International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 4 Issue 10, October-2015

A Privacy Enhanced Geosocial Application

Sony L Johnson
Dept. of computer science and engineering
MBCET
Trivandrum, India

Devi Priya V.S
Dept. of Computer science and engineering
MBCET
Trivandrum, India

***Abstract*—with the widespread use of smart phones, a lot of location based services are quickly becoming popular. FourSquare, Gowalla etc. are geosocial applications which exploit GPS location services. These applications enable a user to communicate collaboratively with their surroundings. For a user to use these applications, their location coordinates should be send to the service provider in plain text. But this creates increased risks to personal privacy. The existing system to provide user privacy applies a distance preserving transformation to user's location coordinates. This transformation is performed with the help of certain secrets (an angle and a shift). Queries can be run on transformed coordinates as on plain location coordinates. In order to further add security, it decouples location coordinates and location data with the help of a random index. So there is no means for the service provider to get user data. But the data stored on the server should be made available to officials in the case of crime investigation purposes. So a system is proposed where user's data is made available for an investigation authority for a limited period of time. This is done by Time Specific Encryption (TSE) where an authority can access user data only for a limited period of time without compromising the privacy properties of the user.**

Keywords- Location Based Services; Coordinate transformation; Time Specific Encryption;

I. INTRODUCTION

Mobile communication has undergone a revolutionary growth in the last decade. The recent advances in mobile communication along with positioning technologies such as Global Positioning System (GPS) and Radio Frequency Identifier (RFID) popularized Location Based Services (LBS). LBS deliver personalised services to mobile users based on their location [4].

Geosocial networking combines social networking with user's offline location. Geosocial applications allow 'check-in' functionality to users to share user's current location with their friends. Geosocial applications such as Foursquare, Gowalla are very popular since they help users connect with their friends, family and other contacts in their local area [12]. Features such as Wi-Fi connectivity and GPS navigation allow more sophisticated capabilities in smart phones like monitoring of road and traffic conditions [13].

For a geosocial application to provide personalized location services to a mobile user, the mobile user's current location should be known to the application server. Then only the server could process different types of location queries such as point queries, nearest neighbor queries and circular range queries efficiently. For current services with minimal privacy mechanisms, this data can be used to infer users detailed activities or to track and predict a user's daily movements [1]. Thus location data at the servers have privacy concerns. The trusted servers can be compromised by different ways such as through computer break-ins, configuration errors at the servers or by software bugs [14, 15]. As per Microsoft safety and security center [8] location services have different risks including the following:

1. The application services that uses a person's location data may sell these to advertisers who may then pushes advertisements to user's mobile phones based on their location.
2. Applications like Foursquare that track a person's location may use this data for activities like spying, stalking or theft.

Locx [1] a privacy providing technique in geosocial applications which ensures a provable amount of privacy to location and location data stored on the server. Using this mechanism server cannot obtain any data regarding a particular user. But user data stored on the server should be made available to officials for crime investigation purposes. Thus the main aim is to design a mechanism that provides user privacy along with providing user data to the authority on demand.

II. LITERATURE REVIEW

Location privacy is defined as the ability to prevent other parties from learning one's current or past location [3]. A system that can obtain position data invades location privacy. Existing systems have different approaches to provide privacy in geosocial applications. 1) Anonymization approaches where approximate location and time is sent to the server instead of exact values. But this will result in inaccurate results. 2) Heavy weight cryptographic mechanisms such as homomorphic encryption, Private Information Retrieval (PIR) etc. But this is an expensive mechanism.

One of the approaches to provide privacy in LBS is by providing anonymity [5] to user data. Location anonymity is defined as the masking of position data. Instead of sending exact values to the location server, this approach sends an approximate location coordinate and time to the server. i.e, instead of disclosing user's exact location information, a bounding box is reported containing at least k people. The intuition here is that this provides security for the users, because within the bounding box a person's location cannot be identified. Therefore a user cannot be individually identified from k-1 other user's location. But location anonymization is not enough to provide privacy. Because the profile of a user contains further details which can be used by

an adversary to individually identify a person. The work proposed in [4] extends the notion of k anonymity to provide privacy in personalized LBS environment. It provides anonymity to a user even when profiles of users are known to adversaries. Instead of generalizing the location only, this approach generalizes both location and user profiles based on user's privacy requirement.

Another approach to provide anonymity with LBS is by using false position data called dummies [2]. Here user sends these position data along with the original position data to the service provider. The service provider cannot identify which is true position data and which is false position data. It replies the user with each of the received position data. From this reply message user simply extracts his necessary information. Thus in this manner by storing a set of position data along with the true position data, user's location gets anonymized. Dummy creation of position data is performed in such a way that they cannot be easily distinguishable from actual position data. If dummies are created randomly, then observer can easily distinguish between dummy data and actual position data. In such cases location anonymity is reduced. To avoid this dummy data should match with the original position data. Two dummy generation algorithms are used for this.MN algorithm [2] and MLN algorithm [2]

When a user sends a set of dummy position data along with the actual position data, the service provider will respond with each of the received location data. The response for the dummy data is ignored by the user. This approach will obviously increase the communication cost. . To reduce the effect of this communication cost, a cost reduction technique for communication between client and server by dividing the position coordinates in to two sets of coordinates X and Y is used.

Reference [7] implements an architecture which provides the property that the service provider of LBS does not learn any information about its customer's location. It exploits trusted computing and Private Information Retrieval. This architecture does not disclose location information to a service provider. A network operator implements an API. The service provider makes use of this API to provide different Location Based Services to the user. A query responder module forwards customer queries to the service provider and service provider's responses back to the customer. A locator module encrypts the location information of the customer before forwarding it to the service provider using public key encryption. The trusted computing module contacts the query scheduler module and customer queries are processed. The trusted computing module is having a property that location information cannot be learnt by the service provider that implements the trusted computing module. The architecture implements a Private Information Retrieval algorithm. Using this PIR algorithm the trusted computing module can retrieve entries from the location database in such a way that, the entry being retrieved is unknown to the administrator of the database (i.e., service provider).

Spatial and temporal cloaking [6] are techniques which achieve k-anonymity by sending approximate time and location information to the service provider. Spatial cloaking achieves anonymity by ensuring that every location sent to a service provider is a cloaking area that contains at least k nodes [5]. ICliqueCloak [16] is a location cloaking algorithm that prevents attacks based on location when user's location are continuously updated. This method maintains maximum cliques needed for location cloaking in an undirected graph. This undirected graph considers the effect of continuous location updates. Thus a clique can quickly be identified and used to generate the cloaking region when a new request arrives. Pseudonym [9] is another mechanism to achieve cloaking. In this mechanism cloaking is achieved by changing the device identifiers frequently. But this mechanism hurts functionality.

Longitude [10] and LocX [1] are mechanisms that provide privacy based on coordinate transformation.

III. OVERVIEW

A. A Basic design

In a location based application, the server needs to implement different types of queries on the location coordinates of the user. For a server to do this the location coordinates should be sent to the server in plain text which raises privacy concerns.

So the solution to this is a method using coordinate transformation [1]. Each user 'u' in the application selects a set of secrets including a rotation angle θ_u, a shift parameter b_u, and a symmetric key and shares these secrets only to their friends and use it to transform all the location coordinates that they stored on the servers. For example when Alice want to store a review regarding a particular location at (x, y), she transforms (x, y) to (x', y') using her secrets. The transformation is performed as follows:

$$(x', y') \leftarrow x\cos\theta - y\sin\theta + b_u\,,\; x\sin\theta + y\cos\theta + b_u\,. \qquad (1)$$

AES encrypted review of the corresponding coordinate is then stored in the server as E(r).

When Bob wants to get review stored by Alice regarding the location (x, y), he applies the same transformation to obtain (x', y'). Using this (x', y') he retrieve the encrypted review E(r) from the server which he decrypts it with Alice's symmetric key to obtain the review 'r'. In the same way all the friends of Alice can obtain all her reviews using the shared secrets and Bob can retrieve all his friend's reviews using each of the friend's secrets. One important limitation with this method is that, here the server can individually recognize the client devices using the IP address. So there are chances that an attacker with access to server data may try to relate a transformed coordinate and corresponding encrypted review with the client device. Thus we need to decouple the correlation between a transformed coordinate and an encrypted data. It can be done by placing it on two different servers, and the dependency can be maintained by using a random index. So one server contains the transformed coordinate and an encrypted index (called as an index server). The other server (data server) contains the index corresponding to the encrypted index in the index server and the encrypted data.

Location and data decoupling:

Here location and data are decoupled by storing them on two different servers. The index server contains a transformed location (x', y') and an encrypted index E(i). The index used is a random number which is generated by a pseudorandom number generator. The data server contains index and encrypted location data E ($data_{(x, y)}$).

Data Storage:

For a user to store some location data, he first transforms his location coordinates (x, y) in to virtual coordinates (x', y') using his rotation angle θu and shift bu. After this user generates a random index i using a pseudorandom number generator. This index generated is encrypted using AES algorithm with his secret symmetric key. Then he stores the transformed coordinates (x', y') and encrypted index E (i) on index server.

Location data $data_{(x,y)}$ is then encrypted with AES encryption algorithm and stores index and corresponding encrypted data E($data_{(x,y)}$) in to other server called data server.

Data Retrieval: Each of a user's friends can access their friend's location data stored on two different servers. For this a user first transforms a location (x, y) for which he wants to get some review to (x', y'). He then queries the index server with this (x', y') to obtain the encrypted index E (i). User having his friend's key can decrypt it to get the index i. Using this index i he queries the data server to obtain the encrypted data E($data_{(x,y)}$). He can decrypt it using friend's key to get the plaintext data.

B. Time Specific Encryption

In the above mechanism user data is highly secured at the server side. There is no means for the server to get user data. But it is useful for police officials to get access to user data for crime investigation purposes. User Privacy properties should still hold. So this is addressed by a Time Specific Encryption (TSE) [11]. Using TSE authenticated police authorities can get access to user data for a specified period of time. Authority can request access to user data. This is done with the help of another application. A request is sent from this application to the server for data access of a particular user. This request contains identity of the user and digital signature of the requestor. The server first verifies the digital signature of the requestor. Server forwards this request to the user only if the digital signature is verified. Upon receiving this request, user can grant permission to access his data. User can specify a time interval for which authority is allowed to access his data. When an authenticated request arrives, user takes the requested data and encrypts it with a newly created key. This key is then encrypted with the public key of the requestor. RSA encryption is used here. This response is sent to the server which is then forwarded to the requestor. The requestor can see the requested data only if he access the response within the time limit specified by the user.

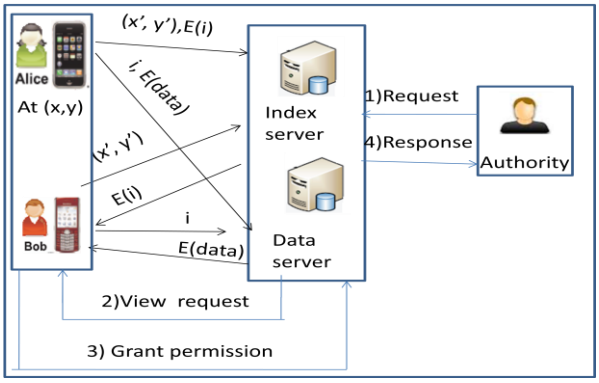

Figure 1:Time Specific Encryption

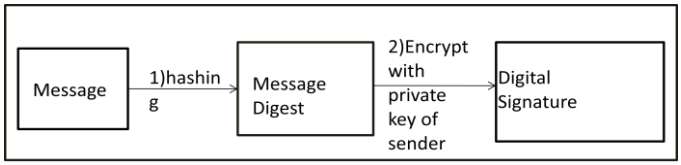

Figure 2.1: Digital Signature: Sender side

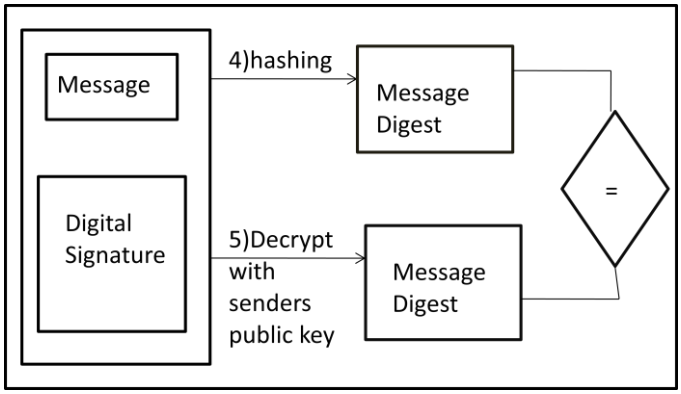

Figure 3.1: Digital signature: Receiver side

C. Algorithm to find whether a given time is within a specified time interval t0 and t1[11].

L ← binary of t0
R← binary of t1
Let S=ϕ
While L<R do
If L≡0 mod 2 then
L=parent(L)
else
S=S U {L}
L=parent(L)+1
endif
if R≡0 mod 2 then
S=S U {R}
R=parent(R)-1

else

R=parent(R)

endif

end while

This algorithm will generate a pattern from a binary tree. The tree construction is as follows.

1. Let r be the root of the tree.
2. Left child of the root will be zero and right child of the root will be 1.
3. Left child of each node can be created by appending '0' to the parent node.
4. Right child of each node can be created by appending '1' to the parent node.

Leaf nodes of the tree represent binary values of time instances. If we want to create a binary tree with leaf nodes up to 'n', then depth of the tree will be' d' such that:

$$2^d = n$$

The path generated for a specified time and pattern generated using the above algorithm has a common element, then it is a valid time. i.e., if the time that the authority tries to access is within the time interval specified by the user, then only access is possible to authority.

IV. Privacy Analysis

A. *Protection from an attacker with data access on the two servers.*

The location data of users stored on two different servers does not reveal any user data to the attacker. Since the coordinates are all transformed in the index server and location data on the data server are encrypted, attackers cannot perform any malicious activities. Even if an attacker has access to two different servers, they cannot link an entry in one server to the other. Only an authenticated user knows how to link an entry in an index server to the corresponding entry in the data server. Since the location coordinates is transformed, an attacker cannot associate a location to a real world location in the world.

V. Conclusion

In this system, users efficiently transform all location coordinates that they shared with the servers with distance preserving transformation. Also location data are all encrypted with 128 bit AES encryption. This transformation and encryption along with decoupling of location and location data ensures privacy to user. Only friends with correct key can decrypt a user's transformed coordinates and encrypted data. Time Specific encryption (TSE) is also implemented where investigation authority can access user data for a specified period of time. Here an authority can access user data only if the user allows access to his data. But here there are chances that, a user who is criminal might not allow access to his data. In such cases, Authority can collect further details of user such as his transformation keys from his friend circle, and thus can get user data from the server. This is left as a future work.

References

[1] Krishna P.N puttaswamy, Shiyuan Wang, Troy Steinbauer, Divyakant Agarwal, Amr El Abbadi,Christopher Kruegel, and Ben Y Zhao. "Preserving Location Privacy in Geosocial appliactions", IEEE transactions on mobile computing, Jan 2014.

[2] Hidetoshib Kido, Yutaka Yanagisawa and Tejsuji Satoh "An anonymous communication technique using dummies for loaction based services".Proc.Int'l Conf on Pervasive Services,2005.

[3] Alastair R. Beresford and Frank Stajano. "Location Privacy in Pervasive Computing" IEEE CS and Communication Society, 2003.

[4] Heechang Shin, Vijayalakshmi Alturi, and Jaideep Vaidya. "A Profile Anonymization Model for Privacy in a Personalised Location Based Service Environment", IEEE computer society, 2008.

[5] M. Grutesar and D. Grunwald, "Anonymous Usage of Location Based Services through Temporal Cloaking", Proc. First Int'l Conf. Mobile System Applications services, 2003.

[6] Xu Zhang, Gyoung Bae-Kim, Hae-Young Bae. "An Adaptive Spatial Cloaking Method for Privacy Protection in Loction Based Service," in information and Communication Technology Convergence (ICTC), 2014 International Conference on IEEE, 2014

[7] Urs Hengartnet and David R. "Hiding Location Information from Location Based Services" IEEE, 2007

[8] http://microsoft.com/security/online-privacy/location-services.aspx

[9] Oliver Jorns, Oliver Jung and Jerald Quirchmayr. "'A platform for the development of Location Based Mobile Applications, COMSWARE,, IEEE, 2015 .

[10] S. Mascetti, C bettini, and Freni, "Longitude: Centralised Privacy Preserving Computation of User's Proximity", in Proc of SDM, 2009

[11] Kenneth G. Paterson and Elizabeth A. Quaglia, "Time-Specifc Encryption" in security and Cryptography for Networks. Springer, 2010

[12] M.Hendrickson, "the state of Location Based Social Networking on the iPhone,"http;//techcrunch.com2008/09/28/the state of location based social networking on the iPhone, 2008.

[13] P.Mohan, V.N.Padmanabhan and R.ramjee," Nericell:Rich monitoring of road and traffic conditions using mobile smart phones," proc.Sixth ACM Conf. Embedded Network Sensor Systems, 2008.

[14] B.Schilit, J hong and M.grutesar,"Wireless location Privacy Protection", computer, vol.36

[15] F.Grace, "Stalker Victims Should Check for GPS," http;//www.cbsnews.com

[16] Xiao Pan, JianLiang Xu, Xiaofeng Meng."Protecting Location Privacy Against Location Dependent Attack in Mobile devices". ACM 2008.