EasyChair Preprint

№ 1249

A Location Privacy Preserving Model based on Geohash

Wei Xiang

June 30, 2019

A Location Privacy Preserving Model based on Geohash

Wei Xiang

(Invited Paper)

Abstract:With the rapid development of location-aware mobile devices, location-based services have been widely used.When LBS (Location Based Services) bringing great convenience and profits, it also brings great hidden trouble, among which user privacy security is one of them.The paper builds a LBS privacy protection model and develops algorithm depend on the technology of one dimensional coding of geohash geographic information. The results of experiments and data measurements show that the model has good performance in avoiding attacking from the leaked information in a continuous query with the user's background knowledge.

Key Words: Location-based service; Geohash coding; Privacy protection model;k-anonymity

I. INTRODUCTION

With the increasing popularity of intelligent terminal application, wireless communication technology and positioning technology,LBS became very popular in Point of Interest （POI).When querying POI, users are usually required to submits its own latitude and longitude information and query content.In this process, there is a risk of user privacy leakage.Enjoying location services and privacy protection is a contradiction, the researchers is committed to find a balance between providing superior location services and user privacy protection.In the current research,it is mainly to change the reality location information to protect privacy.The idea is that publishing false positions, which is dummy; Spatial cloaking,the user's accurate positioning is represented by the fuzzy area.However, these methods still do not prevent attack from query request.

In this paper, a new user privacy protection model is proposed, which uses Geohash one-dimensional coding to achieve fuzzy location.On the LBS service provider, it needs to add a Geohash encoding to the POI record.The results of experiments and data measurements show that the model has good performance in avoiding attacking from the leaked information in a continuous query with the user's background knowledge.

II. TECHNICAL BACKGROUND

2.1 Definition of location privacy

location information refers to the geographical location of mobile users at a certain time.

Location {x, y, t}.(x is longitude, y is latitude, and t is time stamp)

2.2 Location-based service request

Location-based service request refers that the users send request to LBS.

Request {Uid , Location, Query}.(The Uid represents the unique identity of the user,it can be the user's phone number or user name;Location represents the user's Location information, which can be either a single point of location information or a region;Query refers to a user's service request or type of POI, such as "Query for the nearest cinema")

2.3 k-anonymity location privacy protection technology

Most of the space hiding schemes are based on k-anonymity.The idea is to put the user and at least k-1 other users together constitute an anonymous area, so that the probability of the user's real identity being recognized is no more than 1/k.

2.4 Mobile peer-to-peer environments(p2p)

Users broadcast query requests in P2P networks and mobile users spontaneously

organize into anonymous groups by sharing locations after receiving the requests.When a user in an anonymous group is ready to initiate a location service request, it will generate a hidden space based on the location information of other users in the group to initiate a request to the LSP in place of its exact location.

2.5 Geohash

Geohash is an address code that reduces two-dimensional latitude and longitude information to a unique one-dimensional string for each point on earth with a given precision.GeoHash has the following characteristics:GeoHash uses a string to represent longitude and latitude coordinates;GeoHash does not represent a point, but a region;The GeoHash code prefix can represent larger areas. For example, wx4g0ec1, whose prefix wx4g0e indicates a larger range that includes the encoding wx4g0ec1.

III. GEOHASH ENCODING PROCESS

Geohash encoding process has three steps:

(a)Convert latitude and longitude to binary.For example,there is a point(39.923201, 116.390705).The latitude range (-90, 90).The latitude 39.923201 is within the scope of (0,90),so we get first result 1;Then latitude 39.923201 is less than 45,so we get second result 0 and so on.This is followed by a binary representation of latitude.As TABLE I shows.

TABLE I

	latitude range	39.923201
1	(-90,90)	1
2	(0,90)	0
3	(0,45)	1
4	(22.5,45)	1
5	(33.75,45)	1
6	(39.375,45)	0
...	...	...

Finally, the binary representation of latitude is obtained as follows: 10111000110001111001

Similarly, the binary representation of longitude 116.390705 can be obtained as follows: 11010010110001000100

(b)Merge latitude and longitude binary: the merge method is to combine longitude and latitude binary according to parity bits:

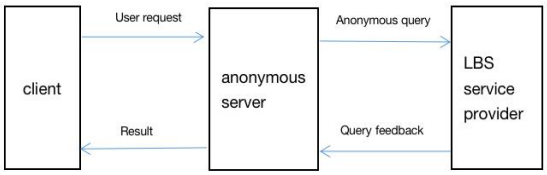
11100 11101 00100 01111 00000 01101 01011 00001

(c)According to Base32,the result of the combined binary encoding is: wx4g0ec1.

IV. SYSTEM MODEL

The system model is shown in GRAPH I and consists of three parts:client,anonymous server and LBS service provider.

GRAPH I



Client refers to intelligent terminal with positioning function, including mobile phone, tablet computer and applications.Anonymous servers are intermediate servers that connect the client and the LBS service provider,which can manage to Geohash code and deal with other processing of the user's location data.LBS service provider refers to the application system that provides POI query.

This model has two advantage：If it is risky to determine whether the LBS service provider is trustworthy, users do not contact with the service provider directly, which reduces the risk of privacy disclosure;All privacy protection coding can finish on anonymous servers which can reduce the load on mobile clients.

V. SYSTEM PROCESS

Users start sending service requests, such as "find a movie theater three kilometers away." The user then submits four parameters via the application to the anonymous server: latitude 81.48374; Longitude: 113.84728; Search range: 3000; Target point (POI) : "cinema".Then start Geohash coding：

Input:Latitude, longitude, coded bits;

Output:Geohash encoding results;

Define the base32 coding character BASE32←‘ 0123456789bcdefghjkmnpqrstuvwxyz’;

Defines an array that converts binary to decimal Numbers bits←array(16,8,4,2,1);

Define the initial latitude interval lat←array(-90.0, 90.0);

Define the initial interval of longitude lon ←array(-180.0, 180.0);

Initialization count constants bit, ch, I are 0;

The switch variable is_even for the flag that initializes the latitude and longitude parameters is 1;

According to the geohash coding rules, the longitude is divided twice, and then latitude and longitude are divided alternately. According to the interval of the value of longitude and latitude, 0 or 1 is taken to merge the binary values of longitude and latitude, and 1 bit base32 encoding is generated according to every 5 bits of binary number.

```
while i<deep
{
 if is_even
 {
 mid ←lon/2
 if longitude >mid
 {
ch ← Take the value of bits array and the result of the original ch variable or operation. The bits array starts from the following table 0, and add 1 to the following table after each processing of a longitude or latitude value. When the subscript reaches 5, the subscript value is set to 0;
 lon ← (mid, lon[1])
}
 if longitude < mid
 {
 lon ← (lon[0], mid)
 }
}
 if is_even
{
 mid ←lat/2
 if latitude>mid
{
 ch ← Take the value of the bits array and the result of the original ch variable or operation. Add 1 to the subscript of the bits array. Add 1 to the following table after each processing of a longitude or latitude value.
lat ← (mid, lat [1])
 }
if latitude< mid
{
lat ← (lat [0], mid)
}
}
```

Every 5 bit binary number is converted into 1 bit base32 encoding.

Determine whether the subscript variable bit of bits array is equal to 4. If it is less than 4, then bit plus 1; otherwise, it means that 5 bits of latitude and longitude have been processed. At this time, generate 1 bit base32 encoding, and reset i plus 1, bit,ch.

After the Geohash coding is completed, the anonymous server sends the results to the LBS service provider for POI query. The LBS service provider returns the query results to the user, and the location of the POI data is displayed in longitude and latitude.

VI. EXPERIMENT

The experiment was implemented by Php+Mysql. The hardware environment was Q9500 2.83GHZ processor and 3GB memory in Windows10, experiments are carried out with real data set.The real data used the Gowalla data set, with a total of 112,491 users and 508,796 mobile locations from 2009 to 2010.1000, 10000 and 100000 inquired POI are randomly selected from the samples to form the set and the Geohash code 5 bits and Geohash code 6 bits are selected for precision respectively to observe the access efficiency of the longitude and latitude based POI information points and the Geohash code. The comparison is shown in the table below.

Table II

POI query sample number	Time based on latitude and longitude （s）	Time based on Geohash （s）
1000	189.193	0.186
10000	1973.837	1.891
100000	18934.647	18.243

Table III

POI query sample number	Time based on latitude and longitude （s）	Time based on Geohash （s）
1000	166.893	0.183
10000	1784.368	1.786
100000	15637.731	16.779

It can be seen from the analysis that: Query requests encoded through Geohash are faster than traditional longitude and latitude based queries;Coding accuracy will affect the speed of the whole query. The existing law is that the higher the precision, the faster the speed.

VII. CONCLUSION

In order to solve the privacy security of LBS service, a new model is proposed. Under the premise of distrust of LBS service provider, anonymous server is introduced to physically cut off the direct connection between the user and the LBS service provider. In addition, Geohash is used to encode, and the two-dimensional longitude and latitude information is reduced to a one-dimensional string, and the encoding comparison operation is used instead of the traditional GPS floating point calculation, which improves the calculation speed and reduces the query time.

REFERENC

[1] Hong J I, Landay J A. An architecture for privacy-sensitiv eubiquitcomputing[C] //Proc of the 2nd Interna Tional Conference on Mobsystems, Applications, and Services. New York: ACM, 2004: 177-18.

[2] Yiu M L, Jensen C S, Huang X, et al. SpaceTwist managing the trade-offs among location privacy, query performance, and query accuracy in mobile services[C] // Proc of the 24thInternational Conference on Data Engineering. 2008: 366-375.

[3] Gruteser M, Grunwald D. Anonymous usage of location-based services through spatial and temporal cloaking[C]//Procof the 1st International Conference on Mobile Systems, Applications and Services. New York:ACM, 2003: 31-42.

[4] Ni Emeyerg. Geohash tips & tricks[EB/OL]. (2015-05-21). http: //geohash.org/-site/tips. html.

[5] Dimiduk N, Khurana A. HBase in action[M]. New York: Manning Publications Co, 2012. 203-235.

[6] Brinkhoff T. A Framework for Generating Network-Based Moving Objects[J].
GeoInformatica, 2002, 6(2):153-180.

[7] Morton G M. A Computer Oriented Geodetic Data Base and ANew Technique in
File Sequencing[J]. Physics of Plasmas, 2015, 24(7):159-173.

[8] Machanavajjhala A, Kifer D, Gehrke J. L-diversity: Privacy Beyond k -anonymity[C]// IEEE International Conference on Data Engineering. IEEE, 2006:24-24.

[9] Chow C Y, Mokbel M F, Liu X. Spatial Cloaking for Anonymous Location-based
Services in Mobile peer-to-peer Environments[J]. GeoInformatica, 2011, 15(2): 351-380.

[10] Chow C Y, Leong H V, Chan A T S. Distributed Group-based Cooperative
Caching in a Mobile Broadcast Environment[C]//ACMInternational Conference
on Mobile Data Management. ACM, 2005: 97-106.

[11] Li T C, Zhu W T. Protecting User Anonymity in Location-based Services with Fragmented Cloaking Region[C]// IEEE International Conference on Computer Science and Automation Engineering. IEEE, 2012:227-231.

[12] Niu B, Zhu X, Li W, et al. Epcloak: An Efficient and Privacy-preserving Spatial
Cloaking Scheme for Lbs's[C]// IEEE International Conference onMobile Ad
Hoc and Sensor Systems (MASS). IEEE, 2014: 398-406.

[13] Palanisamy B, Liu L. Attack-resilient Mix-zones Over Road Networks: Architecture and Algorithms [J]. IEEE Transactions on Mobile Computing, 2015, 14(3): 495-508.

[14] Gruteser M, Grunwald D. Anonymous Usage of Location-Based Services Through Spatial and Temporal Cloaking[C]// International Conference on Mobile Systems. 2003:31-42.