**Geometric Range Search on Encrypted Spatial Data**

**ABSTRACT:**

Geometric range search is a fundamental primitive for spatial data analysis in SQL and NoSQL databases. It has extensive applications in location-based services, computer aided design, and computational geometry. Due to the dramatic increase in data size, it is necessary for companies and organizations to outsource their spatial data sets to third-party cloud services (e.g., Amazon) in order to reduce storage and query processing costs, but, meanwhile, with the promise of no privacy leakage to the third party. Searchable encryption is a technique to perform meaningful queries on encrypted data without revealing privacy. However, geometric range search on spatial data has not been fully investigated nor supported by existing searchable encryption schemes. In this paper, we design a symmetric-key searchable encryption scheme that can support geometric range queries on encrypted spatial data. One of our major contributions is that our design is a general approach, which can support different types of geometric range queries. In other words, our design on encrypted data is independent from the shapes of geometric range queries. Moreover, we further extend our scheme with the additional use of tree structures to achieve search complexity that is faster than linear. We formally define and prove the security of our scheme with indistinguish ability under selective chosen-plaintext attacks, and demonstrate the performance of our scheme with experiments in a real cloud platform (Amazon EC2).

**EXISTING SYSTEM:**

* While most of the searchable encryption schemes focus on common SQL queries, such as keyword queries and Boolean queries, few studies have specifically investigated geometric range search over encrypted spatial data.
* Wang et al. proposed a novel scheme to specifically perform *circular range* queries on encrypted data by leveraging a set of *concentric circles*.
* Some previous searchable encryptions handling order comparisons can essentially manage axis parallel rectangular range search on encrypted spatial data.
* Similarly, Order-Preserving Encryption, which has weaker privacy guarantee than searchable encryption, is also able to perform axis-parallel rectangular range search with trivial extensions.
* Ghinita and Rughinis particularly leveraged certain Functional Encryption with hierarchical encoding to efficiently operate axis-parallel rectangular range search on encrypted spatial data in the application of mobile users monitoring.

**DISADVANTAGES OF EXISTING SYSTEM:**

* Most of the searchable encryption schemes focus on common SQL queries, such as keyword queries and Boolean queries, few studies have specifically investigated geometric range search over encrypted spatial data.
* Inevitably introduces obstacles in terms of search functionalities over encrypted data.
* None of these previous works have particularly studied geometric range queries which are expressed as *non-axis-parallel rectangles* or *triangles*.
* More importantly, there still lacks a *general* approach, which can flexibly and securely support different types of geometric range queries over encrypted spatial data regardless of their specific geometric shapes.

**PROPOSED SYSTEM:**

* In this paper, we propose a *symmetric-key probabilistic* Geometric Range Searchable Encryption. With our scheme, a *semi-honest* (i.e., *honest-but-curious*) cloud server can verify whether a point is inside a geometric range over encrypted spatial datasets. Informally, except learning the necessary Boolean search result (i.e., *inside or outside*) of a geometric range search, the semi-honest cloud server is not able to reveal any private information about data or queries.
* Our main contributions are summarized as follows:
* We present a symmetric-key probabilistic Geometric Range Searchable Encryption, and formally define and prove its security with indistinguishability under Selective Chosen-Plaintext Attacks (IND-SCPA).
* In addition, our search process is *non-interactive* on encrypted data. In terms of search complexity, our baseline scheme incurs linear complexity (with regard to the number of data records), and its advanced version realizes faster than- linear search by integrating with tree structures.
* Our design is a general approach, which can securely support different types of geometric range queries on encrypted spatial data regardless of their geometric shapes. Furthermore, our design is not only suitable for geometric range queries, but also compatible with other regular types of geometric queries, such as *intersection queries* and *point enclosure queries*, over encrypted spatial data.

**ADVANTAGES OF PROPOSED SYSTEM:**

* The security of our scheme is formally defined and analyzed with indistinguishability under Selective Chosen-Plaintext Attacks.
* Our design has great potential to be used and implemented in wide applications, such as Location-Based Services and spatial databases, where the use of sensitive spatial data with a requirement of strong privacy guarantee is needed.

**SYSTEM ARCHITECTURE:**



**MODULES:**

* System Construction Model
* Design Methodology
* Deterministic Scheme
* Probabilistic Scheme

**MODULES DESCSRIPTION:**

**System Construction Model**

The system model of our scheme is developed in this module, which includes a data owner, a data user and the cloud server. A data owner (e.g., a company or an organization) stores its dataset on the cloud server to reduce local cost on data storage and query processing. A data user (e.g., a user of the company or a user of the organization) would like to search over the outsourced spatial dataset in the cloud. The cloud server provides data storage and search services. Note that the data owner itself always has the capability to search over outsourced spatial data.

The cloud server is a *semi-honest* (i.e., *honest-but-curious*) entity, which indicates it provides reliable services, but it will try to learn the private information about data records and geometric range queries. In order to preserve private information leakage, the data owner only stores the encrypted form of its spatial dataset on the cloud server, and a client (a data user or the data owner) only submits the encrypted version (i.e., a search token) of its geometric range query to the cloud server.

**Design Methodology**

Performing *different and continuous* operations over encrypted data makes it challenging to design a general geometric range searchable encryption scheme. In order to flexibly manage different geometric range queries, our main design methodology in this paper is to *preprocess* each type of geometric range queries to a *same form* in the plaintext domain, so that we only need to handle a *single* type of operations in the ciphertext domain. As a consequence, multiple rounds of client-to-server interactions or the impractical assumption of multiple non-colluding servers can be avoided.

**Deterministic Scheme**

With our design methodology in mind, we can first build a Basic scheme with Deterministic Encryption (e.g., a Pseudo Random Function) and Bloom filters. The inherent property of Deterministic Encryption (*i.e., producing a same ciphertext for a same message*) can help us perform equality testing consistently over encrypted data, and the use of Bloom filters can improve the efficiency of set membership testing.

Specifically, a data owner can encrypt each data record with Deterministic Encryption. While for each geometric range query, after a data owner enumerates all the possible points from the data space that are inside the geometric range in the plaintext domain, it encrypts all those possible points separately with Deterministic Encryption, and adds those corresponding ciphertexts to a Bloom filter one after another. The Bloom filter containing the ciphertexts of all the possible points inside the geometric range query will be used as a search token. Finally, the cloud server is able to test whether a point is inside a geometric range query by checking whether an encrypted data record is an element contained in the Bloom filter.

**Probabilistic Scheme**

To overcome limitations in the preceding deterministic scheme, we now build a probabilistic GRSE scheme with the same design. Compared to the deterministic one, this probabilistic scheme can provide both data privacy and query privacy under IND-SCPA. Besides, it is able to preserve query range pattern, which is an inevitable leakage in the preceding deterministic scheme. To achieve these security objectives, the main difference of this probabilistic GRSE scheme (from the high level) is to first add points into a Bloom filter, and then leverage probabilistic encryption to encrypt all the bits in the Bloom filter. However, using probabilistic encryption to protect every bit in a Bloom filter introduces additional challenges of verifying set memberships. Specifically, since probabilistic encryption naturally produces different ciphertexts for a same message, it is computationally impossible for the server to distinguish which bit positions are 1s and which bit positions are 0s in a Bloom filter.

**SYSTEM REQUIREMENTS:**

**HARDWARE REQUIREMENTS:**

* System : Pentium Dual Core.
* Hard Disk : 120 GB.
* Monitor : 15’’ LED
* Input Devices : Keyboard, Mouse
* Ram : 1GB.

**SOFTWARE REQUIREMENTS:**

* Operating system : Windows 7.
* Coding Language : JAVA/J2EE
* Tool : Netbeans 7.2.1

Database : MYSQL

**REFERENCE:**

Boyang Wang, *Student Member, IEEE*, Ming Li, *Member, IEEE*, and Haitao Wang, “Geometric Range Search on Encrypted Spatial Data”, **IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, VOL. 11, NO. 4, APRIL 2016.**