Open access Journal International Journal of Emerging Trends in Science and Technology

Trajectory Privacy in Participatory Sensing using K-anonymity

Authors

Sampada Sarodaya[1], Prof. S. V. Dabhade[2]

[1]Smt. Kashibai Navale College of Engineering, Vadgoan, Pune, India
Email: *sampada.sarodaya184@email.com*

[2]Smt. Kashibai Navale College of Engineering, Vadgoan, Pune, India
Email: *svdabhade@sinhgad.edu*

Abstract

Development of mobile communication, applications represents a challenge for both conceptually and technically so the basic requirements of LBS (location-based services) are numerous. Complex task is to provide user with added value to location information. Previously using Trajectory privacy-preserving framework user's location is preserved using various techniques, but the locations of users trajectories may not sensitive all the time. Even mix-zones are regions where users movement cannot be track by any applications. In this paper mobile users will reveal their location to database services in a periodic or on-demand manner. K-anonymity preserving management of location information by developing efficient and scalable system-level facilities for protecting the location privacy through ensuring location k-anonymity. In the context of LBSs and mobile clients, location k-anonymity refers to the k-anonymity usage of location information. A subject is considered location k-anonymous if and only if the location information sent from a mobile client to an LBS is indistinguishable from the location information of at least k -1 other mobile clients. New spatial cloaking techniques based on real or historical user trajectory is designed to protect user location trajectories and also satisfy the users' specified k-anonymity level.

Keywords: Location-based services, spatial cloaking, k-anonymity, mix-zones, mobile communication.

1. Introduction

With the development of wireless communication technologies such as WLAN, 3G/LTE,WiMax, Bluetooth, Zigbee, and so on, mobile devices are equipped with a variety of embedded sensors surveyed in as well as powerful sensing, storage and processing capabilities. Participatory sensing (urban sensing), which is the process that enables individuals to collect, analyze and share local knowledge with their own mobile devices, emerges as required under these well conditions. Compared with WSNs, participatory sensing offers a number of advantages on deployment costs, availability, spatial- temporal coverage, energy consumption and so forth.. Nowadays, participatory sensing applications mainly depend on the collection of data across wide geographic areas.

The sensor data uploaded by participators are invariably tagged with the spatial temporal information when the readings were recorded. According to the analysis in, the possible threats to a participator's privacy information that include monitoring data collection locations, tracing his/her trajectory, taking photographs of private scenes and recording the intimate chat logs. Once participators realize the serious consequences with the disclosure of their sensitive information, they are unwilling to participate in the campaign and use the services. Since the success of participatory sensing campaign strongly depends on the altruistic process of data collection, if the participators are reluctant to contribute their collected data, it would weaken the popularity and impact of this campaigns deployed at large scale while also reducing the benefits to the users.

Adversary may possibly analyze the trajectories which contain rich spatial-temporal history information to link multiple reports from the same participators and determine certain private information such as the places where the data reports are collected. Thus, it is necessary to unlink the participators' identities from sensitive data collection locations. To best of our knowledge, existing work on privacy in participatory sensing mainly concentrate on data contribution and reporting process. If an adversary has *a priori* knowledge of a participator's trajectory, it is effortless to deanonymize his/her reports.

A trajectory privacy-preserving framework, named TrPF, for participatory sensing is studied. The locations on or nearby participators' trajectories may not all be sensitive, and with this thought, proposal only deals with the sensitive trajectory segments. Compared with existing trajectory privacy-preserving proposals, our proposal has advantages of lower costs and information loss while the privacy level would not decrease. With the advent of mobile technology, the area of participatory sensing (PS) has attracted many researchers in different domains such as public health, urban planning, and traffic. The goal is to leverage sensor equipped mobile devices to collect and share data, which can later be utilized for analysis, mining, prediction or any other type of data processing. While many *unsolicited* PS systems exist (e.g., Flickr, Youtube), in which users participate by arbitrarily collecting data, other PS systems are *campaign*-based, which require a *coordinated* effort of the participants to collect a particular set of data that the server requires for any purpose. Some real-world examples of PS campaigns include where users leverage their mobile devices to collect traffic information. However, privacy concerns are the significant barriers to the success of any participatory sensing campaign, which delay the progress of massive deployment of such systems. Consider a scenario where the goal of the PS campaign is to collect pictures/videos from the anti-government riots at different locations of a city with the coordinated effort of the participants.

2. Motivation

Mobile phones or Smartphone are rapidly becoming the central computer and communication device in people's lives. Until recently mobile sensing research such as activity recognition, where people's activity (e.g., walking, driving, sitting, talking) is classified and monitored, required specialized mobile devices. For this privacy is main issue so to analyzed the realistic architectural assumptions and privacy requirements, and then provided an instantiation that achieved privacy protection in participatory sensing with provable security. When using dummies-based location privacy,[3] First, it is possible to extend it to 2-dimensional space with obstacles, which are regions where service users cannot be located. The dummy generation algorithms must still satisfy the privacy requirements, but must take into account the obstacles. Second, it is relevant to consider privacy-area (or some other privacy metric) aware location privacy in a spatial network setting. Third, while if we consider only snapshot queries, it is of interest to extend it to support also continuous queries that can be issued by mobile users.

A major privacy threat specific to LBS usage is the location privacy breaches represented by space or time correlated inference attacks. Such breaches take place when a party that is not trusted gets access to information that reveals the locations visited by the individual, as well as the times during which these visits took place. An adversary can utilize such location information to infer details about the private life of an individual such as their political affiliations, alternative lifestyles, or medical problems or the private businesses of an organization such as new business initiatives and partnerships.[5] A personalized k-anonymity model for protecting location privacy describes an important observation: location privacy demands personalization. In other words, location privacy is a personalized requirement and is context sensitive. An individual may have different privacy requirements in different contexts, and different individuals may require different levels of privacy in the same context.).

3. Literature Survey

3.1 PAD: Privacy-Area Aware, Dummy-Based Location Privacy in Mobile Services.

Privacy protection techniques that use *k*-anonymity convert an original query into an anonymous query that contains the locations of multiple users. Such techniques, however, generally fail in offering guaranteed large privacy regions at reasonable query processing costs. In this paper, the PAD approach that is capable of offering privacy-region guarantees. To achieve this, PAD uses so-called dummy locations that are deliberately generated according to either a virtual grid or circle. These cover a user’s actual location, and their spatial extents are controlled by the generation algorithms. The PAD approach only requires a lightweight server-side front-end in order for it to be integrated into an existing client/server mobile service system. In addition, query results are organized according to a compact format on the server, which not only reduces communication cost, but also facilitates the result refinement on the client side.

3.2 Trajectory Privacy in Location-based Services and Data Publication

The ubiquity of mobile devices with global positioning functionality (e.g., GPS and AGPS) and Internet connectivity (e.g., 3G and Wi-Fi) has resulted in widespread development of location-based services (LBS). Typical examples of LBS include local business search, e-marketing, social networking, and automotive traffic monitoring. Although LBS provide valuable services for mobile users, revealing their private locations to potentially untrusted LBS service providers pose privacy concerns. In general, there are two types of LBS, namely, snapshot and continuous LBS. For snapshot LBS, a mobile user only needs to report its current location to a service provider once to get its desired information. There are two types of LBS, namely, snapshot and continuous LBS. For snapshot LBS, a mobile user only needs to report its current location to a service provider once to get its desired information. On the other hand, a mobile user has to report its location to a service provider in a periodic or on-demand manner to obtain its desired continuous LBS..

3.3 K-anonymity: A Model For Protecting Privacy

Consider a data holder, such as a hospital or a bank, that has a privately held collection of person-specific, field structured data. Suppose the data holder wants to share a version of the data with researchers. How can a data holder release a version of its private data with scientific guarantees that the individuals who are the subjects of the data cannot be re-identified while the data remain practically useful? The solution provided in this paper includes a formal protection model named k-anonymity and a set of accompanying policies for deployment. A release provides k-anonymity protection if the information for each person contained in the release cannot be distinguished from at least k-1 individuals whose information also appears in the release. This paper also examines re-identification attacks that can be realized on releases that adhere to k-anonymity unless accompanying policies are respected.

3.4 A Formal Model of Obfuscation and Negotiation for Location Privacy

Obfuscation concerns the practice of deliberately degrading the quality of information in some way, so as to protect the privacy of the individual to whom that information refers. In this paper, we argue that obfuscation is an important technique for protecting an individual’s location privacy within a pervasive computing environment. The paper sets out a formal framework within which obfuscated location-based services are defined. This framework provides a computationally efficient mechanism for balancing an individual’s need for high-quality information services against that individual’s need for location privacy. Negotiation is used to ensure that a location-based service provider receives only the information it needs to know in order to provide a service of satisfactory quality.

3.5 Anonymous Usage of Location-Based Services Through Spatial and Temporal Cloaking

Advances in sensing and tracking technology enable location-based applications but they also create significant privacy risks. *Anonymity* can provide a high degree of privacy, save service users from dealing with service providers' privacy policies, and reduce the service providers' requirements for safeguarding private information. However, guaranteeing anonymous usage of location-based services requires that the precise location information transmitted by a user cannot be easily used to re-identify the subject. Paper presents a middleware architecture and algorithms that can be used by a centralized location broker service. The adaptive algorithms adjust the resolution of location information along spatial or temporal dimensions to meet specified anonymity constraints based on the entities who *may* be using location services within a given area. Using a model based on automotive traffic counts and cartographic material, estimation of the realistically expected spatial resolution for different anonymity constraints.

3.6 Traffic-Aware Multiple Mix Zone Placement for Protecting Location Privacy

Privacy protection is of critical concern to Location-Based Service (LBS) users in mobile networks. Long-term pseudonyms, although appear to be anonymous, in fact empower third-party service providers to continuously track users' movements. Researchers have proposed the mix zone model to allow pseudonym changes in protected areas. In this paper, a new form of privacy attack to the LBS system that an adversary reveals a user's true identity and complete moving trajectory with the aid of side information is studied. A new metric to quantify the system's resilience to such attacks, and suggest using multiple mix zones to tackle this problem is proposed. A mathematical model is presented that treats the deployment of multiple mix zones as a cost constrained optimization problem. Furthermore, the influence of traffic density is also taken into account to enhance the protection effectiveness. The placement optimization problem is NP-hard. We therefore design two heuristic algorithms as practical and effective means to strategically select mix zone locations, and consequently reduce the privacy risks of mobile users trajectories. The effectiveness of our proposed solutions is demonstrated through extensive simulations on real-world mobile user data traces.

4. Problem Statement

In participatory sensing system, data reports collected by participators are tagged with spatial-temporal information. Since the location information that attached to the collected data reports are commonly shared, a prominent attack is thus the Trajectory Inference. to achieve more privacy the entire area of monitoring is split to many small zones and if the number of people in the zone is less than a certain threshold say K. The zones are aggregated as region till the number of person is that region is greater than K. By this way , trajectory privacy can be achieved. The K value is configured by admin of the system. If the K value is low , the privacy of user is breached by inference attacks, so K value is high, the accuracy of Applications will be affected.

5. System Architecture

Development of mobile communication, applications represents a challenge for both conceptually and technically so the basic requirements of LBS (location-based services) are numerous. Complex task is to provide user with added value to location information. K-anonymity preserving management of location information by developing efficient and scalable system-level facilities for protecting the location privacy through ensuring location k-anonymity.

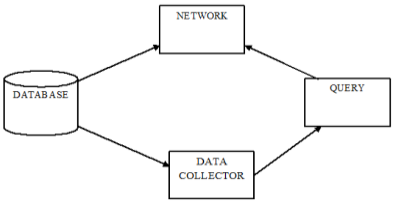

Figure 1: System Architecture

In the context of LBSs and mobile clients, location k-anonymity refers to the k-anonymity usage of location information. A subject is considered location k-anonymous if and only if the location information sent from a mobile client to an LBS is indistinguishable from the location information of at least k -1 other mobile clients.

6. Algorithm

Algorithm is as follows:

```
INPUT: Set Mobile Nodes $N_i$: {$N_1$, $N_2$, $N_3$,.........$N_M$}
   and K-anonymity level: K
OUTPUT: Aggregated area: $Ar_{area}$
1) Start
2) Create network $N_t$ and Sensor node $S_n$
3) N ⟵ $N_i$      Generate n number of node.
4) MOV($N_i$) and COLLECT(data)
5) SECURE($G_r$)    Secure particular grid
6) If SECURE ($G_r$)  Then
     $N_i$ ⟵  H   Hide mobile node
    Else
    Calculate(total number of nodes)
     end If
7)If  N<K Then
            $Ar_{area}$              Aggregated Are
            Z ⟵ $Z_{id}$             Zones with zone Id
      end If
8) End
```

In this algorithm, we give input as mobile node and K-anonymity level then with the help of basic GUI network is created. Then the data is collected from database i.e the data of participator which are present in the network. Their location and details are shown. Even nodes can be moved within the network whenever required. Then we have to secure grid. When particular grid is secured all the participator in that grid are kept hidden. All other participator present in other sensor node can be calculated properly. Then it is checked either the nodes present in sensor node is less then K-anonymity level, if it is then the zones are aggregated.

7. Result

The proposed system is developed to preserve trajectory privacy in participatory sensing. So the entire area of monitoring is split to many small zones and if the number of people in the zone is less than a certain threshold say K. The zones are aggregated as region till the number of person is that region is greater than K. By this way , trajectory privacy can be achieved. The K value is configured by admin of the system. If the K value is low , the privacy of user is breached by inference attacks, so K value is high.

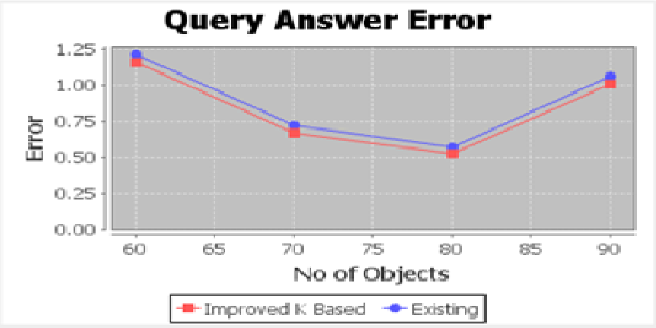

Figure 2: Relative rate between Error n objects

In figure 2 graph is plotted between error and no of objects. the objects ranges from 60-90.

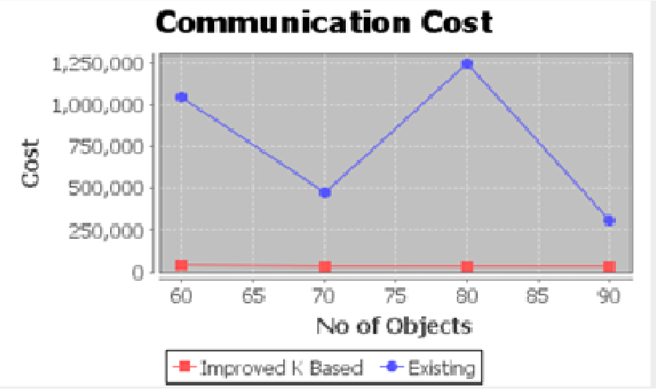

Figure 3: Communication Cost

In figure 3, the graph is plotted between objects and cost of the system.

Conclusion

In the context of LBSs and mobile clients, location k-anonymity refers to the k-anonymity usage of location information. A subject is considered location k-anonymous if and only if the location information sent from a mobile client to an LBS is indistinguishable from the location information of at least k - 1 other mobile clients. If the location information sent by each mobile client is perturbed by replacing the position of the mobile client with a coarser grained spatial range such that there are k - 1 other mobile clients within that range (k > 1), then the adversary will have uncertainty in matching the mobile client to a known location-identity association or an external observation of the location identity binding. This uncertainty increases with the increasing value of k, providing a higher degree of privacy for mobile clients. Previously, the problem with the system is the accuracy of participatory sensing is reduced in mix-zones model and admin have to configure all security zones in the system which in returns require more communication cost and time. To solve this problem the zones are aggregated a region till the number of person is that region is greater than K. The K value is configured by admin of the system. If the K value is low , the privacy of user is breached by inference attacks, so K value is high. By this way , trajectory privacy can be achieved. System automatically determined based on the number of person in zone.

Acknowledgement

The authors would like to thank the researchers as well as publishers for making their resources available and teachers for their guidance. We are thankful to the authorities of Savitribai Phule University, Pune and concerned members of journal, for their valuable guidelines and support. We are also thankful to the reviewer for their valuable suggestions. We also thank the college authorities for providing the required infrastructure and support. Finally, we would like to extend a heartfelt gratitude to friends and family members.

References

1. Sheng Gao, Jianfeng Ma, Weisong Shi, TrPF: A Trajectory Privacy-Preserving Framework for participatory Sensing, IEEE Transactions on Information Security, Vol 8, no.6, JUNE 2013.
2. C. Y. Chow and M. F. Mokbel, "Trajectory privacy in location-based services and data publication," *ACM SIGKDD Explorations Newsletter*, vol. 13, no. 1, pp. 19–29, 2011.
3. H. Lu, C. S. Jensen, andM. L. Yiu, "Pad: Privacy-area aware, dummybased location privacy in mobile services," in *Proc. 7th ACM Int.Workshop on Data Engineering for Wireless and Mobile Access*, 2008, pp. 16–23, ACM.
4. M. Gruteser and D. Grunwald, "Anonymous usage of location-based services through spatial and temporal cloaking," in *Proc. ACM 1st Int. Conf. Mobile Systems, Applications and Services*, 2003, pp. 31–42.
5. B. Gedik and L. Liu, "Protecting location privacy with personalized k-anonymity: Architecture and algorithms," *IEEE Trans. Mobile Comput.*, vol. 7, no. 1, pp. 1–18, Jan. 2008.
6. M. Duckham and L. Kulik, "A formalmodel of obfuscation and negotiation for location privacy," in *Proc. 3rd Int. Conf. Pervasive Computing (PERVASIVE'05)*, 2005, pp. 152–170.
7. X. Liu, H. Zhao,M. Pan, H. Yue, X. Li, and Y. Fang, "Traffic-aware multiple mix zone placement for protecting location privacy," in *Proc. IEEE INFOCOM*, 2012, pp. 972–980
8. L. Kazemi and C. Shahabi, "Towards preserving privacy in participatory sensing," in *Proc. 9th Int. Conf. Pervasive Computing and Communications Workshops (PERCOM Workshops)*, 2011, pp. 328–331.
9. C. Ardagna, M. Cremonini, E. Damiani, S. De Capitani Di Vimercati, and P. Samarati, "Location privacy protection through obfuscationbased techniques," *Data and Applications Security XXI*, pp. 47–60, 2007.
10. S. Gao, J. Ma, W. Shi, and G. Zhan, "Towards location and trajectory privacy protection in participatory sensing," in *Proc.Mobile Computing, Applications and Services*, Los Angles, CA, USA, 2011, pp. 381–386.

11. L. Liu, "From data privacy to location privacy: Models and algorithms," in *Proc. 33rd Int. Conf. Very Large Data Bases (VLDB2007)*, 2007, pp. 1429–1430, VLDB Endowment.
12. .M. Decker, "Location privacy-an overview," in *Proc. IEEE 7th Int. Conf. Mobile Business (ICMB'08)*, 2008, pp. 221–230.
13. P. Samarati and L. Sweeney, "Protecting Privacy when Disclosing Information: k-Anonymity and Its Enforcement through Generalization and Suppression," Proc. IEEE Symp. Research in Security and Privacy, 1998.
14. B. Gedik and L. Liu, "A Customizable k-Anonymity Model for Protecting Location Privacy," Proc. IEEE Int'l Conf. Distributed Computing Systems (ICDCS '05), pp. 620-629, 2005.
15. Meyerson and R. Williams, "On the Complexity of Optimal k-Anonymity," Proc. ACM Symp. Principles of Database Systems (PODS '04), pp. 223-228, 2004.