**Dynamic Keyword Search Scheme over Encrypted Cloud Environment**

**ABSTRACT**:

With the ever-increasing amount of data resided in a cloud, how to provide users with secure and practical query services has become the key to improve the quality of cloud services. Fuzzy searchable encryption (FSE) is identified as one of the most promising approaches for enabling secure query services, since it allows searching encrypted data by using keywords with spelling errors. However, existing FSE schemes are far from the practical use for the following reasons: (1) Inflexibility. It is hard for them to simultaneously support and OR semantics in a multi-keyword query. (2) Inefficiency. They require sequentially scanning a whole dataset to find matched files, and thus are difficult to apply to a large-scale dataset. (3) Limited robustness. It is difficult for them to resist the linear analysis attack in the known-background model. To fix the above problems, this article proposes matrix-based multi-keyword fuzzy search (M2FS) schemes, which support approximate keyword matching by exploiting the indecomposable property of primes. Specifically, we first present a basic scheme, called M2FS-B, where multiple keywords in a query or a file are constructed as prime-related matrices such that the result of matrix multiplication can be employed to determine the level of matching for different query semantics. Then, we construct an advanced scheme, named M2FS-E, which builds a searchable index as a keyword balanced binary (KBB) tree for dynamic and parallel searches, while adding random noises into a query matrix for enhanced robustness. Extensive analyses and experiments demonstrate the validity of our M2FS schemes.

|  |  |
| --- | --- |
| **EXSISTING SYSTEM** | **PROPOSED SYSTEM** |
| * Providing a wide variety of services in a pay-as-you-go fashion, is an extremely successful paradigm of service-oriented computing. With the increasing popularity of cloud-based services, consumers are highly motivated to outsource their data and computing services to cloud platforms for lower costs, higher reliability, and better performance. * However, studies and past experience show that cloud platforms may be unreliable, and vulnerable to various threats that cause data leakage intentionally or unwittingly. * The new challenge of service quality is emerging since it is hard for traditional encryption methods to support common cloud services like keyword-based searches. | * In this article, we propose matrix-based multi-keyword fuzzy search (M2FS) schemes, which exploit the indecomposable property of prime numbers to provide enhanced service quality in cloud computing. Like the work in our M2FS schemes also apply the wildcard technique and edit distance to quantify keywords’ similarity. However, our schemes do not require the construction of a predefined fuzzy set and thus are more scalable and practical. * Our main idea is to encode a file keyword into an index vector filled with primes, such that the result of vectors’ inner product is an integer only when two keywords are similar. * For flexibility, index vectors and query vectors are organized into primerelated matrices so that the result of matrix multiplication can be used to determine whether a file matches a multi-semantic query or not. |
| **EXISTING ALGORITHM**  Multi-Keyword FSE | **PROPOSED ALGORITHM:-**  Single-Keyword FSE |
| **ALGORITHM DEFINITION:-**  Transformed multi-keyword into a single keyword through pre-defined phrases, but rendered the index size to increase with the edit distance. Encoded keywords as bi-grams and quantified keywords similarity based on Euclidean distance. The multi-keyword fuzzy search function was achieved by forward indexes built based on a collection of LSH-based Bloom filters. To improve search accuracy, a keyword transformation method so that keywords with the same root could be queried using a stemming algorithm. Proposed a GPSE scheme that allowed users to query by using generalized wildcard-based string patterns. The construction of is based on homomorphism encryption rather than Bloom filter and hence eliminates the false probability caused by Bloom filter. | **ALGORITHM DEFINITION:-**  The first single-keyword FSE scheme, which quantified the similarity of keywords with edit distance. The shortcoming of their scheme was the large index size which increased exponentially with the edit distance difference. To reduce the proposed a dictionary-based fuzzy set construction that limited the scope of fuzzy keywords. Instead of neighbor sets. However, all these schemes required a predefined dictionary, and thus making it difficult to perform updates on a file collection. The LSH which returns records within a distance of a given query with a high probability is a useful tool for fast similarity search. |
| **DRAWBACKS:-**   * It is hard for them to simultaneously support and OR semantics in a multi-keyword query. * Inefficiency. They require sequentially scanning a whole dataset to find matched files, and thus are difficult to apply to a large-scale dataset. * Limited robustness. It is difficult for them to resist the linear analysis attack in the known-background model. | **ADVANTAGES:-**   * simultaneously support * require sequentially scanning * Limited robustness |

**MINIMUMSYSTEM REQUIREMENTS**

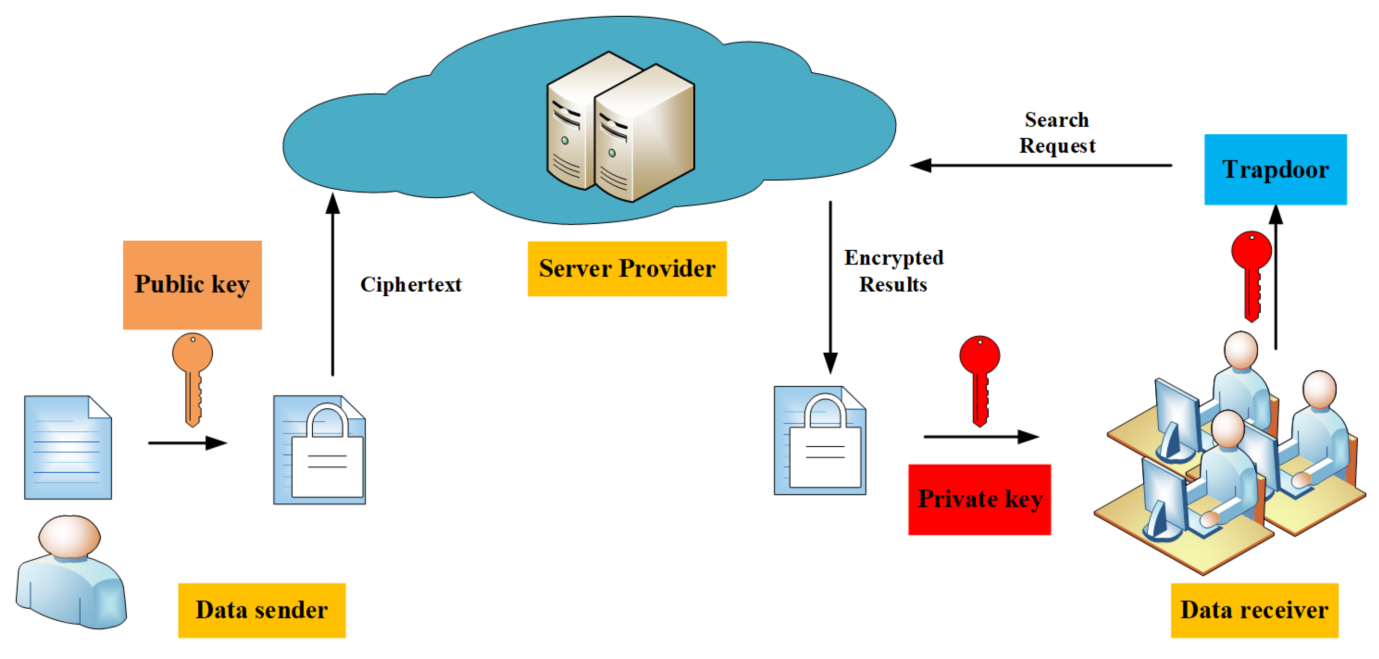
**HARDWARE REQUIREMENTS**

* PROCESSOR : DUAL CORE 2 DUO.
* RAM : 2GB DD RAM
* HARD DISK : 250 GB

**SOFTWARE REQUIREMENTS**

* FRONT END : J2EE (JSP, SERVLET)
* BACK END : MY SQL 5.5
* OPERATING SYSTEM : WINDOWS 7
* IDE : ECLIPSE

**System Architecture:**

****