International Journal of Research
Available at https://edupediapublications.org/journals
e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 05 Issue-01
January 2018

Secured Location Aware Queries with Sensible Knn Keywords

Somayajula Lakshmi Prasanna & Dasari Vinay Kumar

M.Tech (CSE), Department of Computer Science & Engineering, NRI Institute ofTechnology, Guntur, A.P.

Assistant Professor, Department of Computer Science & Engineering, NRI Institute of Technology, Guntur, A.P.

ABSTRACT--*In mobile communication, spatial queries pose a serious threat to user location privacy because the location of a query may reveal sensitive information about the mobile user. In this paper, we study approximate k nearest neighbor (kNN) queries where the mobile user queries the location-based service (LBS) provider about approximate k nearest points of interest (POIs) on the basis of his current location. We propose a basic solution and a generic solution for the mobile user to preserve his location and query privacy in approximate kNN queries. The proposed solutions are mainly built on the Paillier public-key cryptosystem and can provide both location and query privacy. To preserve query privacy, our basic solution allows the mobile user to retrieve one type of POIs, for example, approximate k nearest car parks, without revealing to the LBS provider what type of points is retrieved. Our generic solution can be applied to multiple discrete type attributes of private location-based queries. Compared with existing solutions for kNN queries with location privacy, our solution is more efficient. Experiments have shown that our solution is practical for kNN queries.*

Key Terms--Location based query, location and query privacy, private information retrieval, Paillier cryptosystem, RSA

INTRODUCTION

The embedding of positioning capabilities (e.g., GPS) in mobile devices facilitates the emergence of locationbased services (LBS), which is considered as the next "killer application" in the wireless data market. LBS allow clients to query a service provider (such as Google or Bing Maps) in a ubiquitous manner, in order to retrieve detailed information about points of interest (POIs) in their vicinity (e.g., restaurants,

hospitals, etc.). The LBS provider processes spatial queries on the basis of the location of the mobile user. Location information collected from mobile users, knowingly and unknowingly, can reveal far more than just a user's latitude and longitude. Knowing where a mobile user is can mean knowing what he/she is doing: attending a religious service or a support meeting, visiting a doctor's office, shopping for an engagement ring, carrying out non-work related activities in office, or spending an evening at the corner bar. It might reveal that he is interviewing for a new job or "out" him as a participant at a gun rally or a peace protest. It can mean knowing with whom he/she spends time, and how often. When location data are aggregated it can reveal his/her regular habits and routines - and when he deviates from them. A 2010 survey conducted for Microsoft in the United Kingdom, Germany, Japan, the United States, and Canada found that 94 percent of consumers who had used locationbased services considered them valuable, but the same survey found that 52 percent were concerned about potential loss of privacy.1 In this paper, we study approximate k nearest neighbor (kNN) queries where the mobile user queries the locationbased service provider about approximate k nearest points of interest on the basis of his current location. In general, the mobile user needs to submit his location to the LBS provider which then finds out and returns to the user the k nearest POIs by comparing the distances between the mobile user's location and POIs nearby. This reveals the mobile user's location to the LBS provider.

I. RELATED WORK

Current main techniques to preserve location privacy for LBS are as follows. Information access control: User locations are sent to the LBS provider as usual. This technique relies on the LBS provider to restrict access to stored location data through rule-based polices. It supports three types of location-based queries: 1) user location queries (querying the location of a specific user or users, identified by their unique identifiers); 2) enumeration queries (querying lists of users at specific locations, expressed either in terms of geographic or symbolic attributes); 3) asynchronous queries (querying "event" information, such as when users enter or leave specific areas). This technique requires the LBS provider to maintain all user locations. It is vulnerable to misbehavior of the LBS provider. Mix

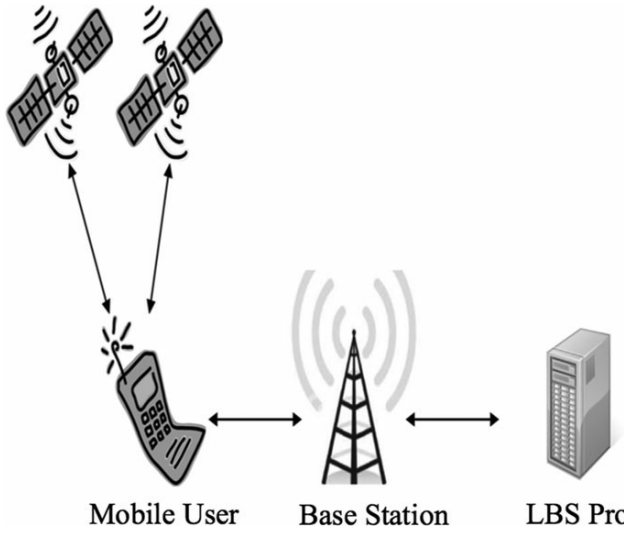
zone: A trusted middleware relays between the mobile users and the LBS provider. Before forwarding the location-based queries of the users to the LBS, the middleware anonymizes their locations by pseudonyms. The basic idea is: when a user enters a mix zone, the middleware assigns him a pseudonym, by which the user queries LBS. The communication between the user and the LBS is through the middleware and the pseudonym changes whenever the user enters the mix zone. Recently, the mixzone has been applied to road networks. This technique requires the middleware to anonymize user locations. It is vulnerable to misbehavior of the middleware.

K-anonymity: This technique ensures that a record could not be distinguished from k-1 other records. Instead of sending a single user's exact location to the LBS, k-anonymity based schemes collect k user locations and send a corresponding (minimum) bounding region to the LBS as the query parameter. The collection of different mobile user locations is done either by a trusted third-party between the users and the LBS, or via a peer-to-peer collaboration among users. Because k-anonymity is achieved, an adversary can only identify a location's user with probability no higher than 1=k. This technique relies on the third party or a peer user to collect different mobile user locations. It is vulnerable to misbehavior of the third party or the peer user. "Dummy" locations: The basic idea is when the mobile user queries the LBS, he sends many random other locations along with his location to the LBS provider to confuse his location such that the server cannot distinguish the actual location from the fake locations. Different from k-anonymity based schemes, this approach include fake or fixed locations, rather than those of other mobile users, as parameters of queries sent to the LBS provider. Fake dummy locations are generated at random, and fixed locations are chosen from special ones such as road intersections. Either way, the exact user locations are hidden from the service provider. Although this technique does not rely on any third party, the LBS provider can restrict the user in a small sub space of the total domain, leading to weak privacy.

II. METHODOLOGY

RSA: RSA is not a probabilistic encryption scheme. To transform RSA to a probabilistic encryption scheme, we need to add some random bits into the message m before

encrypting m with RSA. Optimal Asymmetric Encryption Padding (OAEP) [1] is a padding scheme often used together with RSA encryption.



Paillier Public-Key Cryptosystem: Paillier public-key cryptosystem [19] is composed of three algorithms as follows.

Key Generation: A user randomly chooses two large distinct primes p, q and an element g of $\mathbf{Z}^*_{N2}$ whose order is a nonzero multiple of **N = pq,** publishes the public key **pk=(N,g),** and keeps the private key **sk=(p,q)** secret.

Encryption: Given the public key pk of the user, one can encrypt a message m where m is a positive integer less than N by randomly choosing r from $\mathbf{Z}^*_{N2}$ and computing

$$c = E(m, pk) = g^m r^N \pmod{N^2} \quad (1)$$

where c is the ciphertext of m. Since r is randomly chosen, the ciphertext c of a message m is random. Therefore, Paillier cryptosystem is a probabilistic encryption.

Decryption: The user can decrypt the ciphertext c with the private key $\mathbf{s}^k$ by computing

$$m = \mathsf{D}(c, sk) = \frac{(c^\lambda (mod\ N^2) - 1)/N}{(g^\lambda (mod\ N^2) - 1)/N} (mod\ N), \quad (2)$$

where $\lambda = lcm(p-1, q-1)$.

Homomorphic Properties: Paillier cryptosystem has two homomorphic encryption properties as follows:

$$\mathsf{E}(m_1)\mathsf{E}(m_2) = \mathsf{E}(m_1 + m_2), \quad (3)$$

$$\mathsf{E}(m_1)^a = \mathsf{E}(am_1), \quad (4)$$

for any $m_1, m_2, m, a \in \mathbb{Z}_N$.

Suppose that $\mathsf{E}(m_i) = g^{m_i} r_i^N (mod\ N^2)$ for i=1, 2, it is easy to verify (3) and (4) because

$$\mathsf{E}(m_1)\mathsf{E}(m_2) = g^{m_1+m_2}(r_1 r_2)^N (mod\ N^2) = \mathsf{E}(m_1 + m_2),$$

$$\mathsf{E}(m_1)^a = g^{am_1}(r_1^a)^N (mod\ N^2) = \mathsf{E}(am_1).$$

Basic Private kNN Query Protocol: We assume that POI types are coded into 1; 2; . . .;m whichis published to the public and the mobile user U wishes tofind k nearest POIs of type t around his location. The user Uchooses a cloaking region CR with n n cells, where U islocated in the cell (i; j), and runs the kNN query protocolwith the LBS provider S, composed of Algorithms.

Algorithm 1. Query Generation (User)

Input: $CR, n, m, (i, j), t, k, pk = \{e, N\}$
Output: Q, s

1: Randomly choose two large primes p_1, q_1 such that N $p_1 q_1 > N$.
2: Randomly choose two large primes p_2, q_2 such that N $p_2 q_2 > N$, where $N_2^2 < N_1$.
3: Let $sk_1 = \{p_1, q_1\}, pk_1 = \{g_1, N_1\}$.
4: Let $sk_2 = \{p_2, q_2\}, pk_2 = \{g_2, N_2\}$.
5: For each $\ell \in \{1, 2, \ldots, n\}$, pick a random integer $r_\ell \in \mathbb{Z}$ compute

$$c_\ell = \begin{cases} \mathsf{E}_1(1, pk_1) = g_1{}^1 r_\ell^{N_1} (mod\ N_1{}^2) & \text{if } \ell = i \\ \mathsf{E}_1(0, pk_1) = g_1{}^0 r_\ell^{N_1} (mod\ N_1{}^2) & \text{otherwise} \end{cases}$$

where E_1 denotes the Paillier encryption algorithm w public key $pk_1 = \{g_1, N_1\}$ as described in Section 3.1.
6: For each $\ell \in \{1, 2, \ldots, m\}$, pick a random integer $r'_\ell \in \mathbb{Z}$ compute

Algorithm 2. Response Retrieval RR (User)

Input: $R = \{C_1, C_2, \ldots, C_n\}, s = \{sk_1, sk_2\}, sk = \{d\}$
Output: z

1: The user randomly chooses an integer $r < N$ and comp and sends to the server

$$w = r^e \mathsf{D}_2(\mathsf{D}_1(C_j, sk_1), sk_2)(mod\ N)$$

where $\mathsf{D}_1, \mathsf{D}_2$ are the Paillier decryption algorithm described in Section 3.1.
2: The server computes and replies to the user

$$v = \mathsf{D}(w, sk) = w^d (mod\ N)$$

where D denotes the RSA decryption algorithm as descri in Section 3.2.
3: The user computes

$$z = r^{-1} v (mod\ N)$$

4: **return** z

GENETIC PRIVATE K NEAREST NEIGHBOR QUERIES: In this section, we extend our basic solution to a generic construction of private kNN query protocol. Our generic solution considers a multi-dimension space where each POI is defined with location attributes (i; j) (where 1 _ i; j _ n) and multiple discrete type attributes (t1; t2; . . . ; tT) (where t_ (1 _ _ _ T) is an integer and 1 _ t_ _ m_). For example, a car park (encoded as 3) located at (9, 4) in the cell (8, 5) with "Mid" daily parking fee (encoded as 1) may be represented as d18;5;3;1 ¼ POI ð9; 4Þ, where daily parking fee can be categorized into "Low" (<$10), "Mid" ($10-$30) and "High" (>$30)).

CONCLUSION

In this paper, we have presented a basic and a generic approximate kNN query protocols. Security analysis has shown that our protocols have location privacy, query privacy and data privacy. Performance has shown that our basic protocol performs better than the existing PIRbased LBS query protocols in terms of both parallel computation and communication overhead. Experiment evaluation has shown that our basic protocol is practical. Our future work is to implement our protocol on mobile devices.

REFERENCES

[1] R. Rivest, A. Shamir, and L. Adleman, "A method for obtaining digital signatures and public-key cryptosystems,"

Commun. ACM, vol. 21, no. 2, pp. 120–126, 1978.

[2] R. Schlegel, C. Chow, Q. Huang, and D. Wong, "User-defined privacy grid system for continuous location-based services," IEEE Trans. Mobile Comput., vol. 14, no. 10, pp. 2158–2172, Jan. 2015.

[3] P. Shankar, V. Ganapathy, and L. Iftode, "Privately querying location- based services with SybilQuery," in Proc. 11th Int. Conf. Ubiquitous Comput., 2009, pp. 31–40.

[4] L. Sweeney, "k-anonymity: A model for protecting privacy," Int. J. Uncertain. Fuzziness Knowl.-Based Syst, vol. 10, pp. 557–570, 2002.

[5] S. Wang, X. Ding, R. H. Deng, and F. Bao, "Private information retrieval using trusted hardware," in Proc. 11th Eur. Symp. Res. Comput. Security, 2006, pp. 49–64.

[6] P. Williams and R. Sion, Usable PIR, in Proc. NDSS, 2008.

[7] W. K. Wong, D. W. Cheung, B. Kao, and N. Mamoulis, "Secure kNN computation on encrypted databases," in Proc. ACM SIGMOD Int. Conf. Manage. Data, 2009, pp. 139–152.

[8] B. Yao, F. Li, and X. Xiao, "Secure nearest neighbor revisited," in Proc. IEEE Int. Conf. Data Eng., 2013, pp. 733–744.

[9] X. Yi, R. Paulet, E. Bertino, and V. Varadharajan, "Practical k nearest neighbor queries with location privacy," in Proc. IEEE Int. Conf. Data Eng., 2014, pp. 640–651.

[10] M. L. Yiu, C. Jensen, X. Huang, and H. Lu, "SpaceTwist: Managing the trade-offs among location privacy, query performance, and query accuracy in mobile systems," in Proc. IEEE Int. Conf. Data Eng., 2008, pp. 366–375.

[11] M. Youssef, V. Atluri, and N. R. Adam, "Preserving mobile customer privacy: An access control system for moving objects and custom proles," in Proc. 6th Int. Conf. Mobile Data Manage., 2005, pp. 67–76.