Kingdom of Saudi Arabia
Saudi Arabia Ministry of Higher Education
King Faisal University
College of Computer Sciences and Information Technology

Privacy Protection in Location Based Service (PPLBS)

Course name: Mobile Computing

Course Instructor: Dr.Mona

Section: 62

N.	Student name	ID
1	Salma Alomar	217019858
2	Nabaa Jafar	217017935
3	Maryam Alashwan	217035462

submission date: 26 / 11 /2020

	Assignment Component	Max. Marks	Marks Obtained		
			S1	S2	S3
REPORT	Format (Table of Contents, References, Acknowledgements, Properly bound, Professional style, formatting and font, etc....)	10			
	English language (grammar and spelling)	5			
	Introduction	5			
	Bibliography	10			
	Contents Analysis of related work, and Discussions	30			
	Extra Effort / Additional Research Done / Comprehensive Report	10			
Presentation	Technical contents (comprehensiveness, clarity and accuracy)	10			
	Questions and answers	5			
	Presentation skills	10			
	Style and formatting	5			
	Total Marks	100			
	Grade for this Project	10			
	Plagiarism report more than 10% to 25% Plagiarism report more than 25% to 40% Plagiarism report more than 40%	- 1 -3 -5			

Table of Contents

1.Introduction

These days the production of mobile devices such as smartphones, tablets, and so on are increasing, also, location-based services are becoming increasingly common and popular. Location-based services are considered as a system or a services that provide a geographical location of an entity also provide information about the site and the term entity, that come up with the information that leads to the location of an object, can be human or non-human. Therefore, **L**ocation **B**ased **S**ervices (LBS) applications are beneficial and suitable. However, its mannerism is a serious danger to the privacy of the users as it heads for disclosing their sites to LBS service providers with their queries to all site-based information. LBS researchers can distinguish between location tracking services and location-aware services. Location tracking services provide information about whereabouts of the users to an entity other than the users. On the other hand, a site-aware service provides the users with all required information personal site data. For example, a car navigation system is considered as a location-conscious service. In the location information is provided to the driver, to receive a real-time navigation service. Also, in location tracking services there is the UPS (uninterruptible power supply or uninterruptible power source) truck which tracks a system, where the location information is used for each truck to increase the efficiency of navy management. Furthermore, the site sends promotional offers to the phone of customers when the person becomes close to the users. Moreover, LBS reliance comes with new privacy threats to users; it may be due to the unreliability of LBS providers. The attacker could use the information of the user location to link sensitive information to the identity of users. The request sent to the provider is somehow anonymous so user can get benefit from the service, while canning its privacy in one of the best technologies used anonymity, then the name is used for fudge identity and the ID is hidden so, the service provider not able to link the data with sensitivity data that included in the source request. Unfortunately, the attacker can integrate the user's site with some other external knowledge to be able to identify the source. For example, the user used to make requests from his home. As well, if you use a fake ID to bring anonymity, with available general knowledge, the attacker may locate the source from his or her site. However, the researchers propose many algorithms and models to overcome this series problem. In this report, we discuss some technologies: **K-anonymity, Transformation, and Dummy location** [1,2,3].

2. Pros of LBS

LBS has many benefits regardless it is easy to use. Users will get advertisements depending on their information location, about restaurants and food they like, about trips they are interested in and all other things that user's interested in depending on information that they collect from location that users go to. Target users who use phones in the best time and place by sending for them relate messages. Users can find the nearest hospital or restaurant and other places. Users can go wherever they want to go; they just will search for a place and will get the location to go to it. Also, anyone can reach the users if they share their location. restaurants use LBS to reach to the customer and give them the delivery of food orders. Online shopping delivery for users orders by sharing their location. Some companies make their employees use devices that track location to know where every person in work, if they want to reach some employee immediately, they can by tracking their location. The company can track anyone and where they are on work by using these devices. users can use LBS on their car, so if someone steals their car, they can find the car. Also, if there was an accident traffic police will find where the accident is based on the location of the cars. it is useful to avoid accidents by knowing if another car is near your car. In healthcare, hospitals and emergencies can reach the patients faster by using LBS instead of telling them where the place is, they will just use LBS to reach. Also, the fire brigade will use LBS to reach the fire location faster [3,4].

3. Cons of LBS

Users use LBS that is not trusted, so other people get users information and take benefits of it, for example if users give their personal and sensitive information. Criminals can use LBS to do bad things, spying on users and stalking them, Criminals may know the user's house and their working by tracking the user's location. Bad people or thieves can enter the house if they know there is no one. if users are in far places or places where few people are, criminals can steal the user, kidnapping the user or other bad things in which it is possible to harm the company. For example, one of the employees is traced and they try to get information from them. One of the biggest disadvantages is privacy issues that cause user’s harm [3,4].

4. Review of Related Literature

LBS is one of the newest trends in mobile computing (MC) that provides services to mobile clients based on storing and keeping their data locations. User’s location privacy has been an essential issue that enormous researchers study and discuss widely during recent years. It is one of the challenges in MC that could be threatened and compromised once it is accessed by an unauthorized party. In this section, we will review several conducted studies and researches on the challenges of the user's location privacy in MC.

We can describe LBS as different mobile applications that utilize the physical location of the device in order to afford users with services depending on their actual location. Using LBS, the clients can use services near their current location [5,6]. For instance, searching for neighboring stores, restaurants, or checking traffic conditions on roads. Moreover, a mobile's camera is another way of using LBS. When the clients take a photo with their mobile camera, their geographical location is attached to that photo. What makes it worse is that the clients may share it on their social media accounts. Thus, the attached location is displayed on the system map automatically and could be shared by the client's followers on the social media applications [7]. Protecting the privacy of the client's location from unauthorized usage is crucial. The functions and services that are provided via LBS are what cause the client's location privacy risks. Some of the services gather and store historical data about location in the cloud which could leak personal information of the clients, such as their current location, or the inquiry content. While other services may reveal clients' hobbies and interests to the public and be misused by an attacker or spam advertising [8]. Once the attacker receives the information, the client's privacy is under threat and can be leaked. For instance, if the client's location and personal information are gathered by the social media network, it might be supplied to a third party for advertisement and commercial purposes [9,10]. Table 1 shows how different social media app track their clients' location and how each manages location privacy.

Application	Methodology	Techniques to manage location privacy
Facebook	Facebook knows where the client is from the IP address and GPS signal.	Clients can manage their location privacy by controlling to share their accurate location and allowing particular people to see it.
Google Apps	They are collecting location information through client's IP address, GPS, and sensor data from the client's device.	Using any Google Apps' services allows clients to stop the history of the location at any time and manage what history includes.
Instagram	It collects clients' location once clients post pictures into the application.	Clients manage their location privacy by controlling to share their accurate location and allowing particular people to see it.
Snapchat	One of Snapchat's services is a Snap Map which shows clients' location and shares it with the public.	Snapchat enables clients to hide their location on the Snap Map.
Uber	It gathers location data when the application is open on the screen, and on at the time of requesting a ride or delivery.	The application enables clients to disable tracking location from mobile settings.

Table 1 Methodology and techniques of several social application in tracking user location [11].

5.Technology to Protect the Privacy of the User's Location

In this section a discussion of some approaches/techniques that are used for protecting the privacy of the user's location will be presented. Also, it will include discussion about the strengths and weaknesses for each technology.

5.1 K-Anonymity

k-Anonymity: it hides the exact location of the target user through grouping the users into k user groups. As a result, the likelihood of identifying the target user's location will be 1/k, and that guarantees the targeted user is indistinguishable among k−1 users.

Implementing the K-anonymity model is requiring a trusted third party server (TTP) which is called a location anonymizer .It filters the incoming user requests and makes anonymous counterparts that can safely direct to the service providers to be serviced. In other words, the location anonymizer send the query location of the user to LBS server (that acts as an intermediate between the end-user and the service provider) in an expanded format to form a Cloaking Region (CR) that geographically covers another user (k-1) rather than the specific queried location[11]. The cloaking strategy is the most widely adopted anonymization strategy to satisfy K-anonymity in LBSs. Since the attacks can easily target the untrusted LBS servers that hold user's sensitive information the cloaking area prevents the untrusted LBS servers from identifying the real user's location from the other $k - 1$ dummy locations. Figure 1 demonstrate the most popular cloaking strategies that generalize the actual locations can be divided into two groups as the following:

1. Data–dependent cloaking methodology

It relies on the actual location of the individual user in the system and his/her distance from the location of request to form the area of anonymity.

2. Space–dependent cloaking methodology

This strategy formulates the regions of anonymity while considering the total area covered by the anonymizer [12].

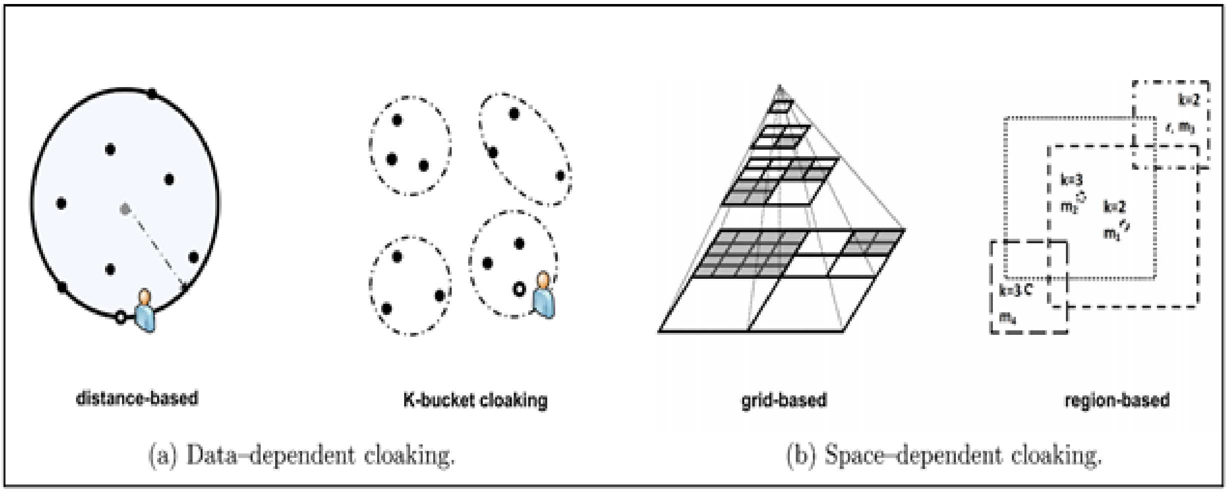

Figure 1: Cloaking Strategies for the offering of K-anonymity in LBSs [12].

The trusted server (anonymizer) has to incorporate algorithms that produce the anonymous counterparts of the original user query which do the following:(1) remove any identifiers of the user (e.g., id, name) inside the request. (2) effectively convert the specific location of the user into a spatiotemporal area that includes several nearby users, which will create other k-1 fake or dummy locations.

However, using the K-anonymity approach has some weaknesses. First, K-anonymity depends on location anonymizer, which suffers from a single point failure. That means the user's privacy will be exposed if the attacker takes the control of it. Second, due to the all delivered queries having to go through the location anonymizer, the k-anonymity suffers from a performance bottleneck. Finally, selecting the dummy location that achieves k-anonymity is a challenge, because most existing methods consider the attackers have no other side information such as the gender and social status of the users when they query about their location. As a result, these dummy generation algorithms may not work well since it will rely on a random walk model or virtual circle/grid model [13].

5.2 Transformation

It securely transforms the query location to prevent the cloud server of the LBS server from defining the user's location. This approach utilizes two techniques. First, Non-a spatial transformation that provides strong privacy by using cryptographic protocols. However, it requires high computational and communication costs. Second, spatial transformation used the geometric transformation to modify user location. For example, scaling, translation, and rotation. Besides, the SpaceTwist framework aims to blinds an untrusted location server. Where it retrieves and requests the Points-Of-Interest (POI) based on their ascending distance from a fake location that is residing nearby the query point. Called an anchor point. Due to this approach is similar to cloaking region-based approaches, it has several drawbacks first, privacy leaks when the exact results are required, also the communication and computational cost will increase as well. Second, the issue of choosing an anchor point. The technique lacked a global cloaking region if the anchor point was too close. On the other hand, if the anchor point was chosen far away the computational and communication cost will obviously increase [11,14].

5.3 Dummy Location

The Dummy location technique conceals and protects the client location by producing multiple unreal locations that are chosen randomly. It considers that the clients have historical traffic stored in a database that enables them to generate dummy locations to not be distinguished from their original locations. These false positions will be submitted to the server along with the true position through mobile devices. The Dummy location technique is powerful because the clients themselves can create dummies without any need for TTP components to transmit the true location of the client. However, choosing a huge number of dummy locations makes this approach suffer from heavy communication and computation cost. Furthermore, hiding and protecting clients' location could reduce in case the client keeps sending continuous data. In particular, when the clients use road navigation services, such as GPS, they have to send their location regularly. Each subject, in general, can move restrictedly in a fixed time. If the unreal locations are produced randomly, the attacker easily can discover the distinction between the real and fake positions. [11, 15, 16]

6.Netflix Case Study

Last year, Netflix 10 rated films through viewers and customers, and this led to better systems than those used by the company. To protect the privacy of the testimonials, their names were replaced with random numbers. Researchers at the University of Texas at Austin have hidden the identity of some Netflix data by comparing the order and timestamps with the general information in the Internet Movie Database, or IMDb. But they not only did not reveal the identity of the entire data set to Netflix but rather What they did was to reverse the anonymity of the Netflix dataset for those sampled users who also entered some movie ratings, under their names, in IMDb. While IMDb logs are public, crawling the site for obtaining them is against IMDb Terms of Service, so use Researchers are a few To prove the algorithm of their own representatives). The aim of this was to clarify the extent to which the information and data required by customers are hidden to conceal the identity, and that is through the information in the Netflix data set. But in the year 2001 the dangers of unknown databases were written and it was in the IEEE journal where the researchers who worked with Anonymous Netflix data of people's identities - as others did with the AOL search database last year - they only compared it with a already defined subset of similar data: standard data extraction technology. Because of the repeated analysis, this has led to a risk to unknown data. For example, a person can be able to access an unknown data set from phone records, hiding part of their identity by linking them to the phone order database for catalog dealers. Also, a person can obtain a key Partially abolishing the identity of a public database of credit card purchases through Amazon book reviews online. Google can through its database hide the identity of a public database for online purchases easily and in an uncomplicated manner.

Also, merchants who keep detailed and accurate information about clients can use some of their data to partially deactivate any data from a large search engine. Therefore, the data broker who maintains databases for many companies may be able to hide the identity of most of the records in these databases. Texas: This process is not difficult because when you cancel out the top 100 films that everyone watches, our customs for watching movies are completely individual. This will certainly be true, as our movie viewing habits are completely individual to our reading books, online shopping habits, phone habits, and web search habits. The clear plans and countermeasures on this matter are not sufficient. Netflix could rationalize its data set by removing a subset of data, but it turned out that this makes the problem a little more difficult, the identity removal algorithm is 54887 surprisingly powerful, and works with partial data, volatile data, even With faulty data. When you rate 8 movies, there may be two of them completely wrong. But they can uniquely identify the records in the data set and then there is very little identifiable data left: from IMDb, from your blog. The moral is that it only takes a small database named for someone to There is no way to remove his anonymity from a much larger anonymous database.

There is also another research that reached the same result, using anonymous data from the 1990 census, they found that 87 percent of the population of the United States, 216 million out of 248 million You can uniquely identify them via the Zip Code Five numbers along with their gender and date of birth. It is possible that about half of the population of the United States will be determined by gender, date of birth or country, and other personal information. On the one hand, anonymous data are an enormous boon for researchers. On the other hand, in the age of wholesale surveillance, where everyone collects data on us all the time, anonymization is very fragile and riskier than it initially seems. Like everything else in security, anonymity systems shouldn't be fielded before being subjected to adversarial attacks.; why should we expect anonymity systems to be any different? And, like everything else in security, anonymity is a trade-off. There are benefits, and there are corresponding risks. Researchers are currently working on developing algorithms and techniques that enable the secure release of anonymous datasets like Netflix's. That is a research result we can all benefit from.

7.Conclusion

The rapid spread of mobile phones has led to the user's privacy concerns related to LBSs Although accessing the data of users provides a lot of benefits for business, it makes them more worried about their own personal life that could be disclosed. In this paper, we suggest some technologies to minimize the leak of personal data and increase their privacy. The future of LBS is becoming better and more popular because a lot of researchers work together to find out the best algorithm and technology to eliminate privacy concerns.

References

[1] Shin, G., Xiaoen, J., Zhigang, C., Xin, H (2012) Privacy protection for users of location-based services, *IEEE Wireless Communications IEEE Wireless Commun. Wireless Communications*, *19*(1), 31-39.

[2] Yutaka, M (2007) Inferring Long-Term User Properties Based on Users Location History, *IJCAI 20th Int'l. Joint Conf. Artificial intelligence, Morgan Kaufmann Publishers*, 2159-2165.

[3] Hutter, D., Gruteser, M (2005) On the Anonymity of Periodic Location Samples, *Security in Pervasive Computing*,179–92.

[4] John, K (2007) Inference Attacks on Location Tracks, *Proc. 5th Int'l Conf. Pervasive Computing, Springer-Verlag*, 301–309.

[5] Marcelino, L., & Silva, C. (2018). Location privacy concerns in mobile applications. In Developments and advances in intelligent systems and applications (pp. 241–249). Cham: Springer.

[6] Shankar, P., Huang, Y. W., Castro, P., Nath, B., & Iftode, L. (2012). Crowds replace experts: Building better location-based services using mobile social network interactions. In IEEE international conference on pervasive computing and communications (PerCom), 2012 (pp. 20–29).

[7] Li, X. Y., & Jung, T. (2013). Search me if you can: Privacy-preserving location query service. In IEEE proceedings of INFOCOM, 2013 (pp. 2760–2768).

[8] Wang, S., Hu, Q., Sun, Y., & Huang, J. (2018). Privacy preservation in location-based services. IEEE Communications Magazine, 56(03), 134–140.

[9] Wang, T., Zeng, J., Bhuiyan, M. Z. A., Tian, H., Cai, Y., Chen, Y., et al. (2017). Trajectory privacy preservation based on a fog structure for Cloud location services. IEEE Access, 05, 7692–7701.

[10] Sun, G., Xie, Y., Liao, D., Yu, H., & Chang, V. (2017). User-defined privacy location-sharing system in mobile online social networks. Journal of Network and Computer Applications, 86, 34–45.

[11] Almusaylim, Z. A., & Jhanjhi, N. (2019). Comprehensive Review: Privacy Protection of User in Location-Aware Services of Mobile Cloud Computing. Wireless Personal Communications, 111(1), 541–564.

[12] Gkoulalas-Divanis, Aris & Kalnis, Panos & Verykios, Vassilios. (2010). Providing K–Anonymity in Location Based Services. SIGKDD Explorations. 12. 3-10. 10.1145/1882471.1882473.

[13] Niu, B., Li, Q., Zhu, X., Cao, G., & Li, H. (2014). Achieving k-Anonymity in Privacy-Aware LocationBased Services. In IEEE Proceedings of INFOCOM, 2014 (pp. 754–762).

[14] Abbas, Fizza & Hussain, Rasheed & Son, Junggab & Oh, Heekuck. (2013). Privacy Preserving Cloud-Based Computing Platform (PPCCP) for Using Location Based Services. Proceedings - 2013 IEEE/ACM 6th International Conference on Utility and Cloud Computing, UCC 2013. 60-66. 10.1109/UCC.2013.26.

[15] CA. Ardagna, M., AR. Beresford, F., CW. Chan, C., CY. Chow, M., CY. Chow, M., ML. Damiani, E., . . . C. Zhang, Y. (1970, January 01). A classification of location privacy attacks and approaches. Retrieved November 24, 2020.

[16] Kido, H., Yanagisawa, Y., & Satoh, T. (n.d.). An anonymous communication technique using dummies for location-based services. Retrieved November 24, 2020.