Hindawi
Mobile Information Systems
Volume 2021, Article ID 5972486, 12 pages
https://doi.org/10.1155/2021/5972486

Research Article

Towards Region Queries with Strong Location Privacy in Mobile Network

Songtao Yang [1] **and Qingfeng Jiang** [2]

[1]*College of Computer Science and Information Engineering, Bengbu University, Bengbu 233030, China*
[2]*College of Computer Science and Engineering, Changshu Institute of Technology, Changshu 225500, China*

Correspondence should be addressed to Songtao Yang; yst@bbc.edu.cn

Received 26 April 2021; Revised 26 August 2021; Accepted 28 October 2021; Published 18 November 2021

Academic Editor: Daniel G. Reina

With the interaction of geographic data and social data, the inference attack has been mounting up, calling for new technologies for privacy protection. Although there are many tangible contributions of spatial-temporal cloaking technologies, traditional technologies are not enough to resist privacy intrusion. Malicious attackers still steal user-sensitive information by analyzing the relationship between location and query semantics. Reacting to many interesting issues, oblivious transfer (OT) protocols are introduced to guarantee location privacy. To our knowledge, OT is a cryptographic primitive between two parties and can be used as a building block for any arbitrary multiparty computation protocol. Armed with previous privacy-preserving technologies, for example, OT, in this work, we first develop a novel region queries framework that can provide robust privacy for location-dependent queries. We then design an OT-assist privacy-aware protocol (or OTPA) for location-based service with rigorous security analysis. In short, the common query of the client in our solution can be divided into two parts, the region query R_q and the content query C_q, to achieve location k-anonymity, location m-diversity, and query r-diversity, which ensure the privacy of two parties (i.e., client and server). Lastly, we instantiate our OTPA protocol, and experiments show that the proposed OTPA protocol is reasonable and effective.

1. Introduction

Location-based services (LBS) are one of the successful mobile applications in our daily life. Armed with the help of LBS, it will be easy for you whether you want to let others track your movements, find or get to somewhere, or simply know your current location and what is around you. Obliviously, LBS along with its corresponding applications greatly improve the public living style in terms of richness and diversity. However, problems of location privacy disclosure have not raised the concern of the public. For instance, some malicious attackers will track the trace of the location using the LBS. Attackers can monitor and identify goal-oriented people, but the goal-oriented people could not be aware of being tracked [1, 2]. In this case, researchers were beginning to engage in how to conceal location and identity of users.

In reality, the user can submit the points of interest (POIs) queries (e.g., "find the nearest mall") to the LBS provider like Google Map. To conceal location and identity, users could mask the query via an anonymity tool, such as k-anonymity and obfuscation. But the attacker can deduce the user's identity from the content of query, background knowledge, and the observation information if we just adopt a simple pseudonym to cloak the location and the identity [3]. To overcome these limitations, these proposed research schemes can be divided into three major types: (*a*) location k-anonymity, (*b*) location obfuscation, and (*c*) private information retrieval (PIR). However, these existing techniques cannot efficiently address the following two major problems in more detail: (*a*) most of the existing proposals assume that all anonymous participants are completely reliable. In contrast, the participants stay at the same level of security. Apparently, this assumption is unrealistic and

inconsistent. It is often questionable with the actual scenario. Collaborator may be disclosing the accurate location information or the accurate queries information, either directly or indirectly. (*b*) In fact, intermediary servers or query issuers obtain a large amount of redundant data during per-query. However, these data are employed in charging customers according to actual use, whether directly or indirectly, and they are valuable assets of the LBS server.

Consider an application scenario shown in Figure 1. Alice wants to get a discount list of this mall located in a certain area or obtain what movie will be released in the nearby cinema. Although there are a lot of POIs around her, she is only concerned with a certain category of these POIs information. For example, she issues a service request, "find the discount price of the mall which is away from my current location about 5 km", to trade with the LBS provider. According to LBS service mode, the NN of Alice is P6, where the set P1, P3, P4, P6, P8 represents some malls.

To our knowledge, previous schemes are not conducive to the embodiment of the commercial value of the LBS information. Hence, we will ask the following question: is it possible to address the two above-mentioned problems using OT-assist privacy-aware protocol? To answer this question, in this paper, we first develop a novel region query framework that supports the private location-dependent query. We then design an OT-assist privacy-aware protocol (or OTPA) for location-based service with rigorous security analysis.

The contributions can be summarized as follows:

(1) A novel region queries framework: We first developed a novel region queries framework that supports private location-dependent queries. Our framework achieves noncooperative privacy preserving via cryptographic techniques, and it does not require a trusted third party. We proposed a new fair exchanging pattern with semitrusted three parties, which includes an intersection with three subjects: users, location cloaking server, and LBS server. Assume that all the participants are semihonest in this architecture; they will honestly follow the protocol but they are curious to find out as much as information from the data that it receives and stores.

(2) An OT-assist privacy-aware protocol: We designed a privacy-aware query protocol, which guarantees the untraceability of user trajectory and unlinkability of the content. A common query is split into a region query and a content query in our solution. Further, we analyzed the user's privacy through theory analysis and demonstrated the effectiveness by experiments.

1.1. Roadmap. The rest of this paper is organized as follows. We reviewed the related work in Section 2. In Section 3, we presented some definitions and gave some terms. In Section 4, we introduced the region queries and designed a system model and expression. The proposed privacy-aware region queries and OTPA protocol are presented in Section 5, followed by the security analysis in Section 6 and the experiment evaluation in Section 7. Finally, we concluded the paper in Section 8.

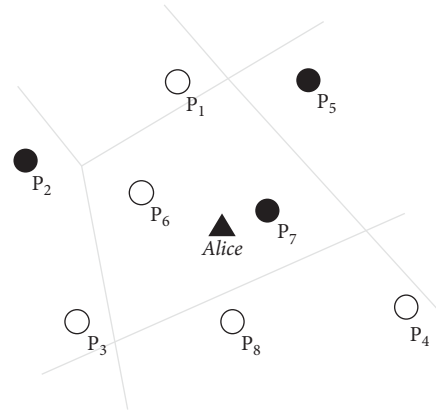


Figure 1: Application scenario of LBS.

2. Related Works

In numerous studies, the location *k*-anonymity [4–6] is always the predominant approach. The essence of location *k*-anonymity is that the probability of identifying the query user cannot exceed $1/k$, which is mainly focused on query privacy. Instead of sending the query to the LBS server, the user interacts with the anonymizer, which cuts off the association between user's identities and query contents to prevent the attacker from analyzing the user's sensitive information. However, *k* is not a representative of the actual location privacy of mobile users. In fact, these cloaking techniques based on the location *k*-anonymity metric could even be counterproductive and give the illusion of a higher location privacy level. Shokri et al. argued that the *k*-anonymity scheme is insufficient for protecting location privacy [7]. For example, if *k* users within an anonymous spatial region (ASR) are located in the same semantic location, the ASR guarantees the requester's query privacy but discloses their location privacy. On the other hand, if *k* users have similar query content and distribute in different locations, the ASR guarantees users location privacy and exposes the user's query privacy. Therefore, some researchers develop further studies for location diversity and context-aware and location semantics [8]. A complementary technique to the location *k*-anonymity is the location obfuscation technique. These location obfuscation techniques are achieved by deliberately reducing the resolution of the user's location to protect user privacy, namely, using a cloaking region instead of the user's actual location. To release ambiguous location is often used as a simple and effective technique [9–11]. Space Twist framework avoids the high computational cost and communication cost caused by the ASR. However, a lower resolution of location may cause coarse-grained service provided by the LBS server. The size of cloaking region is proportional to degree of privacy and is inversely proportional to the quality of service. Therefore, the adversary can deduce the approximate location of the user according to the

context of background environment, which means leading weak privacy [12]. Collusion of LBS leads to complexity of privacy preserving in real-world applications. The correlation between geographic data and social data leads to losing effectiveness for spatial-temporal anonymity technology. As a cryptography-based oblivious transfer method, private information retrieval (PIR) was also adopted to secure the location privacy [13, 14]. OT and PIR are similar: cryptographic protection against information disclosure. The methods which employed PIR protocol or OT protocol provide provable privacy guarantees against correlation attacks and eliminate the requirement for any trusted third party. Computational PIR-based approach utilizes a PIR protocol to implement a simple query pattern, which retrieves a specific database block from the LBS server without discovering which block is retrieved. However, it leads to a prohibitive computational cost and communication cost even for a small POIs databases. Therefore, secure hardware-aided PIR proven efficient is currently considered as a practical mechanism for PIR. Some cryptographic technologies (such as attribute based encryption [15–17] and data integrity checking [18–20]) have potential application in location privacy, which not only guarantees secure data share but also ensures remote data integrity [21].

3. Preliminaries

In this section, we present some definitions for follow-up work, including the framework, privacy-aware protocol, and privacy-aware queries in LBS.

Definition 1 (P). A point of interest (POIs) is a landmark or specific location that someone may feel useful or be interested in, such as a hotel, hospital, and school. It can be formalized into a triple set: $P = \langle l, c, i \rangle$. Here, we denote the $i-$ th POIs as P_i, where, $P_i[l]$ is location coordinates of a P_i, $P_i[c]$ is category of a P_i, $P_i[i]$ is service content of a P_i.

Definition 2 (R_q). Region query can be formalized as follows: $R_q = \langle R, k, m \rangle$. Here, R represents a geographic region illustrated by the query submitter. The k is the user-desired degree of anonymity. m is the user-desired number of different semantic locations within R.

Definition 3 (C_q). Content query can be formalized as follows: $C_q = \langle R', C' \rangle$. Here, R' represents the minimum area meeting users' privacy. C' is a subset of C. It is selected by the user. C is a comprehensive POIs taxonomy set.

Definition 4 ($H(l_i)$). Given an anonymous spatial region, a set of m location points $L = \{l_1, \ldots, l_m\}$. For any location in an anonymous spatial region, the Location Entropy is denoted as $H(l_i) = -\sum_{i=1}^{m} p_i^{(l_i)} \log_2 p_i^{(l_i)}$. Here, $p_i^{(l_i)}$ is the probability of user u_i locating in l_i.

Definition 5 ($H(q_i)$). Given an anonymous set, a set of n users $U = \{u_1, \ldots, u_n\}$. For any use in an anonymous set, the Query Entropy is denoted as $H(q_i) = -\sum_{i=1}^{m} p_i^{(q_i)} \log_2 p_i^{(q_i)}$. Here, $p_i^{(q_i)}$ is the probability of user u_i issuing query q_i.

Definition 6 (POIIR). The POIIR is the abbreviation of "POIs influence range." Let $P = \{p_1, \ldots, p_n\}$ indicate a set of POIs that possess identical datatype in the LBS database. Thus,

$$\mathrm{POIIR}(p_i) = \left\{ p \mid \mathrm{dist}(p, p_i) \leq \mathrm{dist}\left(p, p_j\right), i \neq j \right\}, \quad (1)$$

where p is an arbitrary point in the service range.

For ease of description, we define some terminology about location privacy. The definition of notations in our work is shown in Table 1.

4. Region Queries Framework

4.1. Region Queries. We map the experimental area onto a grid G composed of cells. Each cell corresponds to a Hilbert value, covers an $\alpha \times \alpha$ square area, where α indicates the parameter that defines the cell size of the grid G. Users regularly upload location information to a location cloaking server. The current cell of a user contains the current position of the moving object.

In our solutions, the objective for R_q is to find some Minimum Cloak Regions (MCRs). All these MCRs meet the requirement of user privacy. Similar to the Hilbert Cloak, given a query from the mobile client (MC) with anonymity requirement k, Location Cloaking Server (LCS) ranks the Hilbert Values and splits them into k-buckets. The LCS calculates the start and end positions defining the k-bucket that includes requester and constructs k-ASR using all users in the same bucket. The difference is that our solutions meet the requirements of the location m-diversity, while building k-ASR for each user.

For example, as shown in Figure 2, suppose u_2 issues the query "$R_q = \langle (0, 0), (7, 7), 4, 3 \rangle$." We can easily calculate that one of k-ASR is $\{(0, 0), (3, 3)\}$. Moreover, k-ASR offered by LCS is not unique, which may be $\{(0, 5), (5, 7)\}$. What it is designed to do is to be against inference attacks of LCS. The MC chooses a correct k-ASR that contains its real coordinates as the basis for the C_q.

4.2. System Model and Expression. In a LBS system, a large number of mobile users move within a two-dimensional square unit space. Users can issue location-dependent queries, answered by LBS providers. We adopt the three-tier centralized architecture consisting of three key parts: mobile user, location cloaking server, and LBS server.

Mobile clients (MC) are equipped with a positioning device, for example, GPS or sensor-based local positioning systems, to determine its current location information l. All of the users who held MC in our model enjoy location-dependent service by the LBS server. This device is trusted, and no malicious software component running on the mobile device has access to the location sensor. That can be assured by using a trusted computational approach.

LBS servers (LBSS) are the service providers of the LBS system. These LBSS are nontrusted since an attacker is aware of all the information that users provided to the LBS server and compromise user privacy. In addition, we assume that the attacker has statistical background information about the

Table 1: Notations.

Symbol	Description
l	User location
R	Query region presented by user
R'	A minimum cloaking region selected by MC
k	The degree of k-anonymity
m	Location m-diversity
r	Query r-diversity
C	The set of POIs category
C'	A subset of C
S_n	Candidate set
P_r	Coordinate set of POIs picked by user
M_i	Service information of POIs
K_i	Encryption key

users, although in practice, it is difficult to model the exact knowledge.

Location cloaking servers (LCS) are also the semitrusted party placed between MC and LBSS. All registered mobile users periodically update their location information to the LCS. These LCS construct MCRs, which meets users' requirements of location k-anonymity and location m-diversity.

Users establish a secure connection (e.g., an SSL) with LCS, hiding the query issuer's identity and IP address. As a hypothesis for our model, we further consider that the anonymity algorithm used by LCS is public. We support that the distribution of the population in the geographical space is uneven to conform with laws of nature.

The general procedure of continuous region query processing and specification processing is shown in Figure 3.

(1) A user sends a query R_q that contains the user's privacy requirement to LCS

(2) LCS executes MCRs Finding Algorithm to form MCRs and initiates C_q to LBSS

(3) LBSS retrieve the spatial database and interact with LCS

(4) LCS minifies the candidate set before sending the results to the user

4.3. MCRs Finding Algorithm. The LCS executes MCRs Finding Algorithm to calculate MCRs. We use the notation $G = \{g_1, \ldots g_n\}$ to denote Hilbert curve space; $g_i\,(1 \le i \le n)$ is some of the cells. P represents the criterion of judgment and $P(g_i)$ = TRUE means that g_i only consists of a cell. The MCRs Finding Algorithm consists of three phases.

Firstly, as shown in Figure 4, region segmentation starts from a set of seed points. An alternative is to start with a single region ($R_q[R]$) and subdivide the regions that do not satisfy a condition of $P(g_i)$. In other words, split into four disjoint quadrants any region $P(g_i)$ for which $P(g_i)$ = FALSE. Secondly, region merging is the opposite of region splitting. It starts with small regions and merges the regions that have similar characteristics. The aim of merging any adjacent region g_j and g_k is to find MCRs. Thirdly, we adopt an R-tree to index G. The process of constructing a G is iterative. The processing is repeated until all of MCRs satisfying privacy requirements ($R_q[k]$ and $R_q[m]$) are found. The LCS randomly selects some of MCRs and sends them to the MC.

5. Privacy-Aware Region Queries

5.1. Motivation. Our framework focuses on continuous region query that is distinct from previous studies of single-point top k-nearest-neighbor query. Consider an application scenario shown in Figure 5; the same icons represent that these POIs belong to the same classification; and A, B, and C represent different mobile users, respectively.

The common region queries are classified into three categories: (1) A uses its location as the center of region queries; (2) B uses one certain POI as the center of region queries, but B is not in the particular area; and (3) C uses one certain POI as the center of region queries, but C is in the particular area. Suppose that a user named Alice is moving in a bidimensional road network.

The above description faces two problems. Firstly, users desire to experience both high-quality service and not to expose location and identity. Therefore, users are more concerned about privacy issues. Secondly, the LBSS do not want to publish more information about POIs, which means the LBSS also express concern about the quality of service issues and business profits. From the privacy perspective, both LBSS and MC are attackers. In addition, the IP address issue is orthogonal to our problem. It can be achieved through a widely available anonymous web browsing service.

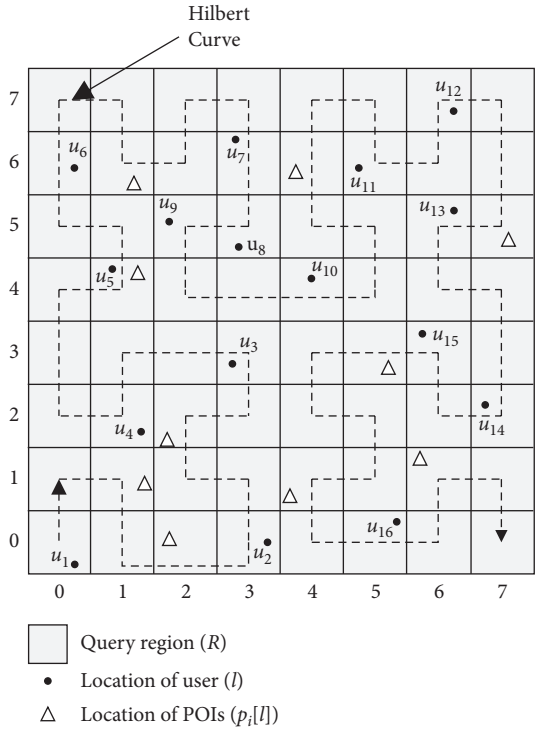
5.2. OT-Assist Privacy-Aware Protocol. Oblivious transfer protocol normally runs as a building block for more complex secure protocols or as a stand-alone protocol for privacy-preserving in LBS. Efficient 1-out-of-r oblivious transfer schemes (OT_r^1) rely on the hardness of the decisional Diffie–Hellman problem to achieve unconditional security. Assume an order-q group G_q with a short description, where q is a large prime number. Let g and h be two generators of G_q. Parameters g, h, q, G_q are publicly accessed by every entity in our protocol, where senders and receivers refer to MC and LBSS, respectively. LBSS have r keys $K_1, \ldots K_r$. The MC knows one of the key $K_a\,(1 \le a \le r)$ is his/her own choice and does not want LBSS to have that data. Meanwhile, the LBSS only provide K_a for the MC but do not want MC to get more information. The implementation process of OT-assist privacy-aware protocol is shown as follows:

(1) MC chooses $a\,(1 \le a \le r)$, generates a random number d, calculates $y = g^d h^a \mod q$, and sends y to LBSS.

(2) LBSS calculate two tuples of sequence D and send D to MC. Here, $D = \{(s_1, t_1), \ldots, (s_r, t_r)\}$, $s_i = g^{k_i} \mathrm{mod} q$, $t_i = K_i\,(y/h^i)^{k_i} \mathrm{mod} q$, $k_i \in Z_q\,(1 \le i \le r)$.

(3) LBSS send D to MC.

(4) MC calculates $K_a = (t_a/(s_a)^d) \mathrm{mod} q$.



User	Hilbert value	Grid no.
u_1	1	0,0
u_2	2	0,3
u_3	3	3,3
u_4	4	2,1
u_5	5	4,1
u_6	6	6,0
u_7	7	6,3
u_8	8	5,3
u_9	9	5,2
u_{10}	10	4,4
u_{11}	11	6,5
u_{12}	12	7,6
u_{13}	13	5,6
u_{14}	14	2,7
u_{15}	15	3,6
u_{16}	16	0,5

Figure 2: Hilbert curve map for a 2D space with 8×8 space partition.

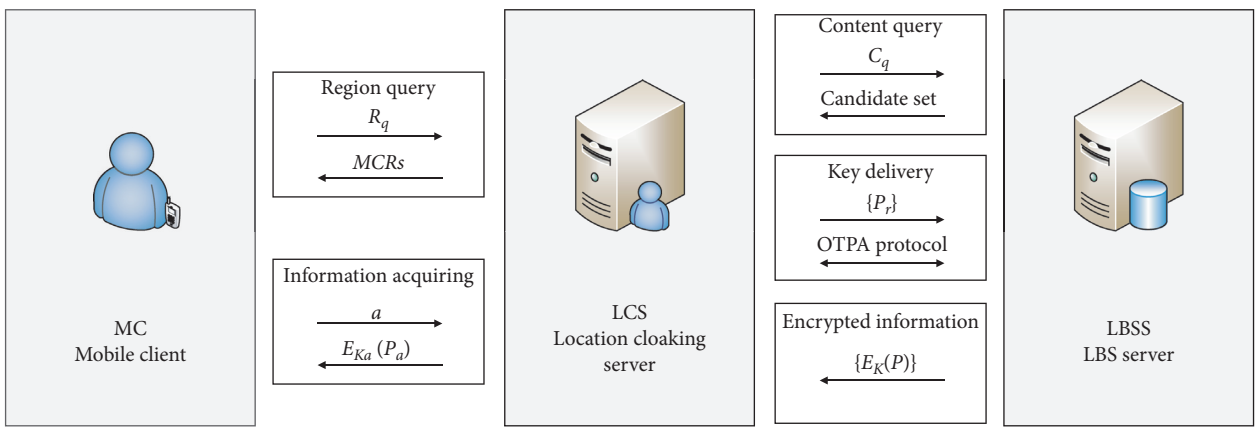


Figure 3: The system model that contains three entities: MC, LCS, and LBSS.

The purpose of the OTPA protocol is to obtain one and only one key from LBSS. This scheme meets the following privacy requirements. For any a, there is d that satisfies $y = g^d h^a \bmod q$. Therefore, LBSS cannot get any information related to a, even if it has unlimited computing power. When MC and LBSS gradually follow the protocol, although MC receives LBSS's secrets $K_1, \ldots K_r$ and cannot get two secrets, there is no way of getting information other than K_a.

5.3. Bidirectional Security Processing. Assume that the LBSS have r POIs information $P = (p_1, \ldots, p_r)$ and randomly generate the r key $K = (k_1, \ldots, k_r)$. Query senders desire p_a, but they do not wish the LBSS to know what they will get. Moreover, the LBSS also employ k_a to prevent users from accessing unauthorized content. We define this query process as bidirectional security processing. We implement our solutions with secure multiparty computation theories. It is reasonable to make an assumption about which the LCS does not collude with the LBSS since the LCS stores query examples of the MC. Otherwise, it will completely subvert any method for location privacy preserving if the LCS is allowed to collude with LBSS. We consider that all the participants in a query session are semihonest. The MC and the LCS try to obtain more data than authorized. The LBSS tries to associate a user with a location or some POIs. More details of bidirectional security processing are depicted as follows, as shown in Figure 6:

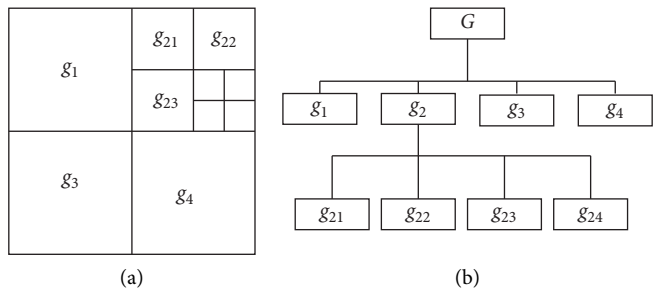


Figure 4: An example of MCRs Finding Algorithm. (a) Splitted region. (b) R-tree.

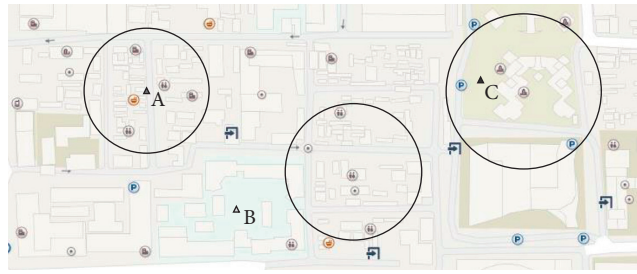


Figure 5: An example of a realistic road network environment.

(1) The MC submits a region query R_q to the LCS to find some MCRs.

(2) The LCS responds to the request of the user according to the privacy requirements of the user, executes MCRs Finding Algorithm, and sends some of MCRs to the MC.

(3) The MC randomly selects MCRs as $C_q[R']$ and submits a content query C_q to the LBSS for obtaining POIs candidate set S_n. $C_q[C']$ contains actual POIs category (c_i) about this query.

(4) The LBSS calculate all candidate POIs of $C_q[R']$' and send candidate set S_n to MC. S_n is formulated as the following form:

$$S_n = \left\{p_i[l] | p_i[l] \subset C_q[R'] \text{ and } p_i[c] \in C_q[C']\right\}. \quad (2)$$

Further, we can also express S_n as $S_n = \{S^{(1)}, \ldots, S^{(r)}\}$, where $S^{(j)}\,(1 \le j \le r)$ is a location set retrieved by c_j.

(5) MC calculates the obstacle distance between its current coordinate and each element of $S^{(i)}$ and adds the nearest point to the set P_r. MC randomly also extracts an element from $S^{(j)}\,(1 \le j \le r, i \ne j)$ and adds it to the set P_r. The MC disrupts the order of P_r and sends it to the LBSS.

(6) The LBSS retrieve the spatial database and find all of POIs information in terms of P_r. It is referred to as M.

(7) The MC and the LBSS perform OTPA protocol.

(8) The LBSS can encrypt $M = \left\{E_{K_1}(p_1), \ldots, E_{K_r}(p_r)\right\}$ to prevent LCS from reading it and send it back to the LCS.

(9) The MC retrieves a particular record for $E_{K_a}(p_a)$, which is precisely what the user needs.

6. Security Analysis

Data security and user's privacy have the absolute critical priority for a LBS system. There is much more risk of sensitive data being stolen or leaked because LBSS gather mass data from social media users. In this section, firstly, we explain the privacy threats caused by location and measurement of the privacy leakage. Moreover, we compare the proposed solution with existing works in terms of location k-anonymity, location m-diversity, and query r-diversity.

6.1. Attack Expression and Privacy Metric. Location privacy is the nature of an individual to control access to their current and past location information. Figure 7 shows the importance of location. There are four key factors affecting personal privacy in LBS system: identity, location, time stamp, and candidate POIs. As long as it is not associated with the particular user's identity, query context does not lead to privacy disclosure. However, the user's trajectory is the key link in query context and user's identity. For example, continuous location samples have been tracked by attacks and then used to infer a user's identity. The relationship feature between trajectory and POIs can also be used to define a user's behavior. The combination of identity and behavior exposed the sensitive data of the user.

All research related to location privacy stems from the assumption that untrusted LBS providers are the most critical threat to privacy. The LBS attacks involve two aspects: location tracing and user identification. Meanwhile, the prior knowledge of the attacker is unable to measure, and

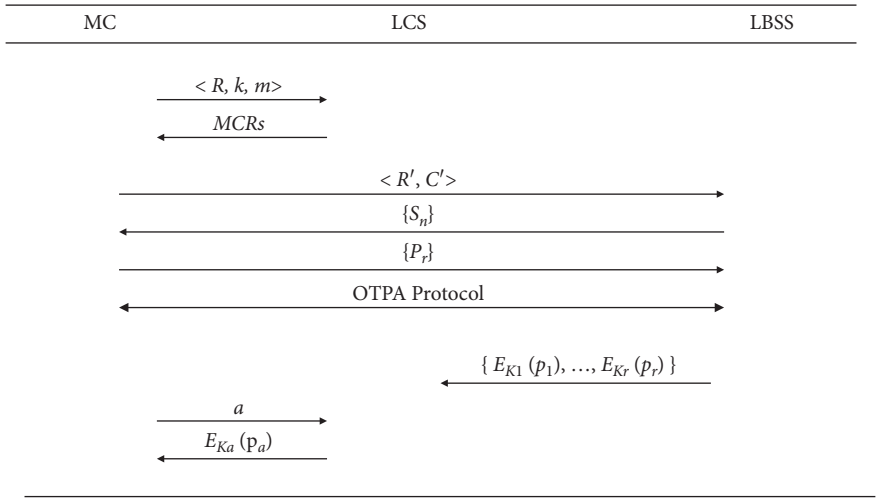


Figure 6: The interactions between MC, LCS, and LBSS.

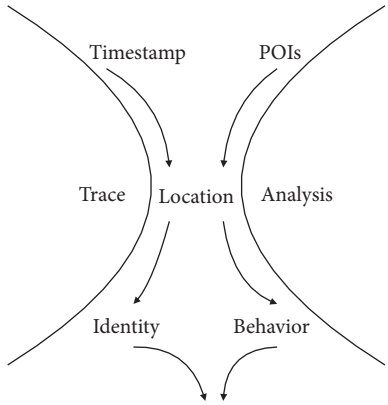


Figure 7: Location is a key factor in the LBS system.

the invade mode taken by the attacker is unpredictable. As the diversity of profiles, such as user profile or user velocity, are not the same, the spatial cloaking faces continuous multiquery attacks, inference attacks, and correlation attacks.

Theorem 1. *The combination of identity, location, timestamp, and candidate POIs poses a serious threat to user privacy, and location plays a significant role in the LBS system.*

The concept of entropy was rooted in Shannon entropy. It gives an accurate metric of the uncertainty that an attacker infers for the user's information. Shannon entropy also can be used to evaluate location privacy or query privacy. Before a user submits a query, the uncertainty over location obtained by LBSS has been called Priori Location Entropy. However, we can improve the degree of privacy using some techniques such as anonymity, fuzzy, and obfuscation. The uncertainty over location obtained by LBSS has been called Posterior Location Entropy after applying these techniques. The inherent feature of Location Entropy is mainly embodied in the following aspects. Firstly, when the LBSS have real-time location information of users, the Priori Location Entropy $H(l_i) = 0$. *Secondly, when the LBSS do not have any background knowledge, the maximum Priori Location Entropy* $H(l_i) = \log_2^m$. *Thirdly, when the LBSS have some background knowledge which is achieved through statistical analysis, the Priori Location Entropy* $0 < H(l_i) < \log_2^m$.

We can easily recognize that higher entropy is associated with three things: location k-anonymity degree, location m-diversity, and query r-diversity. In our solutions, the probabilities of location anonymity, location diversity, and query diversity are $1/k$, $1/m$, *and* $1/r$, *respectively. Users can freely control their privacy requirements because all of these parameters are determined by themselves. Therefore, our approaches achieved the purpose of hiding user privacy. Obviously, it is inevitable that each query provides some new knowledge for LBSS, which is more conducive to inferring the user's sensitive data. However, our solutions improved the complexity of the invasion of privacy, although they do not overcome the inherent limitations of spatial and temporal cloaking methods. We will be establishing a privacy measure model in subsequent studies.*

6.2. Comparison of OTPA, Spatial Cloaking, and PIR. Because it is required to submit a k-ASR to LBSS in spatial cloaking methods, the user issuing queries must be appearing in the area. The anonymous area has been gradually diminished by attackers according to user profiles, road network restrictions, and moving speed. Thus, the user's trajectory is traceable. The quality of trajectory details relies heavily on the power of an attacker. At the same time, the candidate result set is a vital component for LBS providers to infer the user's sensitive data. There is a direct correlation between queries content and user identity. An attacker can deduce who is most likely to issue the query. In our solutions, the region submitted to LBSS satisfies four properties: location k-anonymity, location m-diversity, query r-diversity, and reciprocal relationship of ASR. Therefore, our solutions can resist the inference attack for spatial cloaking. Firstly, the user submitting queries does not

reveal the accurate location to LCS and LBSS since the calculation program of the nearest neighbor runs on the client device. However, LBSS can calculate the minimum inference region, which is the intersection of the R' and all of disclosed POIs influence regions (POIIR). Consequently, larger value of R' means higher location privacy for the user. The POIIR of disclosed POIs is discrete and random, which makes it difficult to trace the sequence of trajectories. Moreover, the user submitting queries confuses the query content with a plurality of POIs that are selected by themselves. Therefore, the probability of LBSS inference user query content is $1/r$. Consequently, LBSS cannot associate the user with the identity by specific POIs.

Theorem 2. *Assume that all of these attributes of location k-anonymity, location l-diversity, query r-diversity, and reciprocal relationship of ASR can guarantee privacy, which makes our solution have the untraceability and the unlinkability.*

OTPA is parallel to PIR. Both of them are based on encryption techniques to protect user privacy. Computational PIR relies on the quadratic residuosity problem. However, it cannot avoid a linear scan of the entire database for processing each query. The communication complexity of each query is roughly $\sqrt{n}$. The symbol n represents the size of the database. Therefore, the PIR techniques require extreme computational efficiency, where the usage of resources, such as run-time, storage, or data samples, is sublinear in the size of the candidate module. In contrast, OTPA does not have such requirements. Our solutions are superior to PIR techniques because the typical PIR framework does not limit the number of POIs obtained by the user. Thus, it does not provide an effective way to protect the valuable resources of LBS server.

Theorem 3. *Assuming that the OTPA scheme is unconditionally secure, our solution achieves server-oriented security. It can be hard to maliciously get precious data of the LBS server.*

7. Experiment Results and Discussion

We implement a prototype system by extending an existing work of C# program that supports OT protocol. The database is one of the widest and most interesting public data sets to analyze user trajectory which is generated by Brinkhoff's network-based generator of moving objects. We conduct the experiments on a machine with Intel(R) Core(TM) i7-10510U CPU and 40 GB memory and some smartphones with Android 10 OS as the client. Our experimental default parameters are summarized in Table 2. We simulate 1000 users sending queries randomly to the LBS provider through a wireless network. Default values for these parameters constrain the scope of the following experiments; see Table 1 for specific meanings.

In the following experiments, we mainly focus on the communication cost and the computational cost, which is the dominating factor for the proposed solutions. In OT protocol, the cost of computation is often criticized with the comparison of communication cost. OT protocol is characteristically implemented using modular exponentiations, which are involved in the intensive computing. Therefore, researchers are more concerned about the effectiveness and availability of these algorithms in cryptographic applications.

Table 2: Parameters and default values.

Parameters	Default values
R	1 km^2
k	50
m	20
r	10
C	50
M_i	10 kbit
K_i	64 bit

The first experiment aims at studying the time consumption with different numbers of candidate POIs. The efficiency of our approaches depends on parameters R and R'. Without loss of generality, we assume that the number of candidate POIs is directly proportional to the size of R'. The time consumption in two query phases is shown in Figure 8. The result shows that the CPU time of content query is large since the number of modular exponentiation is proportional to the number of candidate POIs.

As shown in Figures 9–11, the CPU time is influenced by these parameters (R, R', k, m, and r) in the region query and content query. We can find that more stringent privacy requirements take longer time.

Figure 12 shows the result of the comparison with the typical method Casper and PIR. Experimental results indicate that the average processing time of the above three methods is linear to the number of candidate POIs. From computation efficiency, modular exponentiation is the most expensive. Therefore, Casper performs better than the other two methods in the average computation time.

The second experiment focuses on studying the communication cost in the two-query phase. Figure 13 shows that the communication cost in the region query is lower since the main communications are composed of some coordinates of POIs transferred from server to clients. The communication cost of the content query will just keep growing. However, its upper limit is around 550 kb since the category of POIs is no more than 50. The R and α affected communication cost in region query stage, and R and α are larger, which makes the traffic greater. R' and C' affected the communication cost in the content query stage. The larger the R' and C', the greater the traffic loads. At the same time, k and m have decided the area of R', and r have limited the dimensions of C'. Therefore, the higher the user's privacy requirement, the greater the traffic loads.

Finally, we observe the number of POIs that users obtain from each query since users are often charged by the LBS provider according to the number of retrieved POIs. We conduct experiments to compare with other techniques. Figure 14 shows that the number of candidate POIs is linear to the number of users. The difference is due to the diversity of the querying methods. These results indicate that, in order to maintain an appropriate number of disclosed POIs,

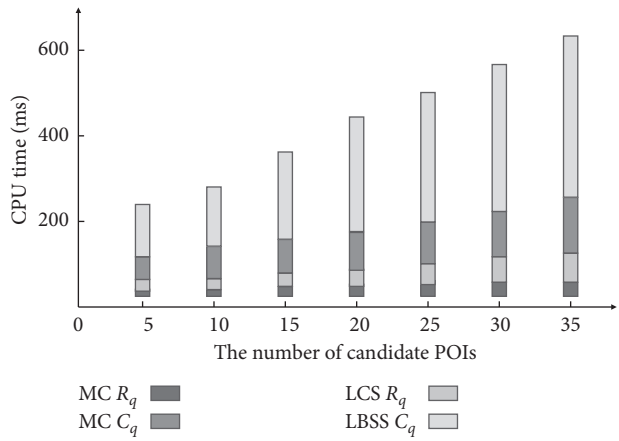


Figure 8: The time consumption in two query phases.

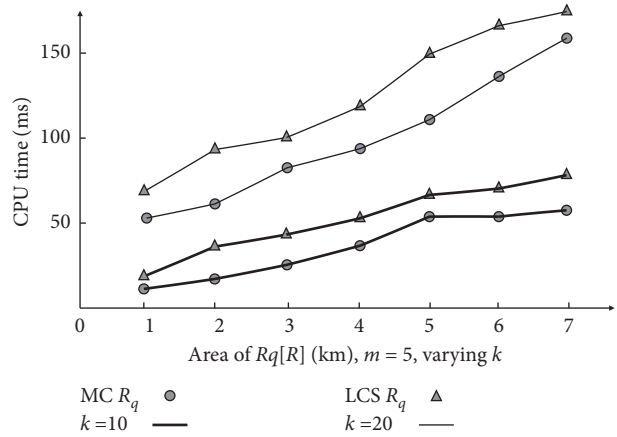


Figure 9: The R and k influence on CPU time.

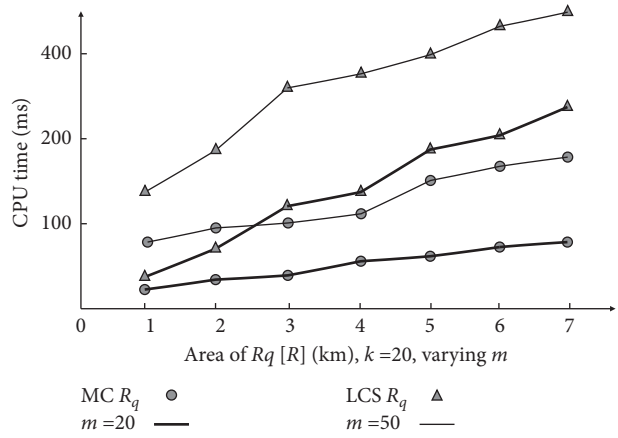


Figure 10: The R and m influence on CPU time.

cloaking-based methods have to collect a large number of users. These result in a high cost of location updates and pose privacy concerns since all users must be trustworthy. The number of disclosed POIs is constant for PIR methods because no other users are required to construct a cloaking set. The number of candidate POIs gradually decreases from 50 to 1 as the user number increases in our solutions. However, only one candidate POI is exposed to the user submitting query. Therefore, we provide security guarantees for the resources of the LBS server.

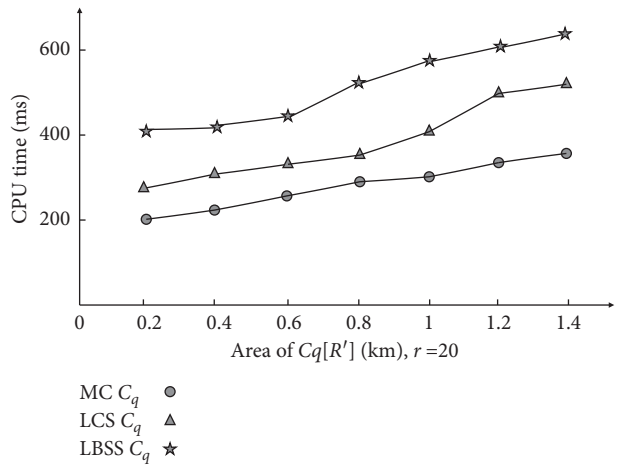


Figure 11: The R' and r influence on CPU time.

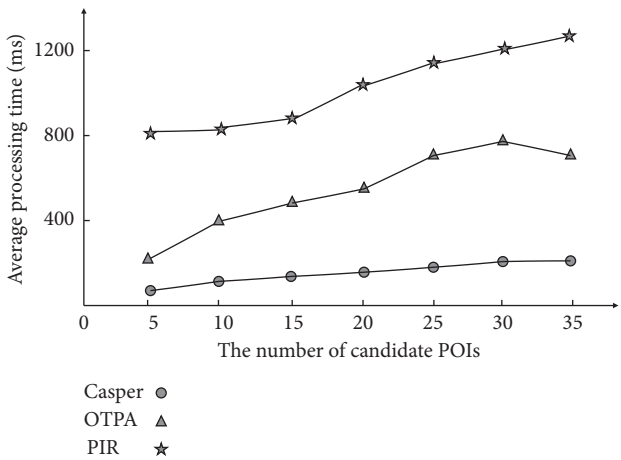


Figure 12: The comparison with the Casper and PIR.

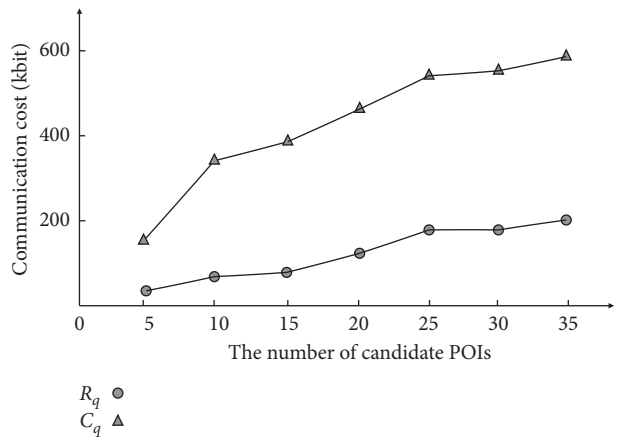


Figure 13: The communication cost in the region query and content query.

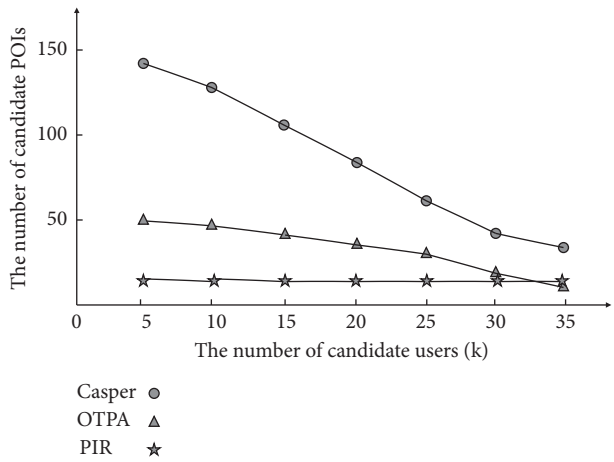


Figure 14: The relationship between candidate POIs and the number of users.

8. Conclusion

Our awareness of privacy has been heightened lately because some platforms abuse our personal data gathered by LBSS or LCS. Two prominent issues need to be further explored in the field of LBS privacy. Many studies assumed that the parties involved in anonymity are entirely trustworthy. In reality, participants could reveal the other location information because of the inconsistency of privacy degree of anonymous. In addition, the strategy that the LBSS confuse attackers with a plurality of redundant POIs information is not conducive to the operation of the LBS market and hinders the development of LBS. We developed a region queries framework and designed a privacy-aware query protocol-based oblivious transfer protocol, mainly to solve the aforementioned problems. Our solution has met the requirement of untraceability and unlinkability under the premise of preserving personal privacy. Therefore, it is certified that authenticated users can only obtain service information what they need, but malicious users cannot steal LBS server resources. Simulation results show a mutual influence and interactive relationship between the query processing time, the communication cost, the privacy degree, and the candidate POIs. Although it is inevitable that strict privacy requirements must confront a sacrifice of service quality, we will enhance our understanding of LBS to strengthen future work from reducing operating costs to improving efficiency and reinforcing privacy.

Data Availability

The location data used to support the findings of this study may be released upon application to the Microsoft GeoLife GPS Trajectories, who can be contacted at http://research.microsoft.com/en-us/downloads/b16d359d-d164-469e-9fd4-daa38f2b2e13/default.aspx.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was partially supported by grants from the Scientific Research Fund for Advanced Talents of Bengbu University (no. BBXY2021KYQD01); the Scientific Research Projects of Universities in Anhui Province of China; the Humanity and Social Science Youth Foundation of Ministry of Education of China (no. 18YJCZH068); and the Natural Science Foundation of the Jiangsu Higher Education Institutions of China (no. 18KJB520002).

References

[1] A. Das, M. Degeling, D. Smullen, and N. Sadeh, "Personalized privacy assistants for the Internet of Things: providing users with notice and choice," *IEEE Pervasive Computing*, vol. 17, no. 3, pp. 35–46, 2018.

[2] K. H. Mohammadani, K. A. Memon, I. Memon, N. N. Hussaini, and H. Fazal, "Preamble time-division multiple access fixed slot assignment protocol for secure mobile ad hoc networks," *International Journal of Distributed Sensor Networks*, vol. 16, no. 5, Article ID 155014772092162, 2020.

[3] H. Li, H. Zhu, S. Du, X. Liang, and X. Shen, "Privacy leakage of location sharing in mobile social networks: attacks and defense," *IEEE Transactions on Dependable and Secure Computing*, vol. 15, no. 4, pp. 646–660, 2018.

[4] R.-H. Hwang, Y.-L. Hsueh, and H.-W. Chung, "A novel time-obfuscated algorithm for trajectory privacy protection," *IEEE Transactions on Services Computing*, vol. 7, no. 2, pp. 126–139, 2014.

[5] J. Wang, Z. Cai, Y. Li, D. Yang, J. Li, and H. Gao, "Protecting query privacy with differentially private k-anonymity in location-based services," *Personal and Ubiquitous Computing*, vol. 22, no. 3, pp. 453–469, 2018.

[6] S. Zhang, X. Li, Z. Tan, T. Peng, and G. Wang, "A caching and spatialK-anonymity driven privacy enhancement scheme in continuous location-based services," *Future Generation Computer Systems*, vol. 94, pp. 40–50, 2019.

[7] R. Shokri, C. Troncoso, C. Diaz, J. Freudiger, and J. P. Hubaux, "Unraveling an old cloak: K-anonymity for location privacy,"

in *Proceedings of the 9th Annual ACM Workshop on Privacy in the Electronic Society-WPES'10*, 2010.

[8] X. He, R. Jin, and H. Dai, "Leveraging spatial diversity for privacy-aware location-based services in mobile networks," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 6, pp. 1524–1534, 2018.

[9] R. Schlegel, C.-Y. Chow, Q. Huang, and D. S. Wong, "User-defined privacy grid system for continuous location-based services," *IEEE Transactions on Mobile Computing*, vol. 14, no. 10, pp. 2158–2172, 2015.

[10] P. Perazzo and G. Dini, "A uniformity-based approach to location privacy," *Computer Communications*, vol. 64, no. 64, pp. 21–32, 2015.

[11] A. S. Saxena, D. Bera, and V. Goyal, "Modeling location obfuscation for continuous query," *Journal of Information Security and Applications*, vol. 44, no. 44, pp. 130–143, 2019.

[12] A.-M. Olteanu, K. Huguenin, R. Shokri, M. Humbert, and J.-P. Hubaux, "Quantifying interdependent privacy risks with location data," *IEEE Transactions on Mobile Computing*, vol. 16, no. 3, pp. 829–842, 2017.

[13] R. Paulet, M. G. Kaosar, Y. Xun, and E. Bertino, "Privacy-preserving and content-protecting location based queries," *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 5, pp. 1200–1210, 2014.

[14] H. Jannati and B. Bahrak, "An oblivious transfer protocol based on elgamal encryption for preserving location privacy," *Wireless Personal Communications*, vol. 97, no. 2, pp. 3113–3123, 2017.

[15] N. Chen, J. Li, Y. Zhang, and Y. Guo, "Efficient CP-ABE scheme with shared decryption in cloud storage," *IEEE Transactions on Computers*, p. 1, 2020.

[16] J. Li, Y. Wang, Y. Zhang, and J. Han, "Full verifiability for outsourced decryption in attribute based encryption," *IEEE Transactions on Services Computing*, vol. 13, no. 3, pp. 478–487, 2020.

[17] J. Li, Y. Zhang, J. Ning, X. Huang, G. S. Poh, and D. Wang, "Attribute based encryption with privacy protection and accountability for CloudIoT," *IEEE Transactions on Cloud Computing*, p. 1, 2020.

[18] J. Li, H. Yan, and Y. Zhang, "Identity-based privacy preserving remote data integrity checking for cloud storage," *IEEE Systems Journal*, vol. 15, no. 1, pp. 577–585, 2021.

[19] J. Li, H. Yan, and Y. Zhang, "Certificateless public integrity checking of group shared data on cloud storage," *IEEE Transactions on Services Computing*, vol. 14, no. 1, pp. 71–81, 2021.

[20] J. Li, H. Yan, and Y. Zhang, "Efficient identity-based provable multi-copy data possession in multi-cloud storage," *IEEE Transactions on Cloud Computing*, p. 1, 2019.

[21] L. Zhang, H. Xiong, Q. Huang, J. Li, K.-K. R. Choo, and J. Li, "Cryptographic solutions for cloud storage: challenges and research opportunities," *IEEE Transactions on Services Computing*, p. 1, 2020.