache is represented by the access point. In this approach, the LBS user who issues a query can obtain his/her query answer through one of two ways: 1) Searching in the cache; 2) if the query answer is not found in the cache the LBS user is forced to connect to the LBS server. To enhance the probability of finding the query answer in the cache, a data freshness function is used.

A ***privacy-supportive LBS server*** structure was proposed in [55], helping the user to make his own privacy decision. The basic idea depends on constructing an LBS server structure, which provides auxiliary information to the user to support his privacy decision so that he will be aware of risks about his achieved privacy level. This helped the user to create his/her queries carefully. Authors of [56] developed point-to-point access along with building an air index (NPI) list within the connecting channel. The mission of the server is to index the data segments before broadcasting. The indexing data carry information about the cells of the region the users are located within, where the periodic transmission of the indexed data ensures privacy protection.

The main advantage of collaboration-based approaches is minimizing the number of queries sent to the LBS server, which in turn reducing the dependency on the LBS server due to the reduced number of connections. However, the approaches presented in this category still depend on the mission provided by the LBS server. In the worst case, the LBS users are forced to rely on the LBS server totally, which means going back to the disadvantages of server-based approaches group mentioned in Table 2.

3. PROTECTION GOALS FROM LBS USER PERSPECTIVE

In this section, we provide the protection goals that the LBS user concerns about and could be misused to harm his/her privacy. Three main protection goals are presented which are LBS user identity (ID), location information (LI), and temporal information (TI). Various scenarios supported with examples are explored to highlight the importance of protecting the previous three protection goals. In addition, we determine which one of the three protection goals is achieved in the approaches presented in the previous section.

3.1 LBS User Identity

The identity of a user can be his/her name, a unique identifier, or any set of properties uniquely identifying the user. Upon this, the user aims at hiding his/her identity for privacy protection purpose. As an application scenario, consider a user of an advanced navigation system providing real-time traffic information and points of interest

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

information based on the current location of the user. As a decision maker in a bank committee or a member of a politic party, it is important for the user to protect his/her identity. That is because revealing such sensitive data can be misused and exploited by an attacker for extortion as an example, which in turn affects the decision making or may threaten the life.

3.2 Spatial Information

Spatial information refers to the accurate location of the user where the query is issued from, or to the locations of the POIs the user searches for. Tracking these uncovered locations allows the attacker to collect a personal information about the user, constructing a malicious profile, and then launching the actual attack (such as stealing, mugging, or blackmailing) against the user based on the content of the maliciously constructed profile. Presenting a historical travel as an evidence in front of a court based on the tracked locations of the user may lead to a loss of the case, putting a person for many years in a prison. This simple scenario shows the importance of protecting the spatial information of the user.

3.3 Temporal Information

Temporal information refers to the point in time or time period when the spatial information of the user is valid. In some scenarios, spatial information is only considered critical if it is associated with the temporal information. For example, consider a traveling user on a speed road, where the trajectory of the user is tracked by an attacker. Revealing the temporal information, allows the attacker to calculate the speed easily. This personal information (i.e., the speed) could be sent to the insurance company, forcing the user to pay a penalty of exceeding the limited speed.

Table 4 below shows a summary of the approaches discussed above in terms of both achieving ID, LI, and TI protection goals and achieving full privacy protection.

Discussion. Table 4 shows that most of the privacy protection approaches in LBS field belong to server-based approaches class. The user-based approaches class comes in the second-ranking according to the number of privacy protection approaches. Compared to the two previous classes, collaboration-based approaches class includes a very few of proposed approaches, which reflects the lack of researches presented in this class. In addition, very few approaches achieved a full privacy protection. Under satisfying the full protection goals term, there are gaps in the most of the checked approaches. Dummies-based and mix zones-based approaches are considered sophisticated methods to satisfy the whole protection goals. In spite of the lack of the approaches provided in collaboration-based class, one of the proposed approaches achieved a full privacy protection and guaranteed satisfying the three protection goals (i.e., ref [52]). The idea of integrating between the cache-based technique and dummies-based technique is one of the promising methods to ensure a full privacy protection such as [53] and [54]. That is because 1) It avoids dealing with untrusted LBS server (the main drawback of server-based class); 2) it gives the user a complete control on the privacy protection method (the main benefit of user-based class); 3) it satisfies the whole protection goals besides to a full privacy protection as shown in [52].

After exploring the different classes of the privacy protection approaches proposed in LBS field and identifying the protection goals, we need to have a look at the privacy metrics used to quantify the privacy. The next section provides the details of the used privacy metrics.

4. PRIVACY METRICS

Various privacy metrics have been proposed in the LBS privacy protection field, where [57] presented a wide spectrum of privacy metrics. The purpose of any privacy metric is estimating how much the privacy of the user is broken by an attacker. In general, since there are two types of privacy, there are two corresponding types of privacy metrics (i.e., location privacy metrics and query privacy metrics). Each privacy metric provided in this paper can be standard or specific to a certain approach. In addition, some of the privacy metrics can be used for both location privacy and query privacy. Figure 6 gives a comprehensive look at the privacy metrics introduced in this section.

4.1 Query Privacy Metrics

Location entropy is the most widely used privacy metric. Originally, this metric is inspired from Shannon's entropy in information theory [58]. It is used to measure the uncertainty associated with location information in LBS queries by quantifying the information an attacker can obtain from one (or a series) of location update(s). Some researchers employed the concept of location entropy, proposing a new privacy metrics. The authors of [28] defined the popularity of a spatial area as 2^E, where E is the location entropy. Accurately, the popularity of the public region is measured using

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 4: Achieving Protection Goals in Privacy Protection Approaches.

Class	Technique / Approaches \ Protection Goals	ID	LI	TI	Full Privacy Protection	
					Location Privacy	Query Privacy
Server-based approaches	Anonymization:					
	- [19]	×	√	√	√	×
	- [20]	√	√	×	×	√
	Mix zones:					
	- [21]	×	√	√	√	×
	- [22]	√	√	√	√	×
	- [23]	×	√	√	√	×
	Obfuscation (Clocking):					
	- [24]	×	√	×	√	×
	- [25]	×	√	×	√	×
	- [26]	×	√	×	√	×
	- [27]	×	√	×	√	×
	- [28]	×	√	√	√	√
	- [29]	√	√	×	√	×
	- [30]	×	√	×	√	×
	- [31]	√	√	×	√	×
	PIR:					
	- [32]	√	√	×	√	√
	- [33]	√	√	×	√	×
	Path confusion:					
	- [34]	×	√	√	√	√
	- [35]	√	√	×	√	×
	Perturbation:					
	- [36]	√	√	×	√	√
	- [37]	√	√	×	√	×
	- [38]	√	√	×	√	×
	- [39]	√	√	×	√	×
	- [40]	√	√	×	√	×
	- [41]	√	√	√	√	×
	- [42]	√	√	×	√	×
	- [43]	×	√	√	√	×
User-based approaches	Dummies:					
	- [44]	×	√	×	√	×
	- [45]	√	√	×	√	×
	- [46]	√	×	√	×	√
	- [47]	√	√	√	√	×
	Coordinate transformation:					
	- [48]	×	√	√	√	×
	- [49]	×	√	√	√	×
	Cryptography:					
	- [50]	×	√	×	√	×
	- [51]	×	√	√	√	×
Collaboration-based approaches	Caching:					
	- [52]	√	√	√	√	√
	- [53]	×	√	×	√	×
	- [54]	√	√	×	√	√
	Supportive LBS server:					
	- [55]	×	√	×	√	×
	- [56]	×	√	×	√	×

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

entropy based on its visitors' footprints inside it. Hoh et al. [59] introduced time to confusion as a privacy metric. Their idea is built on the fact that the degree of privacy risk strongly depends on the duration for which the user can be tracked by an attacker. Based on generated traceable trajectories, they defined linkability term related to the location samples that form a trajectory, taking into consideration a predefined location entropy threshold (called an uncertainty threshold). Then, they computed the time to confusion as the difference between the timestamp of the first location sample and that of the last one.

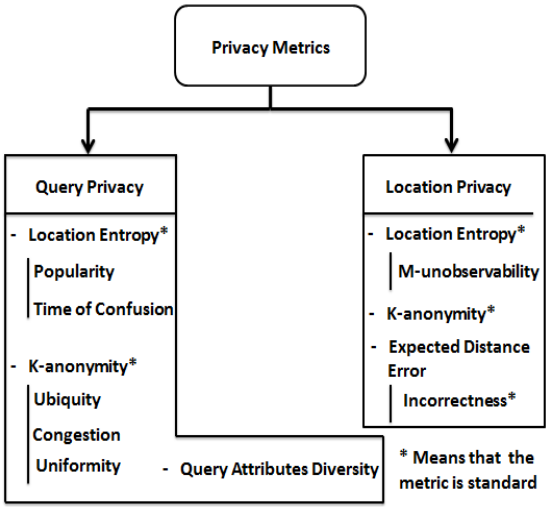


Figure 6: Types of Privacy Metrics.

K-anonymity is another popular metric used for LBS query privacy protection. The concept of k-anonymity was first introduced in the database research field [60] and then quickly became a common privacy metric in LBS privacy protection community. In the context of LBSs, k-anonymity metric refers to the state at which the location information included in an LBS query corresponds to an area where the query issuer is indistinguishable from at least (k–1) other users also present in that area. Depending on generating an enhanced anonymity set, some researchers employed the concept of k-anonymity to develop a new privacy metrics. The key idea exploits the number of users located in an area to express the strength of k-anonymity level. ***Ubiquity*** is introduced when the users are located in the entire area, increasing the location anonymity of the entire area [61]. Limiting ubiquity in a specific area, motivated the authors of [44] to present congestion term as a privacy metric. Congestion means that there is a concentration of users in a local region, increasing the anonymity of that specific region. The authors of [62] provided uniformity term as a privacy metric, which requires that each distributed region contains the same number of users.

Since the attacker can apply some analysis on the sent query, the degree of ***query attributes diversity*** is used as a query privacy metric. This metric is used to measure the confusion of the attacker when trying to infer some sensitive information about the user [46].

4.2 Location Privacy Metrics

Location entropy is used to measure the privacy of the LBS user's location by quantifying the diversity, and therefore difficulty in identifying a user's personal preferences, parameters, and whereabouts [63]. Again, location entropy is adopted by researchers to develop a new privacy metrics. The authors of [64] proposed ***M-unobservability*** privacy metric to measure the uncertainty of associating POIs with a user's positions, where M refers to the POIs that are most likely visited by the user. In more details, the unobservability term refers to the state of Items of Interest (IOIs) being indistinguishable from any IOI at all. In the context of LBSs, IOIs are POIs. Hence a user's location privacy can be defined as unobservability of his/her location.

K-anonymity is also used as a privacy metric for location privacy. Here, the location of the user is indistinguishable among at least k-1 users' locations, not the query issuer.

Expected distance error was introduced in [34] to measure how accurately an attacker can estimate a user's position, taking into account the differences among the locations observed by the attacker. ***Incorrectness*** is a standard location privacy metric inspired from the latter privacy metric. In this metric, the distance is defined as the difference between the attacker's estimation and the true value related to the accurate location of the user [65].

5. INFERENCE ATTACKS AGAINST LBS PRIVACY PROTECTION APPROACHES

In the three previous sections, we navigated wide sets of LBS privacy protection approaches, showing which protection goal is ensured by each discussed approach, followed by the types of privacy metrics used for privacy quantifying. However, there are some tactics that can be used by the attackers to weaken the robustness of the privacy protection approaches. These tactics called inference attacks, and target the protection goals explained previously by using smart ways to derive private information about the users. In this section, we discuss the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

concepts of the most advanced inference attacks, highlighting the negative impact of each one on the protection goals.

5.1 Types of Inference Attacks

In ***Location Homogeneity Attack (LHA)*** [15, 66], the attacker analyzes the positions of all k-cluster members. If their positions are almost identical (Figure. 7a), the position information of each member is revealed. If the cluster members are distributed over a larger area, the position information is protected (Figure. 7b). An advanced location homogeneity attack can utilize map knowledge to reduce the effective area size where users can be located (Figure. 7c). As an example, all the k-cluster members' locations are limited in a hospital or university landmarks.

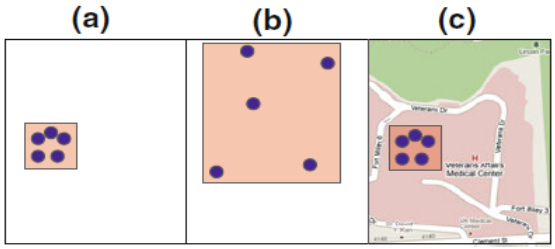


Figure 7: Location Homogeneity Attack.

Another advanced inference attack is called ***Map Matching Attack (MMA)*** [67], where the attacker can use external background knowledge about a user to decrease the privacy. In-depth, this attack is used to restrict the obfuscation area to certain locations where users can be located by removing all the irrelevant areas. For as an example, the map is employed to remove areas such as lakes from the obfuscation region, which effectively shrinks the size of the obfuscated region. This way enables the attacker to define the location of the user as shown in Figure 8.

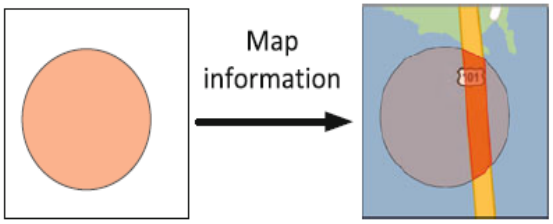


Figure 8: Map Matching Attack.

Query Sampling Attack (QSA) [68, 69, 70] is where the attacker employs the unfair location distribution of the LBS users for his own malicious purpose. This type of inference attack targets isolated users in a sparse region as illustrated in Figure 9. Therefore, it relies on the gathered traffic statistics of the environment where the users are located. In details, the attacker tries to calculate a probability distribution function of the user location over a given area. If the probability is not uniformly distributed, the attacker can determine the areas where the user is located with a high probability.

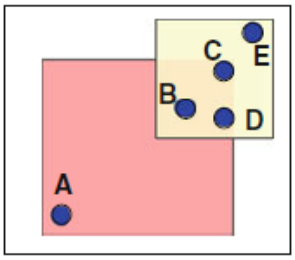


Figure 9: Query Sampling Attack.

Semantic Location Attack (SLA) [71, 72] is where the attacker can infer semantic meanings related to the behavior of the user by exploiting the amount of time a user stays in one place, such as a school, a fast food restaurant, or a cafe. Consequently, the key idea is estimating the probability of the stay duration or usage time in a frequently visited POI by the user as shown in Figure 10.

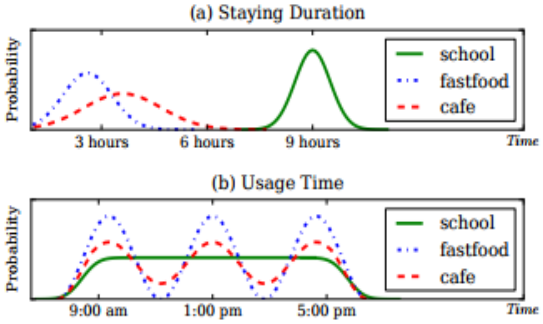


Figure 10: Semantic Location Attack.

5.2 Impact of Inference Attacks in The LBS Privacy Protection Approaches

To show the negative impact of inference attacks on the privacy of LBS users, we considered each inference attack presented in the previous sub section (i.e., LHA, MMA, QSA, and LSA) as the main criteria factor, meanwhile each one of the protection goal (i.e., ID, LI, and TI) is considered as an affected aspect. All approaches contained in Table 4 above are evaluated. Our evaluation relies on three options to measure the negative impact of the criteria factors. Table 5 gives a description of the three used options.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 5: Options of Measurement.

Option	Description
√	When the factor has high negative impact and guides to penetrate the privacy.
×	When the factor has a low negative impact and doesn't guide to penetrate the privacy.
P	When the factor has a partially negative impact on the success of the privacy penetration.

5.3 Analysis and Discussion

Table 6 below can be read horizontally or vertically as illustrated in Figure 11.

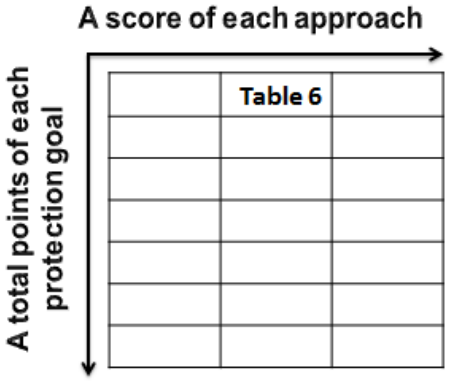


Figure 11: Horizontal and Vertical Reading of Table 6.

If it is read horizontally, then the numbers on the table represent the total points that each approach has got from all of the protection goals for each one of the above three options. Each option has a score varies in the range of [0, 1, 2, 3]. For instance, the corresponding numbers of the approaches [21, 22] proposed under mix zones technique show that (√) option got a score equals (2) points, referring that both ID and LI protection goals are highly suffering from location homogeneity attack. That's because both of them could be determined by an attacker (even if no location updates are performed when moving from one zone to another) in case of all mix zones belong to the same landmark. In addition, (×) option got a score equals (1) point, referring that TI protection goal has a low suffering from location homogeneity attack. The reason behind this is that the attacker can have a chance of observing the time moments at which the location updates are performed (i.e., when the LBS user changed his/her zone passing to another one). Compared to [21, 22], the work [23] proposed under the same technique can have a better protection for both ID and LI protection goals, where (×) and (P) options got scores equal (2) and (1) point respectively. This, in turn, reflects a higher privacy protection level. That is because the approach proposed in [23] is considered as a development of both approaches presented in [21] and [22], where a pseudonym concept is performed taking into consideration the context information of the area where the LBS users are located. Another example of horizontal reading of Table 6 is the approach proposed in ref [56], where the score of (√) option equals (3), reflecting a very high negative effect of location homogeneity attack.

Since the LBS user can choose any technique that belongs to any of the proposed classes, we focused on reading Table 6 vertically with a further statistical analysis and limited the horizontal reading in the two examples discussed above. If Table 6 is read vertically, then the numbers represent the total points that each protection goal has got for each one of the above three options and related to the all approaches provided in the all classes.

From the numbers that appear in Table 6, it can be noticed that the total number of points that the (√) option got is 45 point. These points distribute on the (ID, LI, and TI) protection goals with (10, 23, and 12) values respectively. For (×) option, the protection goals got (19, 3, and 13) values from the total points which are 35, and the corresponding values related to the protection goals for the (P) option are (6, 9, and 10) from the total points which is 25. Figure 12 shows the negative effect percentage on the protection goals according to the three options.

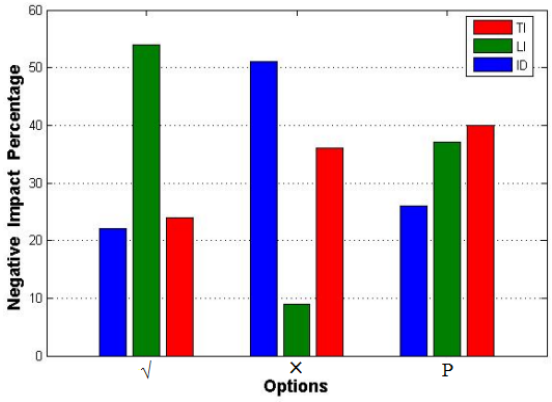


Figure 12: The Negative Impact of LHA on The Protection Goals.

Among the three protection goals, it is obvious that LI protection goal is the most negatively affected one by LHA, meanwhile ID protection goal is the lowest one. That's because the attacker mainly depends on the contextual location information to break the privacy protection approach. In addition,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 6: Effect of Location Homogeneity Inference Attack.

Class	Technique/Approaches \ Protection Goals	Location homogeneity attack			Sub Totals		
		ID	LI	TI	√	×	***P***
Server-based approaches	Anonymization:						
	- [19]	×	√	P	1	1	1
	- [20]	√	√	P	2	0	1
	Mix zones:						
	- [21], [22]	√	√	×	2	1	0
	- [23]	×	P	×	0	2	1
	Obfuscation (Clocking):						
	- [24],[25]	P	√	√	2	0	1
	- [26]	√	×	×	1	2	0
	- [27]	P	√	×	1	1	1
	- [28]	P	√	√	2	0	1
	- [29]	×	P	√	1	1	1
	- [30]	√	√	×	2	1	0
	- [31]	√	√	×	2	1	0
	PIR:						
	- [32], [33]	×	√	×	1	2	0
	Path confusion:						
	- [34]	√	√	P	2	0	1
	- [35]	√	×	×	1	2	0
	Perturbation:						
	- [36]	×	√	P	1	1	1
	- [37]	×	√	P	1	1	1
	- [38]	×	√	P	1	1	1
	- [39]	P	P	√	1	0	2
	- [40]	P	×	√	1	1	1
	- [41]	×	√	P	1	1	1
	- [42]	×	P	√	1	1	1
	- [43]	×	√	×	1	2	0
User-based approaches	Dummies:						
	- [44]	×	√	√	2	1	0
	- [45]	×	P	√	1	1	1
	- [46]	×	√	√	2	1	
	- [47]	×	P	P	0	1	2
	Coordinate transformation:						
	- [48]	√	√	×	2	1	0
	- [49]	√	P	×	1	1	1
	Cryptography:						
	- [50]	×	√	×	1	2	0
	- [51]	×	P	×	0	2	1
Collaboration-based approaches	Caching:						
	- [52]	×	P	P	0	2	1
	- [53]	×	√	√	2	1	0
	- [54]	×	√	P	1	1	1
	Supportive LBS server:						
	- [55]	P	√	√	2	0	1
	- [56]	√	√	√	3	0	0
Totals:							
√	high negative impact	10	23	12	45		
×	low negative impact	19	3	13	35		
P	partially negative impact	6	9	10	25		
Total:					105		

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

TI protection goal pays attention to be protected against LHA by a considerable negative percentage impact.

Our previous results are more supported when it comes to calculating the percentages achieved by each one of the three options. Table 6 shows that the (√) option got 45 points from the total number which is 105 with a percentage equals around 43%. For (×) and (P) options, they got 35 and 25 points from 105 points respectively. Figure 13 illustrates the percentage achieved by each option over all the protection goals.

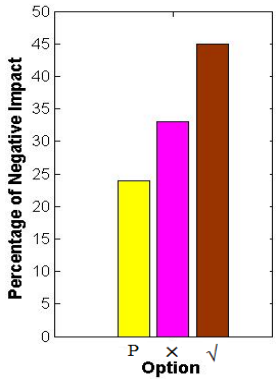


Figure 13: The Negative Impact Percentage of The Three Options on all Protection Goals.

Regarding MMA, QSA, and ALA attacks, we rebuilt Table 6, scanned it vertically, and find out the percentages as shown in Figure 14 (a, b, and c) respectively.

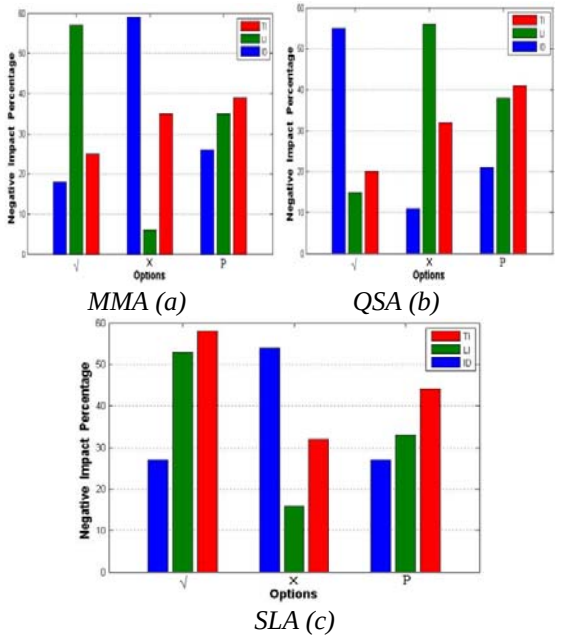


Figure 14: The Negative Impact of MMA, QSA, and SLA on The Protection Goals.

Similar to LHA attack, Figure 14.a shows that LI protection goal is the most affected one by MMA attack. That's because the attacker tries to exploit his/her knowledge (in both LHA and MMA inference attacks) to increase the probability of existing a user in a certain location, or to filter some areas where the existing of the user is impossible.

Figure 14.b shows that the majority of the negative impact of QSA attack is related to ID protection goal. That's because the process of isolation a user from the others enables the attacker to focus on a specific user. Notice that compared to MMA attack, QSA attack got a higher partial negative percentage on TI and LI protection goals due to the same reason (i.e., isolation process).

For SLA attack, Figure 14.c illustrates that both LI and TI protection goals are very highly affected with percentages around 50 and 58 respectively. That's because of the nature of this attack, where the attacker observes and focus on the (time usage) spent by a certain user within a (specific area). Moreover and compared to LHA and MMA attacks, SLA attack includes both of them since the attacker can employ his/her intuition to minimize the area under observed time usage term. So, the percentages related to (P) option were very close to each other for LHA, MMA, and SLA inference attacks.

Since the major concern is related to (√) option (high negative impact), we rank the strength of the discussed inference attacks according to the percentages that (√) option got over the protection goals. Table 7 summarizes the ranking.

Table 7: Ranking of Inference Attacks According to The Strength.

Protection goal / Inference attack	ID	LI	TI	Ranking
SLA	27%	53%	58%	1
MMA	18%	57%	25%	2
LHA	22%	51%	26%	3
QSA	55%	15%	20%	4

6. CHALLENGES IN LBS PRIVACY PROTECTION FIELD

In this section, we shed light on the challenges that should be taken into account by the researchers in privacy protection of LBS area. These future challenges are represented by valuable research questions, which can inspire and guide the researchers in their future works.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

6.1 Achieving a Comprehensive Privacy Protection

We have mentioned that in LBS privacy protection field researchers deal with two types of privacy protection, which are location privacy and query privacy. To achieve a full privacy protection, the two previous types of privacy must be protected. When it comes to talking about query privacy protection type, query privacy must be guaranteed during three phases (i.e., query sending, query processing, and query responding). Achieving location privacy protection with query privacy protection during the three previous phases leads to a comprehensive privacy protection. It is worth to mention that achieving a comprehensive privacy protection in LBS field was not previously provided to the best of our knowledge. So, the issue can be represented by the following research question: ***"How to achieve a comprehensive privacy protection, taking into consideration the query privacy protection during sending, processing, and responding?".***

6.2 Robustness Against Inference Attacks

In the previous section, we explored the most advanced inference attacks that could be used by an attacker to weaken the privacy protection technique, addressing the negative impact severally. However, the attacker can use a mixture of those inference attacks, scaling up his/her ability to hijack the privacy of LBS users. So, ***"how to make defenses against a mixture of advanced inference attacks become a very high challenge".***

6.3 Privacy Quantifying

It is important to evaluate the efficiency of an approach that targets to protect the privacy. Without a privacy metric, it is difficult to measure the privacy protection approach. Despite its importance, there is a lack in proposing a standard privacy metrics, where most approaches have their own way of measuring or quantifying the privacy as shown in Figure 6 above. Existing such standard privacy metrics enables researchers to compare different privacy protection approaches with each other in a sufficient way. So, the research question related to privacy quantifying can be provided as follows: ***"How to quantify the privacy by a standard privacy metric that suits different privacy protection approaches?".***

6.4 Trust Management of Collaboration-based Privacy Protection Approaches

Compared to server-based and user-based privacy protection approaches, a few researchers have just begun to investigate collaboration-based privacy protection approaches. That's because of two main issues, which are:

1. Since LBS users can help each other's to answer their queries during a collaboration session (avoiding dealing with untrusted LBS server), there are no guarantees that prevent any LBS user to convert into an attacker.
2. No robust base of trust is provided by group members to ensure the success of the session.

So, the corresponding research question that combines the two above points is: ***"How to build a robust trust base that guarantees the honor or credibility of the LBS users in collaboration-based approaches?".***

6.5 Open Problems Related to User-based Approaches

Besides the inference attacks that can weaken the privacy approaches of both user-based and server-based classes, user-based approaches suffer from additional critical issues.

Dummies generation is considered an open problem in dummy-based approaches. The underlying reason behind this is that generating weak dummies can present the LBS user as an easy victim, where weak dummies can easily filter by the attackers. So, ***"how to generate strong dummies to ensure a high level of privacy protection"*** is considered one of the most critical research questions on user-based approaches side.

Since the privacy protection method is installed on the mobile device of the LBS user, it is very important to take into consideration, the different platforms used by LBS users. In other words, ***"how to design a privacy protection system that can be installed in the mobile devices regardless the used operating system"*** is an another research question in LBS privacy protection field. This research question, referred by portability quality attribute presented in Table 3 above, was not answered to the best of our knowledge.

6.6 Moving Objects (MO) and Manipulating K-Nearest Neighbor (K-NN) Queries

In the context of mobile computing, K-NN queries concept means that the LBS user can send a query asking for a set POIs that are limited within a certain range of spatial domain. For instance, retrieve the four nearest petrol station that is limited within 2 KM range from my real location. In the previous query, the query issuer is a MO, while the queried POIs are stationary ones. In more complexity, manipulation of the previous query

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

needs accuracy and efficiency of the retrieved results when dealing with MOs and Moving Queries (MQs) [73, 74, 75]. That's because of the continuous location updates of both MOs and MQs. For instance, as a MO, retrieve the friendly helicopters that are expected to enter the region within the next 10 minutes. This issue is clearly highlighted when it is applied to roads network. However, mobile users demand not only accurate and efficient results but results that do not threaten their privacy. According to this demand, ***"how to process K-NN queries with guarantees related to accuracy, efficiency, and privacy"*** is an important research question that must be answered.

6.7 Mobility Modeling and Uncertainty Term

Mobility modeling is tightly-linked when it comes to manipulate and answer K-NN queries in an efficient way. To show the importance of mobility modeling, we need to explain uncertainty term first. In the field of LBS, uncertainty term refers to the amount of uncertainty in identifying the real position of the user by an attacker, so the attacker always works hard to minimize this uncertainty [76]. However, dealing with MOs and MQs leads to another meaning of uncertainty term and related to the accurate answer of the query which will be returned to the query issuer (i.e., the LBS user himself/herself) [71, 77]. In other words, if we have a MO that issued an MQ asking for another MO, the real position of the queried MO is changed during sending, processing, and answering the query. That is because of the continuous motion of the object in the real-time (temporal and spatial domain). As a result, the query issuer will receive unmatched value to the actual real position of the queried MO. So, ***"how to Propose an effective mobility modeling that contributes to enhancing uncertainty real-time issue is a problematic aspect in LBS field".***

7. RULE-BASED RECOMMENDATIONS FOR LBS USERS

In this section, we provide some rule-based recommendations that can help LBS users to select the best way to protect their privacy. The provided rules are inspired by our real life. One of these rules is related to the surrounding that the LBS resides, while the rest of the rules are related to the LBS user himself/herself.

7.1 Intensity-based Rule

The first rule relies on the intensity of the people who interact with the LBS user in his/her society such as a workplace, university, or any daily-life social activity. The rule states that ***"if (the intensity of the people surrounded the LBS user is high), it will be better to choose one of the approaches from collaboration-based class"*** such as hiding in the mobile crowd [52]. That's because the LBS user can take the benefit of avoiding dealing with untrusted LBS server. On another hand, the LBS users that are located in the same place mostly have the same tends about the kinds of POIs they search for. This, in turn, increases the probability of existing the answers of the future queries in the mobile devices caches.

7.1 Frequent Usage-based Rules

The rest of the rules is strongly-linked with the usage of LBS-enabled applications. In the study presented in [78], the authors provided a statistical information about the usage of LBS-enabled applications by the users, using a web-based survey. We took the benefit of this statistical information to classify the LBS users into four main categories as shown in Figure 15.

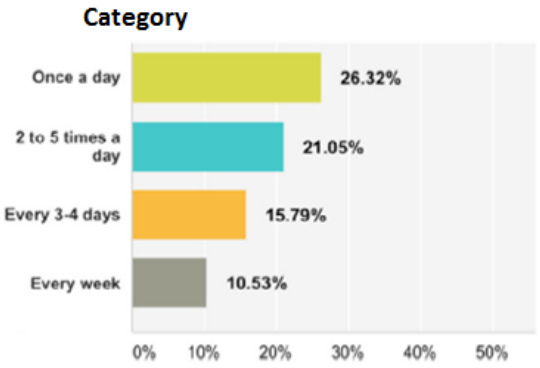


Figure 15: The Usage Frequency of LBS-enabled Applications.

Based on both the usage frequency and the performance, the second rule states that ***"if the LBS user (belongs to the last two categories) and (concerns about the speed of the query response), it will be better to choose one of the approaches from server-based class".*** That's because the server-based approaches perform better than user-based approaches due to the computation capabilities of the LBS server. Moreover, since the usage frequency is very low, the LBS user will be considered as an idle user in the eyes of the attackers. So, the privacy will be far away from attacks. In regards to performance also, the third rule states that ***"if the LBS user (belongs to the first two categories) and (does not concern about the speed of the query response) and (aware of privacy attacking risks), it will be better to choose one of the approaches from user-based class".*** For

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

an example, cryptography-based approaches such as [50, 51] are perfect in this case.

Based on usage frequency, having some control on the privacy protection mechanism, and performance, the forth rule states that ***"if the LBS user (belongs to the first category) and (aware of privacy attacking risks) and (concerns about the speed of the query response), it will be better to choose one of the approaches from server-based class".*** That's because the user can provide strict conditions to ensure a high level of privacy protection. An anonymization-based technique such as [19, 20] is perfect in this case.

Based on both the usage frequency and having a full control on the privacy protection mechanism, the fifth rule states that ***"if the LBS user (belongs to the second category) and (aware of privacy attacking risks), it will be better to choose one of the approaches from user-based class".*** Dummies-based approaches such as [45, 47] are perfect in this case. That's because the LBS user can generate a large number of dummies, protecting the privacy by achieving a high level of K-anonymity.

8. CONCLUSION

When it comes to exploring the privacy issue in LBS research area, we present a survey that studies various LBS privacy protection approaches. A three protection goals (ID, LI, and TI) are defined and checked, according to each presented privacy protection approach, for evaluation purpose. Our study showed that the majority of the addressed approaches did not protect all protection goals. In addition, there is a clear lack in achieving a full privacy protection (location and query privacy protection). To highlight the negative impact of the inference attacks, we studied the concepts of most advanced attacks (LHA, MMA, QSA, and SLA attack) and evaluated them. Our evaluation strategy depended on considering each studied inference attack as a critical factor, and each protection goal as an affected aspect. Our results showed that among the four inference attacks, SLA attack was the most strong one with a negative impact (27% for ID, 53% for LI, and 58% for TI). MMA, LHA, and QSA inference attacks came in second, third, and fourth degree respectively. As a kind of hand help for both researchers and LBS users, we extracted an eight research questions and provided some rule-based recommendations. Our research questions concerned about achieving a comprehensive privacy protection, robustness against a mixture of inference attacks, presenting a standard privacy metrics, and some open problems such as (generating strong dummies and manipulating K-nearest neighbor queries). As for recommendations, we advised LBS users to select specific approaches to ensure a high level of privacy protection, taking into consideration (intensity of people, usage frequency, performance, having control on privacy protection method, and awareness of privacy risks).

In the future work, the research questions, including the open problems such as dummy generation, achieving comprehensive privacy protection, and robustness against mixture of inference attacks, will be answered. In addition, this work will be improved to evaluate the privacy protection approaches from another point of view such as the optimal k-anonymity level that suit each approach.

REFRENCES:

[1] Chang, Victor, Verena Kantere, and Muthu Ramanchadran. "Emerging Services for Internet of Things." (2017).

[2] Leminen, Seppo, Mervi Rajahonka, and Mika Westerlund. "Actors in the Emerging Internet of Things Ecosystems." International Journal of E-Services and Mobile Applications (IJESMA) 9.1 (2017): 57-75.

[3] Ghanbari, Amirhossein, et al. "Business development in the Internet of Things: A matter of vertical cooperation." IEEE Communications Magazine 55.2 (2017): 135-141

[4] Taleb, Tarik, et al. "On Multi-Access Edge Computing: A Survey of the Emerging 5G Network Edge Architecture & Orchestration." IEEE Communications Surveys & Tutorials (2017).

[5] He, Wu, Gongjun Yan, and Li Da Xu. "Developing vehicular data cloud services in the IoT environment." IEEE Transactions on Industrial Informatics 10.2 (2014): 1587-1595.

[6] foursquare, online, available: http://foursquare.com/, retrieved Apr 15. 2017.

[7] General Motors, "Onstar.", online, available: http://www.onstar.com/ web/portal/home, Retrieved Apr 15. 2017.

[8] Zheng, Yu, et al. "GeoLife2. 0: a location-based social networking service." Mobile Data Management: Systems, Services and Middleware, 2009. MDM'09. Tenth International Conference on. IEEE, 2009.

[9] Zheng Y, Zhou X (2011) Computing with spatial trajectories. Springer.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[10] Dardari, Davide, Pau Closas, and Petar M. Djurić. "Indoor tracking: Theory, methods, and technologies." IEEE Transactions on Vehicular Technology 64.4 (2015): 1263-1278.

[11] Zhang, Lan, et al. "Montage: Combine frames with movement continuity for realtime multi-user tracking." IEEE Transactions on Mobile Computing 16.4 (2017): 1019-1031.

[12] Paulet, Russell, et al. "Privacy-preserving and content-protecting location based queries." IEEE Transactions on Knowledge and Data Engineering 26.5 (2014): 1200-1210.

[13] Fung, Eric, Georgios Kellaris, and Dimitris Papadias. "Combining Differential Privacy and PIR for Efficient Strong Location Privacy." International Symposium on Spatial and Temporal Databases. Springer International Publishing, 2015.

[14] Lin, Chi, Guowei Wu, and Chang Wu Yu. "Protecting location privacy and query privacy: a combined clustering approach." Concurrency and Computation: Practice and Experience 27.12 (2015): 3021-3043.

[15] Pan, Xiao, et al. "Protecting personalized privacy against sensitivity homogeneity attacks over road networks in mobile services." Frontiers of Computer Science 10.2 (2016): 370-386.

[16] Saravanan, Shanthi, and Balasundaram Sadhu Ramakrishnan. "Preserving privacy in the context of location based services through location hider in mobile-tourism." Information Technology & Tourism 16.2 (2016): 229-248.

[17] Piao, Chunhui, et al. "User privacy protection for a mobile commerce alliance." Electronic Commerce Research and Applications (2016).

[18] Gruteser M, Grunwald D (2003) Anonymous usage of location based services through spatial and temporal cloaking. In: Proceedings of the 1st international conference on mobile systems, applications and services (MobiSys '03), San Francisco, California, pp 31–42.

[19] Gedik, Bugra, and Ling Liu. "Protecting location privacy with personalized k-anonymity: Architecture and algorithms." IEEE Transactions on Mobile Computing 7.1 (2008): 1-18.

[20] Mokbel, Mohamed F., Chi-Yin Chow, and Walid G. Aref. "The new Casper: query processing for location services without compromising privacy."Proceedings of the 32nd international conference on Very large data bases. VLDB Endowment, 2006.

[21] Beresford AR, Stajano F (2004) Mix zones: user privacy in location-aware services. In: Proceedings of the second IEEE annual conference on pervasive computing and communications workshops (PerCom '04 Workshops), pp 127–131.

[22] Beresford, Alastair R., and Frank Stajano. "Location privacy in pervasive computing." IEEE Pervasive computing 2.1 (2008): 46-55.

[23] Palanisamy B, Liu L (2011) Mobimix: protecting location privacy with mix-zones over road networks. In: Proceedings of the 27th IEEE international conference on data engineering (ICDE '11), pp 494–505.

[24] Ardagna C, Cremonini M, Damiani E, De Capitani di Vimercati S, Samarati P (2007) Location privacy protection through obfuscation-based techniques. In: Proceedings of the 21st annual IFIP WG 11.3 working conference on data and applications security, Redondo Beach, CA, USA, pp 47–60.

[25] Duckham M, Kulik L (2005) A formal model of obfuscation and negotiation for location privacy. In: Proceedings of the third international conference on pervasive computing (Pervasive '05), Munich, Germany, pp 152–170.

[26] Damiani ML, Bertino E, Silvestri C (2010) The probe framework for the personalized cloaking of private locations. Trans Data Priv 3(2):123–148.

[27] Gruteser, Marco, and Dirk Grunwald. "Anonymous usage of location-based services through spatial and temporal cloaking." Proceedings of the 1st international conference on Mobile systems, applications and services. ACM, 2003.

[28] Xu, Toby, and Ying Cai. "Feeling-based location privacy protection for location-based services." Proceedings of the 16th ACM conference on Computer and communications security. ACM, 2009.

[29] Lin, Chi, Guowei Wu, and Chang Wu Yu. "Protecting location privacy and query privacy: a combined clustering approach." Concurrency and Computation: Practice and Experience 27.12 (2015): 3021-3043.

[30] Shao, Zhou, David Taniar, and Kiki Maulana Adhinugraha. "Range-kNN queries with privacy protection in a mobile environment." Pervasive and Mobile Computing 24 (2015): 30-49.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[31] Saravanan, Shanthi, and Balasundaram Sadhu Ramakrishnan. "Preserving privacy in the context of location based services through location hider in mobile-tourism." Information Technology & Tourism 16.2 (2016): 229-248.

[32] Paulet, Russell, et al. "Privacy-preserving and content-protecting location based queries." IEEE Transactions on Knowledge and Data Engineering 26.5 (2014): 1200-1210.

[33] Fung, Eric, Georgios Kellaris, and Dimitris Papadias. "Combining Differential Privacy and PIR for Efficient Strong Location Privacy." International Symposium on Spatial and Temporal Databases. Springer International Publishing, 2015.

[34] Hoh, Baik, and Marco Gruteser. "Protecting location privacy through path confusion." First International Conference on Security and Privacy for Emerging Areas in Communications Networks (SECURECOMM'05). IEEE, 2005.

[35] Meyerowitz, Joseph, and Romit Roy Choudhury. "Hiding stars with fireworks: location privacy through camouflage." Proceedings of the 15th annual international conference on Mobile computing and networking. ACM, 2009.

[36] Yiu ML, Jensen CS, Møller J, Lu H (2011) Design and analysis of a ranking approach to private location-based services. ACM Trans Database Syst 36(2):1–42.

[37] Ghinita G, Kalnis P, Skiadopoulos S (2007) Prive: anonymous location-based queries in distributed mobile systems. In: Proceedings of the 16th international conference on world wide web (WWW '07), Banff, Alberta, Canada, pp 371–380.

[38] Talukder N, Ahamed SI (2010) Preventing multi-query attack in location-based services. In: Proceedings of the third ACM conference on wireless network security (WiSec '10), Hoboken, New Jersey, USA, pp 25–36.

[39] Bamba B, Liu L, Pesti P, Wang T (2008) Supporting anonymous location queries in mobile environments with privacygrid. In: Proceeding of the 17th international conference on world wide web (WWW '08), Beijing, China, pp 237–246.

[40] Li N, Li T, Venkatasubramanian S (2007) t-closeness: privacy beyond k-anonymity and l-diversity. In: Proceedings of the 23rd IEEE international conference on data engineering (ICDE '07), pp 106–115.

[41] Solanas A, Sebe´ F, Domingo-Ferrer J (2008) Micro-aggregationbased heuristics for p-sensitive k-anonymity: one step beyond. In: Proceedings of the 2008 international workshop on privacy and anonymity in information society (PAIS '08), Nantes, France, pp 61–69.

[42] Mascetti S, Bettini C, Wang XS, Freni D, Jajodia S (2009) Providenthider: an algorithm to preserve historical k-anonymity in lbs. In: Proceedings of the 10th IEEE international conference on mobile data management (MDM '09), pp 172–181. Taipei, Taiwan.

[43] B. Gedik and L. Liu, "Protecting Location Privacy With Personalized k-Anonymity: Architecture and Algorithms," IEEE Trans. Mobile Computing, vol. 7, Jan. 2008, pp. 1–18.

[44] H. Kido, Y. Yanagisawa, and T. Satoh, ―An Anonymous Communication Technique Using Dummies for Location- based Services,‖ IEEE Proc. Int'l. Conf. Pervasive Services, ICPS '05, July 2005.

[45] B. Niu, Q. Li, X. Zhu, G. Cao, and H. Li, "Achieving k-anonymity in privacy-aware location-based services," in Proc. of IEEE INFOCOM 2014.

[46] A. Pingley et al., ―Protection of Query Privacy for Continuous Location Based Services,‖ IEEE INFOCOM'11, Apr. 2011.

[47] Hara, Takahiro, et al. "Dummy-Based User Location Anonymization Under Real-World Constraints." IEEE Access 4 (2016): 673-687.

[48] Gutscher A (2006) Coordinate transformation—a solution for the privacy problem of location based services? In: Proceedings of the 20th international conference on parallel and distributed processing (IPDPS '06), Rhodes Island, Greece, pp 354–354.

[49] Ardagna, Claudio Agostino, et al. "Location privacy protection through obfuscation-based techniques." IFIP Annual Conference on Data and Applications Security and Privacy. Springer Berlin Heidelberg, 2007.

[50] Mascetti S, Freni D, Bettini C, Wang XS, Jajodia S (2011) Privacy in geo-social networks: proximity notification with untrusted service providers and curious buddies. VLDB J 20(4):541–566.

[51] Marias G, Delakouridis C, Kazatzopoulos L, Georgiadis P (2005) Location privacy through secret sharing techniques. In: Proceedings of the 1st international IEEE WoWMoM

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

workshop on trust, security and privacy for ubiquitous computing (WOWMOM '05), pp 614–620.

[52] Shokri, Reza, et al. "Hiding in the mobile crowd: Location privacy through collaboration." Dependable and Secure Computing, IEEE Transactions on 11.3 (2014): 266-279.

[53] X. Zhu, H. Chi, B. Niu, W. Zhang, Z. Li, and H. Li, ―Mobicache: When k-anonymity meets cache,‖ in Proc. of IEEE GLOBECOM 2013.

[54] Niu, Ben, et al. "Enhancing privacy through caching in location-based services." Proc. of IEEE INFOCOM. 2015.

[55] Dewri, Rinku, and Ramakrisha Thurimella. "Exploiting service similarity for privacy in location-based search queries." Parallel and Distributed Systems, IEEE Transactions on 25.2 (2014): 374-383.

[56] Sun, Weiwei, et al. "An Air Index for Spatial Query Processing in Road Networks." Knowledge and Data Engineering, IEEE Transactions on 27.2 (2015): 382-395.

[57] Pfitzmann, Andreas, and Marit Hansen. "A terminology for talking about privacy by data minimization: Anonymity, unlinkability, undetectability, unobservability, pseudonymity, and identity management." (2010).

[58] Shannon, Claude E. "A mathematical theory of communication." ACM SIGMOBILE Mobile Computing and Communications Review 5.1 (2001): 3-55.

[59] Hoh, Baik, et al. "Achieving guaranteed anonymity in gps traces via uncertainty-aware path cloaking." IEEE Transactions on Mobile Computing 9.8 (2010): 1089-1107.

[60] Sweeney, Latanya. "k-anonymity: A model for protecting privacy." International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 10.05 (2002): 557-570.

[61] Andrés, Miguel E., et al. "Geo-indistinguishability: Differential privacy for location-based systems." Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security. ACM, 2013.

[62] Perazzo, Pericle, and Gianluca Dini. "A uniformity-based approach to location privacy." Computer Communications 64 (2015): 21-32.

[63] Voulodimos, Athanasios S., and Charalampos Z. Patrikakis. "Quantifying privacy in terms of entropy for context aware services." Identity in the Information Society 2.2 (2009): 155-169.

[64] Chen, Zhigang, et al. "LISA: Location information scrambler for privacy protection on smartphones." Communications and Network Security (CNS), 2013 IEEE Conference on. IEEE, 2013.

[65] Humbert, Mathias, et al. "Reconciling utility with privacy in genomics." Proceedings of the 13th Workshop on Privacy in the Electronic Society. ACM, 2014.

[66] Machanavajjhala, Ashwin, et al. "L-diversity: Privacy beyond k-anonymity." ACM Transactions on Knowledge Discovery from Data (TKDD) 1.1 (2007): 3.

[67] Krumm J (2007) Inference attacks on location tracks. In: Proceedings of the 5th international conference on pervasive computing (Pervasive '07). Springer, Toronto, pp 127–143.

[68] Shokri R, Theodorakopoulos G, Le Boudec J, Hubaux J (2011) Quantifying location privacy. In: Proceedings of the 31st IEEE symposium on security and privacy (SP '11), Berleley/Oakland, California, USA, pp 247–262.

[69] Lin, Chi, Guowei Wu, and Chang Wu Yu. "Protecting location privacy and query privacy: a combined clustering approach." Concurrency and Computation: Practice and Experience 27.12 (2015): 3021-3043.

[70] Saravanan, Shanthi, and Balasundaram Sadhu Ramakrishnan. "Preserving privacy in the context of location based services through location hider in mobile-tourism." Information Technology & Tourism 16.2 (2016): 229-248.

[71] Li, Yanhui, et al. "Semantic-Aware Location Privacy Preservation on Road Networks." International Conference on Database Systems for Advanced Applications. Springer International Publishing, 2016.

[72] Lee, Byoungyoung, et al. "Protecting location privacy using location semantics." Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2011.

[73] Mahin, Mehnaz Tabassum, Tanzima Hashem, and Samia Kabir. "A crowd enabled approach for processing nearest neighbor and range queries in incomplete databases with accuracy guarantee." Pervasive and Mobile Computing 39 (2017): 249-266.

[74] Cahsai, Atoshum, et al. "Scaling k-Nearest Neighbours Queries (The Right Way)." Distributed Computing Systems

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

(ICDCS), 2017 IEEE 37th International Conference on. IEEE, 2017.

[75] Mahadi, Swapnali M., and Sucheta M. Kokate. "A Fast nearest Neighbor Search Using KD Tree and Inverted Files." International Journal 4.7 (2016).

[76] Zuberi, Rubina S., and Syed N. Ahmad. "Secure Mix-Zones for Privacy Protection of Road Network Location Based Services Users." Journal of Computer Networks and Communications 2016 (2016).

[77] Dai, Jian, Zhi-Ming Ding, and Jia-Jie Xu. "Context-Based Moving Object Trajectory Uncertainty Reduction and Ranking in Road Network." Journal of Computer Science and Technology 31.1 (2016): 167-184.

[78] Basiri, Anahid, et al. "Indoor location based services challenges, requirements and usability of current solutions." Computer Science Review24 (2017): 1-12.

[79] Wernke, Marius, et al. "A classification of location privacy attacks and approaches." Personal and Ubiquitous Computing 18.1 (2014): 163-175.

[80] Zhang, Xu, and Hae Young Bae. "Location Positioning and Privacy Preservation Methods in Location-based Service." International Journal of Security & Its Applications 9.4 (2015).

[81] Shin, Kang G., et al. "Privacy protection for users of location-based services."Wireless Communications, IEEE 19.1 (2012): 30-39.