AN ADAPTIVE HOMOMORPHIC CRYPTOGRAPHY APPROACH FOR LOCATION BASED QUERIES

Hazar Naziya Shaik1, Aravapalli Rama Satish2

1 Post Graduate Student, Department of CSE, DVR & DHS MIC College of Technology, Kanchikacherla
2 Associate Professor, Department of CSE, DVR & DHS MIC College of Technology, Kanchikacherla

ABSTRACT:

The widespread utilization of Location Based Services has been rising. It leads to two difficulties, one, protection of user personal location and two, protection of data in location server. Above two objectives were accomplished by two-stage approach. The first objective comes when user wants to query the location database to get location information, which is known as Points of Interest (POI), and doesn’t want to reveal his/her location to the location server. The second issue concentrates on protection of data in location server. Generally, a location server gives results for custom-made queries to users as per their locations. First step of Two-Stage approach is Oblivious Transfer Protocol and second step is Private Information Retrieval Protocol. These two stages achieved client and server securities, but did not concentrate on time parameter. So, to give quick responses to the user, in this paper, a new idea is proposed, i.e. the Adler-32 cryptographic technique used along with Oblivious Transfer phase.

Keywords: Location based Queries, Adler-32 Checksum, Oblivious Transfer, and Private Information Retrieval

[1]INTRODUCTION

Location Server provides Location Based Service (LBS) [11] for user to process Location Based Queries [10]. Location Server is a server which maintains location database. Location Database maintains Public Grid. In public grid every location on earth is noticed. For Example, in Figure 1 Krishna District public grid is shown.

A LBS provides utility, information and entertainment services, used by portable mobile devices and are operating through Mobile network. All these services works on user’s geographical position. Location based query is generated by user to search for his/her points of interest (POI).

Due to increase in mobile device usage, privacy of user and protecting location server’s data became major issue. This paper, deals with, one, preserving privacy of user by hiding his location details from location server, two, protecting location server’s data, by revealing only related data to user i.e., Location server offers users of various queries and it should be able to send exact POI to each query i.e. no user should get unnecessary POIs. To achieve these
two goals two-stage approach is used: one, Oblivious Transfer Phase is utilized for Server's data security, two, Private Information Retrieval Protocol is utilized for user's Location privacy. The frame work for Location Based Queries shown in Figure 2.

In recent times the number of client queries made against LBS server is increased in integral number of times.

[1.1]RELATED WORK

The primary solution for the issue client protection was proposed by Beresford, in which the client's protection is set up by continuously changing the client's name or blend zone [3]. It can be demonstrated that, the incessant changing of the client's name gives little assurance to the client's protection. A later examination of the blend zone methodology has been connected to Road Networks [14]. But this solution should maintain control over users of blend zone which is difficult in practice.

A complementary technique to the blend zone methodology depends on k-Anonymity [15]. The major goal of k-Anonymity is, protecting data when circulating sensitive data records [4]. This was accomplished by speculation and concealment calculations to guarantee
that a record couldn't be recognized from \((k - 1)\) different records. By utilizing, a trusted anonymizer to give obscurity to the area information, a user can't be recognized from \((k - 1)\) different clients. Even when sufficient care is taken, \(k\)-anonymity can still be vulnerable to attacks [8].

An improved technology to \(k\)- Anonymity [8] is trusted anonymizer methodology, which permits the users to define their level of privacy based on value of \(k\). User should estimate the value of \(k\), if user chooses small value then efficiency will be increased and if user chooses large value then privacy will be improved. So, identifying value for \(k\) became unnatural. To reduce the overhead of user, Feeling Based Privacy [12] introduced. In this concept, rather than choosing \(k\), users will select their cloaking region [6], and system calculates number of cells for the cloaking region based on popularity of the area, the popularity of the public region (e.g. shopping mall), measured using entropy based on its visitor's footprints inside it. Popularity is then used as the user’s desired level of privacy protection. This technique could not applicable for areas like house, office; where footprints are very less.

Next research was done to create confusion against path and position of the user. Hoh and Gruteser [13] have proposed the concept of Path Confusion. This concept is to fade the user path by adding some false points to the original path of user, which enable the server to find user location. Position confusion [16] is related to provide more privacy. The trusted Anonymizer, groups users based on their cloaking regions. So that LS find difficulty to identify an individual. CR Techniques used in this concept contain semantic information of user which reveals user’s geographical information.

As solutions based on central anonymizer which are not practical. Hashem and Kulik proposed a scheme, in which groups of trusted users construct an Ad-hoc network [19] and querying LS was allotted to a single user. If the user goes offline, this responsibility will transfer to another user. But in real time constructing ad-hoc networks always not possible.

Another alternative method for Central Anonymizer is Dummy Records [20]. The concept of this methodology is to confuse the server by sending some other locations along with original location. The main drawback in this method is, increased overhead of choosing false locations.

All previous discussed problems were solved with the introduction of Private Information Retrieval (PIR) scheme. The basic idea of PIR [5] is to provide security for user queries. This scheme allows query processing, data fetching from database without disclosing user’s location details. PIR provides only user security, as well as user; server data should be protected, so along with PIR, Oblivious Transfer protocol also introduced in Russels Paulet's solution [1]. Although user and server securities are achieved, calculated experimental time is too large. To reduce experimental time, in Oblivious Transfer [2] initialization phase Hash functions are replaced with Adler-32 cryptographic technique [18].

[1.2] MOTIVATION

This paper has major performance and time efficiency with respect to Russell Paulet’s [1] approach. This paper deals with two-stage approach which works on two- protocols. Major contributions made in this paper are:
AN ADAPTIVE HOMOMORPHIC CRYPTOGRAPHY APPROACH FOR LOCATION BASED QUERIES

- Reduced Key Size
- Reduced experimental time

[2] PROPOSED APPROACH

A Two-Stage approach is used and each stage is operated through a protocol. First we discuss Notations, then System Architecture.

[2.1] NOTATIONS

The cyclic group $G_0$ is a multiplicative subgroup of finite field $F_p$ where $p$ is a large prime number. $g_0$ is a generator of multiplicative subgroup $G_0$ with order $q$, where $q$ is prime number and $(p-1)$ divisible by $q$. $G_1$ is multiplicative subgroup of finite field $F_p$. $g_1, g_2$ are generators of multiplicative subgroup $G_1$ with same order $q'|(q-1)$. Multiplicative subgroups $G_0, G_1$ are linked together to generate key. The generated key is in the form of $g_x^{y \oplus z}$, where $x, y$ are variable integers.

The operator $\oplus$ is the exclusive OR operator, $a || b$ to be the concatenation of $a$ and $b$. We require for security reasons, that $|q'| = 1024$ and $p$ has the form $p = 2q' + 1$. The parameters $G_0, G_1, g_0, g_1, g_2, p, q'$ are fixed for the duration of a round of our protocol and be made publicly accessible to every entity in our protocol.

[2.2] SYSTEM ARCHITECTURE

The architecture of this system is shown in Figure 3. This architecture has three entities, each entity has a role. These roles are shown in Table 1.

![System model for privacy preserving data protection](image)

<table>
<thead>
<tr>
<th>S. No</th>
<th>Entity</th>
<th>Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Users</td>
<td>Access location data</td>
</tr>
<tr>
<td>2</td>
<td>Service Provider</td>
<td>Bridge for LS &amp; user</td>
</tr>
<tr>
<td>3</td>
<td>Location Server</td>
<td>Maintains location</td>
</tr>
</tbody>
</table>

Table. 1: Roles of Entities
For example, if user will make a query like “Banks in my location?” Next, Service provider purpose is to maintain connection between LS and User. Then, Location Server which maintains database of POI records \((r_i)\) for \(1 \leq r_i \leq p\), (where ‘p’ is a large prime number), which produces output of POI records. Each POI record contains GPS co-ordinates of location and description or name of the location.

Users felt that LS, SP as a single Server, but in practice, SP and LS are not a single server, they are not same. Users will make location based queries to location server by utilizing Location Based Service.

**[3] PROTOCOLS**

This section explains Protocol structure and Protocol Description. The over view of Protocols shown in Figure 4.

**[3.1] PROTOCOL STRUCTURE**

Structure of protocols described as follows: Initially, Oblivious Transfer Protocol will be executed. The output of this phase is location of private grid ID and key. Next, PIR Protocol will start execution, takes output of OT protocol and retrieves particular grid of encrypted POI. Further, Decryption performed on this POI to get user understandable result.

**[3.1.1] Oblivious Transfer:**

This protocol is composed of two phases,

i) Initialization phase,

ii) Transfer Phase.

i) Initialization Phase:

Initialization phase is done by Location Server. LS maintain a database of location data and two dimensional key matrix \(K_{m \times n}\). Each element in key matrix is represented by \(k_{i,j}\). Elgamal Cryptography [7] chooses a key \(k_{i,j}\) from Key matrix \(K_{m \times n}\), which encrypts unique POI record. Along with Key Matrix, Location server maintains public grid which is collection of cells with equal number of POI records in each cell. A POI record information of location including GPS Co-ordinates of location, name of location and description on location.
ii) Transfer Phase:

Transfer Phase is built on six algorithms: QG₁, RG₁, RR₁, QG₂, RG₂, and RR₂. First three algorithms computed in Oblivious Transfer Phase, and Last three algorithms computed in Private Information Retrieval Phase. These algorithms are described below:

a) QueryGeneration₁ (Client System) (QG₁):
This algorithm fetches the query from cloaking region and append secret key to maintain uniqueness, to perform this, algorithm takes inputs i, j indices of key from key matrix, and the dimensions of the Key matrix m, n, then gives outputs a query Q₁ and secret key s₁, denoted as \((Q₁, s₁) = QG₁(i, j, m, n)\).

b) ResponseGeneration₁ (Server System) (RG₁):
This algorithm generates response of k grids out of p grids from Public Grid by Adaptive Oblivious Transfer, to do this, algorithm takes input as the key matrix \(K_{mn} \times n\), and the query \(Q₁\), then produces outputs as response \(R₁\), denoted as \((R₁) = RG₁(K_{mn}, Q₁)\).

c) ResponseRetrieval₁ (Client System) (RR₁):
This algorithm receives k cells out of p cells of Public grid, to do this, algorithm takes inputs indices i, j, the dimensions of the Key matrix m, n, the query \(Q₁\) and the secret \(s₁\), and the response \(R₁\), and outputs a cell key \(k_{ij}\) and cell-id \(ID_{ij}\) in public grid, denoted as \((k_{ij}, ID_{ij}) = RR₁(i, j, m, n, (Q₁, s₁), R₁)\). This will be executed after PIR phase.

[3.1.2] Private Information Retrieval Phase:

In this stage, user executes communicational efficient PIR [9] to retrieve a single grid from private grid. PIR retrieves total database from server (in present scenario retrieves data from OT Protocol) and then process the user query to get exact information required by user. This protocol doesn’t allow user to retrieve unrelated information, this will retrieve only one grid per query. By this way it provides security for server data and user queries. To process query and give exact information, PIR executes three algorithms. To process PIR algorithms PIR makes use of set of prime powers \(S\). They are explained below:

d) QueryGeneration₂ (Client System) (QG₂):
This algorithm takes output of RR₁ and generates user location, to generate id, algorithm takes inputs the cell-id \(ID_{ij}\), and the set of prime powers \(S\), and outputs a query \(Q₂\) and secret key \(s₂\), denoted as \((Q₂, s₂) = QG₂(ID_{ij}, S)\).

e) ResponseGeneration₂ (Server System) (RG₂):
This algorithm generates required cell information from k cells; and takes inputs the database \(D\), the query \(Q₂\), and the set of prime powers \(S\), and outputs a response \(R₂\), denoted as \((R₂) = RG₂(D, Q₂, S)\).

f) ResponseRetrieval₂ (Client System) (RR₂):
This algorithm retrieves exact cell of POI records along with secret key, to do this, algorithm takes inputs the cell-key $k_{ij}$ and cell-id $ID_{ij}$, the query $Q_2$ and secret $s_2$, the response $R_2$, and outputs the data $d$, denoted as $(d) = RR2(k_{ij}, ID_{ij}, (Q_2, s_2), R_2) = e$

This transfer phase repeatedly fetch records form DB to fulfill the requirement of User.

**[3.2] PROTOCOL DESCRIPTION**

Main aim of this paper will be satisfied by receiving POI records from location Database without compromising privacy of users and data of server.

**[3.2.1] OBLIVIOUS TRANSFER:**

In two-stage approach (Figure. 4), Oblivious Transfer is first stage protocol. Oblivious Transfer Protocol uses $k$ out of $N$ Adaptive Oblivious Transfer Protocol [17] to retrieve $k$ records out of $N$ records of database by which server security is initiated. As discussed earlier, by performing three algorithms ($QG_1, RG_1, RR_1$) location ID, and corresponding secret $(s)$ are generated. This combination of cell ID and key are used in second-stage in which Private Information Retrieval Protocol is used, by which exact POI cell retrieved.

Oblivious Transfer protocol further classified into two phases:

i) Initialization

ii) Transfer Phase

i. Initialization Phase:

Consider a user ‘u’ from set of users ‘U’, creates a cloaking region [6] to start initialization phase. All queries of user should make with respect to the cloaking region. Cloaking region contains user queries and user location co-ordinates. With this information SP converts Cloaking region to private grid means, it appends all available locations of user’s present location, which will be submitted to server later (Figure. 5). Server could not able to identify user location because each cloaking region is collection of cells which maintain equal number of POI records and cloaking region is in uncertain format to confuse server. After receiving the private grid of user, public grid of server will superimposes it (Figure. 6).

Then as shown in Algorithm Initialization, server starts encrypting each POI record in each cell $Q_{ij}$, (where $Q_{ij}$ represents a cell in Public Grid). Encrypted data and its symmetric key stored in small database table in its corresponding cell in public grid $P$ (Figure. 7). Encrypted form of POI is a number. Each encrypted data will represent with a integer ‘C’ [1] with respect to cloaking region. The value of ‘C’ is quite large, which takes more time to perform mathematical operations. So, to reduce the value of $C_i$, to get small integer, Transfer phase performs $e= C_i \mod \pi_i$, where $\pi_i$ is set of prime powers $\pi_i=\pi_i^{ci}$, such that $C_i< \pi_i$ are selected and $c_i$ in the exponent must be small for the protocol work efficiently.
Then, while fetching record, server has to find small integer 'e', it uses Chinese Remainder Theorem to find e, where e=C_i (mod π_i), for each C_i. Then 'e' represents the database.

Algorithm: Initialization:

Input: X_{1,1} \ldots \ldots X_{m,n}, where X_{i,j} = ID_{x_{i,j}} k_{i,j}
Output: Y_{1,1} \ldots \ldots Y_{m,n}

Step 1: K_{i,j} \leftarrow K_{i,j}g_1^R_i g_2^C_j, for 1 \leq i \leq n and 1 \leq j \leq n, where R_i and C_j are randomly chosen
Step 2: Y_{i,j} \leftarrow X_{i,j} \oplus A(K_{i,j}), for 1 \leq i \leq n and 1 \leq j \leq n, where A is Adler-32 check sum
Step 3: return Y_{1,1} \ldots \ldots Y_{m,n} \{Encryptions of X_{1,1} \ldots \ldots X_{m,n} using K_{i,j}\}

In the above algorithm 'A' represents Adler-32 check sum, where by using Adler-32 checksum, key size will be reduced because Adler 32 checksum will produce 16-bit result which is encrypted key. Whereas, Hash Function Produce 32-bit result. By reducing key size, encryptions and mod operations can be easily performed. So, that the user can receive results.
more quickly. After completing Initialization phase, SP can fetch grid references from database by performing Transfer Phase.

ii. Transfer:

The main goal of Transfer Phase [1] is to fetch single record from location server. To execute this phase, transfer phase takes help of Two-Dimensional oblivious Transfer [23] and Adaptive Oblivious Transfer [17]. In this phase using Elgamal Encryption [7] coordinates of user will be encrypted.

After execution of Initialization Algorithm, once per one query, then user query will start execution. Then Transfer Phase takes output of previous phase (i.e. Initialization) and assassinates QG1, RG1 and RR1 algorithms. After completion of OT protocol, cell id of user required POI and its key ID, key will send to PIR protocol.

Figure 7: Association of Public and Private grids

[3.2.2] PRIVATE INFORMATION RETRIEVAL PHASE

PIR performs remaining three algorithms by receiving output of RR1, those are QG2, RG2, RR2 algorithms to get adequate database integer (C). After processing RR2 output will be database integer 'e', to get C corresponding to 'e', should perform C = log(b)e. This C represents required database. Then user's SP gets appropriate POI record in encrypted form C. On this Elgamal Decryption will be performed to get user understandable result, then result will be sent to user.

[4] SECURITY ANALYSIS

This section describes how two-stage approach achieves user and server securities.

[4.1] User's Security:

In the Two protocols, Oblivious Transfer Protocol provides security for user and Location Server. Before Oblivious Transfer's initialization phase, cloaking region created
with user’s location and query details. To this cloaking region, all available locations in user location will be appended to confuse LS to find User Location.

Then cloaking region will converted to a grid known as private grid. This grid will be encrypted before submitting to Location server by Initialization, Transfer Algorithms of Oblivious transfer. By sending encrypted for user details, the server cannot recognize user.

Consider the key format $\mathcal{G}_a = \mathcal{G}_j \cdot \mathcal{G}_j$ is protected by discrete logarithm problem. As k out of N adaptive Oblivious Transfer is used then only k grids retrieved from server in Transfer Phase. By this server's data security is initiated.

[4.2] Server's Security:

Server security achieved by both Oblivious Transfer Phase and Private Information Retrieval Phase.

Initially, Oblivious Transfer's Transfer Phase works on k out of N Adaptive Oblivious Transfer, which retrieves k grids out of N grids from Location Database. By which security for data initiated.

Private Information Retrieval protocol used to fulfill server's security requirement. Consider PIR protocol's three algorithms QG$_2$, RG$_2$, RR$_2$. QG$_2$ algorithm done in client system to gain user private grid id. RG$_2$ will execute on QG$_2$ algorithm's output and process user query to identify user interested POI and generate required POI's grids. RR$_2$ algorithm receives output of RG$_2$ and retrieves required encrypted POIs from received grids. To decrypt these POIs, symmetric key received after completion of all these phases. So that no decryption possible in between the process by RR$_2$. Then client system able to receive only that information which it is actually required. Hence, server's data is protected.

[5] EXPERIMENTAL RESULTS

Our application designed for LBS can be applicable for: portable mobile devices. The desktop machine that was utilized for testing is outfitted with an Intel Core 2 Duo E8200 2.66GHz processor and 2GB of RAM. The usage on this stage was composed utilizing Visual C++ under the Windows XP working framework. We utilized the Number Theory Library (NTL) for calculations obliging large whole numbers.

While comparing previous solution with current solution, time utilization is reduced. This is happened by changing SHA-1 hash function in Oblivious Transfer Protocol's Initialization algorithm by Alder-32 check sum function, by which key size has been reduced. So that calculating logarithms, exponents became faster (Graph 1). Average performance time difference for both solutions shown in Table 2.
Graph. 1: Comparison between current solution and Russell Paulet's solution.

The represented table and graph show difference between Russell Paulet's solution and current solution.

<table>
<thead>
<tr>
<th>Performance Time</th>
<th>Current Solution (avg in sec)</th>
<th>Russell Paulet's Solution (avg in sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.033</td>
<td>0.347</td>
</tr>
</tbody>
</table>

Table. 2: Compares time in both solutions

6. CONCLUSION

A Location based query problem was considered i.e., whenever a user queries a location database, server should not able to identify the user's location although user location already saved in server. To achieve this, Oblivious Transfer phase is utilized, and PIR is utilized for Protection of data in location server, i.e., user should get appropriate and related POIs. To improve efficiency SHA-1 hash function is replaced with Adler-32 check sum by which key size is reduced to 16-bit. By this experimental time is reduced.

REFERENCES


AN ADAPTIVE HOMOMORPHIC CRYPTOGRAPHY APPROACH FOR LOCATION BASED QUERIES


Author[s] brief Introduction

We Hazar Naziya Shaik from Nandigama and A. Rama Satish from Vijayawada are working currently on Location Based Services. We are interested in Data Mining and Networking.