**Achieving Secure, Universal, and Fine-Grained Query Results Verification for Secure Search Scheme over Encrypted Cloud Data**

**LITERATURE SURVEY**

**Security Challenges for the Public Cloud**

In this talk, we will first discuss a number of pressing security challenges in Cloud Computing, including data service outsourcing security and secure computation outsourcing. Then, we will focus on data storage security in Cloud Computing. As one of the primitive services, cloud storage allows data owners to outsource their data to cloud for its appealing benefits. However, the fact that owners no longer have physical possession of the outsourced data raises big security concerns on the storage correctness.

Hence, enabling secure storage auditing in the cloud environment with new approaches becomes imperative and challenging. In this talk, we will present our recent research efforts towards storage outsourcing security in cloud computing and describe both our technical approaches and security & performance evaluations.

**Practical Techniques for Searches on Encrypted Data**

It is desirable to store data on data storage servers such as mail servers and file servers in encrypted form to reduce security and privacy risks. But this usually implies that one has to sacrifice functionality for security. For example, if a client wishes to retrieve only documents containing certain words, it was not previously known how to let the data storage server perform the search and answer the query without loss of data confidentiality. In this paper, we describe our cryptographic schemes for the problem of searching on encrypted data and provide proofs of security for the resulting crypto systems. Our techniques have a number of crucial advantages. They are provably secure: they provide provable secrecy for encryption, in the sense that the untrusted server cannot learn anything about the plaintext when only given the ciphertext; they provide query isolation for searches, meaning that the untrusted server cannot learn anything more about the plaintext than the search result; they provide controlled searching, so that the untrusted server cannot search for an arbitrary word without the user’s authorization; they also support hidden queries, so that the user may ask the untrusted server to search for a secret word without revealing the word to the server.

We have described new techniques for remote searching on encrypted data using an untrusted server and provided proofs of security for the resulting crypto systems. Our techniques have a number of crucial advantages: they are provably secure; they support controlled and hidden search and query isolation; they are simple and fast (More specifically, for a document of length , the encryption and search algorithms only need O(*n*) stream cipher and block cipher operations); and they introduce almost no space and communication overhead. Our scheme is also very flexible, and it can easily be extended to support more advanced search queries. We conclude that this provides a powerful new building block for the construction of secure services in the untrusted infrastructure.

**Public Key Encryption with keyword Search**

We study the problem of searching on data that is encrypted using a public key system. Consider user Bob who sends email to user Alice encrypted under Alice’s public key. An email gateway wants to test whether the email contains the keyword “urgent” so that it could route the email accordingly. Alice, on the other hand does not wish to give the gateway the ability to decrypt all her messages. We define and construct a mechanism that enables Alice to provide a key to the gateway that enables the gateway to test whether the word “urgent” is a keyword in the email without learning anything else about the email. We refer to this mechanism as Public Key Encryption with keyword Search. As another example, consider a mail server that stores various messages publicly encrypted for Alice by others. Using our mechanism Alice can send the mail server a key that will enable the server to identify all messages containing some specific keyword, but learn nothing else. We define the concept of public key encryption with keyword search and give several constructions.

We defined the concept of a public key encryption with keyword search (PEKS) and gave two constructions. Constructing a PEKS is related to Identity Based Encryption (IBE), though PEKS seems to be harder to construct. We showed that PEKS implies Identity Based Encryption, but the converse is currently an open problem. Our constructions for PEKS are based on recent IBE constructions. We are able to prove security by exploiting extra properties of these schemes.

**Searchable Symmetric Encryption: Improved Definitions and Efficient Constructions**

Searchable symmetric encryption (SSE) allows a party to outsource the storage of his data to another party in a private manner, while maintaining the ability to selectively search over it. This problem has been the focus of active research and several security definitions and constructions have been proposed. In this paper we begin by reviewing existing notions of security and propose new and stronger security definitions. We then present two constructions that we show secure under our new definitions. Interestingly, in addition to satisfying stronger security guarantees, our constructions are more efficient than all previous constructions. Further, prior work on SSE only considered the setting where only the owner of the data is capable of submitting search queries. We consider the natural extension where an arbitrary group of parties other than the owner can submit search queries. We formally define SSE in this multi-user setting, and present an efficient construction.

In this article, we have revisited the problem of searchable symmetric encryption, which allows a client to store its data on a remote server in such a way that it can search over it in a private manner. We make several contributions including new security definitions and new constructions. Motivated by subtle problems in all previous security definitions for SSE, we propose new definitions and point out that the existing notions have significant practical drawbacks: contrary to the natural use of searchable encryption, they only guarantee security for users that perform all their searches at once. We address this limitation by introducing stronger definitions that guarantee security even when users perform more realistic searches. We also propose two new SSE constructions. Surprisingly, despite being provably secure under our stronger security definitions, these are the most efficient schemes to date and are (asymptotically) optimal (i.e., the work performed by the server per returned document is constant in the size of the data). Finally, we also consider multi-user SSE, which extends the searching ability to parties other than the owner.

**Deterministic and Efficiently Searchable Encryption**

We present as-strong-as-possible definitions of privacy, and constructions achieving them, for public-key encryption schemes where the encryption algorithm is deterministic. We obtain as a consequence database encryption methods that permit fast (i.e. sub-linear, and in fact logarithmic, time) search while provably providing privacy that is as strong as possible subject to this fast search constraint.

One of our constructs, called RSA-DOAEP, has the added feature of being length preserving, so that it is the first example of a public-key cipher. We generalize this to obtain a notion of efficiently-searchable encryption schemes which permit more flexible privacy to search-time trade-offs via a technique called bucketization. Our results answer muchasked questions in the database community and provide foundations for work done there.

**Public-Key Encryption with Fuzzy Keyword Search: A Provably Secure Scheme under Keyword Guessing Attack**

A lot of interest has been drawn recently into public-key encryption with keyword search (PEKS), which keeps publickey encrypted documents amendable to secure keyword search. However, PEKS resist against keyword guessing attack by assuming that the size of the keyword space is beyond the polynomial level. But this assumption is ineffective in practice. PEKS are insecure under keyword guessing attack. As we observe, the key to defend such attack is to avoid the availability of the exact search trapdoor to adversaries. Accordingly, we compromise the exactness of search trapdoor by mapping at least two different keywords into a fuzzy search trapdoor. We propose a novel concept called public-key encryption with fuzzy keyword search (PEFKS), by which the un-trusted server only obtains the fuzzy search trapdoor instead of the exact search trapdoor, and define its semantic security under chosen keyword attack (SS-CKA) and indistinguishability of keywords under non-adaptively chosen keywords and keyword guessing attack (IK-NCK-KGA). For the keyword space with and without uniform distribution, we respectively present two universal transformations from anonymous identity-based encryption to PEFKS, and prove their SSCKA and IK-NCK-KGA securities. To our knowledge, PEFKS is the first scheme to resist against keyword guessing attack on condition that the keyword space is not more than the polynomial level.

In PEKS, a proxy server, who responds the keyword queries of a receiver, can know the content of keywords by implementing KGA. Moreover, it is efficient under the practical condition that the size of the keyword space is not more than the polynomial level. In order to resist against KGA, we novelly defined public-key encryption with fuzzy keyword search (PEFKS) and its IK-NCK-KGA security. And we proposed two universal transformations from IBE to PEFKS under different conditions. Under the condition that the keyword space has uniform distribution, we proposed a SS-CKA and IK-NCK-KGA secure transformation PEFKS-UD, and provided an instance based on BF01 scheme. Under the condition that the keyword space has non-uniform distribution, we proposed another SS-CKA and IK-NCK-KGA secure transformation PEFKS-ND, and provided two methods to sort keywords, which is the key to realize PEFKS-ND. Beyond the perspective of cryptosystem, we discussed the biased advantage of KGA on PEFKSND, which is caused only by the non-uniform distribution of the keyword space. We illuminate that the biased advantage has been decreased as much as possible. So we made PEFKS-ND secure in a broad sense.

**Parallel and Dynamic Searchable Symmetric Encryption**

Searchable symmetric encryption (SSE) enables a client to outsource a collection of encrypted documents in the cloud and retain the ability to perform keyword searches without revealing information about the contents of the documents and queries. Although efficient SSE constructions are known, previous solutions are highly sequential. This is mainly due to the fact that, currently, the only method for achieving sub-linear time search is the inverted index approach which requires the search algorithm to access a sequence of memory locations, each of which is unpredictable and stored at the previous location in the sequence. Motivated by advances in multi-core architectures, we present a new method for constructing sub-linear SSE schemes. Our approach is highly parallelizable and dynamic. With roughly a logarithmic number of cores in place, searches for a keyword w in our scheme execute in o(r) parallel time, where r is the number of documents containing keyword w (with more cores, this bound can go down to O(log n), i.e., independent of the result size r). Such time complexity outperforms the optimal Θ(r) sequential search time—a similar bound holds for the updates. Our scheme also achieves the following important properties: (a) it enjoys a strong notion of security, namely security against adaptive chosen-keyword attacks; (b) compared to existing sub-linear dynamic SSE schemes, updates in our scheme do not leak any information, apart from information that can be inferred from previous search tokens; (c) it can be implemented efficiently in external memory (with logarithmic I/O overhead). Our technique is simple and uses a red-black tree data structure; its security is proven in the random oracle model.

**Dynamic Searchable Encryption via Blind Storage**

Dynamic Searchable Symmetric Encryption allows a client to store a dynamic collection of encrypted documents with a server, and later quickly carry out keyword searches on these encrypted documents, while revealing minimal information to the server. In this paper we present a new dynamic SSE scheme that is simpler and more efficient than existing schemes while revealing less information to the server than prior schemes, achieving fully adaptive security against honest-but-curious servers. We implemented a prototype of our scheme and demonstrated its efficiency on datasets from prior work. Apart from its concrete efficiency, our scheme is also simpler: in particular, it does not require the server to support any operation other than upload and download of data. Thus the server in our scheme can be based solely on a cloud storage service, rather than a cloud computation service as well, as in prior work. In building our dynamic SSE scheme, we introduce a new primitive called Blind Storage, which allows a client to store a set of files on a remote server in such a way that the server does not learn how many files are stored, or the lengths of the individual files; as each file is retrieved, the server learns about its existence (and can notice the same file being downloaded subsequently), but the file’s name and contents are not revealed. This is a primitive with several applications other than SSE, and is of independent interest.

In this work, we introduced a new cryptographic construct called Blind Storage, and implemented it using a novel, yet light-weight protocol SCATTERSTORE. We also showed how a dynamic SSE scheme can be constructed using Blind Storage, in a relatively simple manner. The resulting scheme is more computationally efficient, require simpler infrastructure, and is more secure than the existing schemes.

**Secure Ranked Keyword Search over Encrypted Cloud Data**

As Cloud Computing becomes prevalent, sensitive information are being increasingly centralized into the cloud. For the protection of data privacy, sensitive data has to be encrypted before outsourcing, which makes effective data utilization a very challenging task. Although traditional searchable encryption schemes allow users to securely search over encrypted data through keywords, these techniques support only boolean search, without capturing any relevance of data files. This approach suffers from two main drawbacks when directly applied in the context of Cloud Computing. On the one hand, users, who do not necessarily have pre-knowledge of the encrypted cloud data, have to postprocess every retrieved file in order to find ones most matching their interest; On the other hand, invariably retrieving all files containing the queried keyword further incurs unnecessary network traffic, which is absolutely undesirable in today’s pay-as-you-use cloud paradigm. In this paper, for the first time we define and solve the problem of effective yet secure ranked keyword search over encrypted cloud data. Ranked search greatly enhances system usability by returning the matching files in a ranked order regarding to certain relevance criteria (e.g., keyword frequency), thus making one step closer towards practical deployment of privacy-preserving data hosting services in Cloud Computing. We first give a straightforward yet ideal construction of ranked keyword search under the state-of-the-art searchable symmetric encryption (SSE) security definition, and demonstrate its inefficiency. To achieve more practical performance, we then propose a definition for ranked searchable symmetric encryption, and give an efficient design by properly utilizing the existing cryptographic primitive, order-preserving symmetric encryption (OPSE). Thorough analysis shows that our proposed solution enjoys “as-strong-as-possible” security guarantee compared to previous SSE schemes, while correctly realizing the goal of ranked keyword search. Extensive experimental results demonstrate the efficiency of the proposed solution.

In this paper, as an initial attempt, we motivate and solve the problem of supporting efficient ranked keyword search for achieving effective utilization of remotely stored encrypted data in Cloud Computing. We first give a basic scheme and show that by following the same existing searchable encryption framework, it is very inefficient to achieve ranked search. We then appropriately weaken the security guarantee, resort to the newly developed crypto primitive OPSE, and derive an efficient one-to-many order-preserving mapping function, which allows the effective RSSE to be designed. Through thorough security analysis, we show that our proposed solution is secure and privacy-preserving, while correctly realizing the goal of ranked keyword search.

**Privacy Preserving Multi-Keyword Ranked Search with Anonymous ID Assignment over Encrypted Cloud Data**

The advancement in cloud computing has motivated the data owners to outsource their data management systems from local sites to commercial public cloud for great flexibility and economic savings. But people can enjoy full benefit of cloud computing if we are able to address very real privacy and security concerns that come with storing sensitive personal information. For real privacy, user identity should remain hidden from CSP (Cloud service provider) and to protect privacy of data, data which is sensitive is to be encrypted before outsourcing. Thus, enabling an encrypted cloud data search service is of great importance. By considering the large number of data users, documents in the cloud, it is important for the search service to allow multikeyword query and provide result similarity ranking to meet the effective need of data retrieval search and not often differentiate the search results. In this system, we define and solve the challenging problem of privacy-preserving multikeyword ranked search over encrypted cloud data (MRSE), and establish a set of strict privacy requirements for such a secure cloud data utilization system to be implemented in real. We first propose a basic idea for the Multi-keyword Ranked Search over Encrypted cloud data (MRSE) based on secure inner product computation and efficient similarity measure of coordinate matching, i.e., as many matches as possible, in order to capture the relevance of data documents to the search query, then we give two significantly improved MRSE schemes to achieve various stringent privacy requirements in two different threat models. Assignment of anonymous ID to the user to provide more security to the data on cloud server is done. To improve the search experience of the data search service, further extension of the two schemes to support more search semantics is done.

The previous work mainly focused on providing privacy to the data on cloud in which using multi-keyword ranked search was provided over encrypted cloud data using efficient similarity measure of co-ordinate matching. The previous work also proposed a basic idea of MRSE using secure inner product computation. There was a need to provide more real privacy which this paper presents. In this system, stringent privacy is provided by assigning the cloud user a unique ID. This user ID is kept hidden from the cloud service provider as well as the third party user in order to protect the user’s data on cloud from the CSP and the third party user. Thus, by hiding the user’s identity, the confidentiality of user’s data is maintained.