

PRIVACY PRESERVING RANKED MULTI-KEYWORD SEARCH FOR MULTIPLE DATA OWNERS IN CLOUD COMPUTING

K.Madhavi¹, Katepalli Suresh Kumar²

¹Asst. Professor, Department of IT, Vr Siddhartha Engg College, Itsmadhavi12@gmail.com

²M.Tech, Department of IT, Vr Siddhartha Engg College, rajakatepalli@gmail.com

ABSTRACT

With the advent of cloud computing, it has become increasingly popular for data owners to outsource their data to public cloud servers while allowing data users to retrieve this data. For privacy concerns, secure searches over encrypted cloud data have motivated several research works under the single owner model. However, most cloud servers in practice do not just serve one owner; instead, they support multiple owners to share the benefits brought by cloud computing. In this paper, we propose schemes to deal with Privacy preserving Ranked Multi-keyword Search in a Multi-owner model (PRMSM). To enable cloud servers to perform secure search without knowing the actual data of both keywords and trapdoors, we systematically construct a novel secure search protocol. To rank the search results and preserve the privacy of relevance scores between keywords and files, we propose a novel Additive Order and Privacy Preserving Function family. To prevent the attackers from eavesdropping secret keys and pretending to be legal data users submitting searches, we propose a novel dynamic secret key generation protocol and a new data user authentication protocol. Furthermore, PRMSM supports efficient data user revocation. Extensive experiments on real-world datasets confirm the efficacy and efficiency of PRMSM.

Keywords: cloud computing, multiple owners, privacy preserving.

INTRODUCTION

Cloud computing is a subversive technology that is changing the way IT hardware and software are designed and purchased. As a new model of computing, cloud computing provides abundant benefits including easy access, decreased costs, quick deployment and flexible resource management, etc. Enterprises of all sizes can leverage the cloud to increase innovation and collaboration. Despite the abundant benefits of cloud computing, for privacy concerns, individuals and enterprise users are reluctant to outsource their sensitive data, including emails, personal health records and government confidential files, to the cloud. This is because once sensitive data are outsourced to a remote cloud, the corresponding data owners lose direct control of these data. Cloud service providers (CSPs) would promise to ensure owners' data security using mechanisms like virtualization and firewalls. However, these mechanisms do not protect owners' data privacy from the CSP itself, since the CSP possesses full control of cloud hardware, software, and owners' data. Encryption on sensitive data before outsourcing can preserve data privacy against CSP. However, data encryption makes the traditional data utilization service based on plaintext keyword search a very challenging problem. A trivial solution to this problem is to download all the encrypted data and decrypt them locally. However, this method is obviously impractical because it will cause a huge amount of communication overhead. Therefore, developing a secure search service over encrypted cloud data is of paramount importance.

Secure search over encrypted data has recently attracted the interest of many researchers. Song et al. first define and solve the problem of secure search over encrypted data. They propose the conception of searchable encryption, which is a cryptographic primitive that enables users to perform a keyword-based search on an encrypted dataset, just as on a plaintext dataset. Searchable encryption is further developed. However, these schemes are concerned mostly with single or boolean keyword search. Extending these techniques for ranked multikeyword search will incur heavy computation and storage costs. Secure search over encrypted cloud data is first defined by Wang et al. and further developed. These researches not only reduce the computation and storage cost for secure keyword search over encrypted cloud data, but also enrich the category of search function, including secure ranked multi-keyword search, fuzzy keyword search, and similarity search. However, all these schemes are limited to the single-owner model. As a matter of fact, most cloud servers in

practice do not just serve one data owner; instead, they often support multiple data owners to share the benefits brought by cloud computing. For example, to assist the government in making satisfactory policies on health care service, or to help medical institutions conduct useful research, some volunteer patients would agree to share their health data on the cloud. To preserve their privacy, they will encrypt their own health data with their secret keys. In this scenario, only the authorized organizations can perform a secure search over this encrypted data contributed by multiple data owners. Such a Personal Health Record sharing system, where multiple data owners are involved, can be found at mymedwall.com. Compared with the single-owner scheme, developing a full-fledged multi-owner scheme will have many new challenging problems. First, in the single owner scheme, the data owner has to stay online to generate trapdoors (encrypted keywords) for data users. However, when a huge amount of data owners are involved, asking them to stay online simultaneously to generate trapdoors would seriously affect the flexibility and usability of the search system. Second, since none of us would be willing to share our secret keys with others, different data owners would prefer to use their own secret keys to encrypt their secret data. Consequently, it is very challenging to perform a secure, convenient, and efficient search over the data encrypted with different secret keys. Third, when multiple data owners are involved, we should ensure efficient user enrolment and revocation mechanisms, so that our system enjoys excellent security and scalability.

We propose PRMSM, a privacy preserving ranked multi-keyword search protocol in a multi-owner cloud model. To enable cloud servers to perform secure search without knowing the actual value of both keywords and trapdoors, we systematically construct a novel secure search protocol. As a result, different data owners use different keys to encrypt their files and keywords. Authenticated data users can issue a query without knowing secret keys of these different data owners. To rank the search results and preserve the privacy of relevance scores between keywords and files, we propose a new additive order and privacy preserving function family, which helps the cloud server return the most relevant search results to data users without revealing any sensitive information. To prevent the attackers from eavesdropping secret keys and pretending to be legal data users submitting searches, we propose a novel dynamic secret key generation protocol and a new data user authentication protocol. As a result, attackers who steal the secret key and perform illegal searches would be easily detected. Furthermore, when we want to revoke a data user, PRMSM ensures efficient data user revocation. Extensive experiments on real-world datasets confirm the efficacy and efficiency of our proposed schemes.

IMPLEMENTATION

MODULES:

- ❖ System Model
- ❖ Data User Authentication
- ❖ Illegal Search Detection
- ❖ Search over Multi-owner

MODULES DESCRIPTION:

System Model

- ✓ In the first module, we develop the System Model to implement our proposed system. Our System model consists of Admin, users, data owners, and Cloud Servers. Admin provides the accessibility to Data-owners. Initially Data-owner needs to register and admin approves the each data owner request. The respective Password and login credentials will be sent to the Email ID of Data owner.
- ✓ In Users sub-module, Each user has a global identity in the system. A user may be entitled a set of attributes which may come from multiple attribute authorities. The user will receive a secret key associated with its attributes entitled by the corresponding attribute authorities.
- ✓ In data owners sub-module, the proposed scheme should allow new data owners to enter this system without affecting other data owners or data users, i.e., the scheme should support data owner scalability in a plug-and-play model.

- ✓ In Cloud Server sub-module of system model, the owner sends the encrypted data to the cloud server through Admin. They do not rely on the server to do data access control. But, the access control happens inside the cryptography. That is only when the user's attributes satisfy the access policy defined in the cipher text; the user is able to decrypt the ciphertext. Thus, users with different attributes can decrypt different number of content keys and thus obtain different granularities of information from the same data

Data User Authentication

- ✓ To prevent attackers from pretending to be legal data users performing searches and launching statistical attacks based on the search result, data users must be authenticated before the administration server re-encrypts trapdoors for data users. Traditional authentication methods often follow three steps. First, data requester and data authenticator share a secret key, say, k_0 . Second, the requester encrypts his personally identifiable information d_0 using k_0 and sends the encrypted data $(d_0)k_0$ to the authenticator. Third, the authenticator decrypts the received data with k_0 and authenticates the decrypted data.
- ✓ The key point of a successful authentication is to provide both the dynamically changing secret keys and the historical data of the corresponding data user.

Illegal Search Detection

- ✓ In our scheme, the authentication process is protected by the dynamic secret key and the historical information. We assume that an attacker has successfully eavesdropped the secret key. Then he has to construct the authentication data; if the attacker has not successfully eavesdropped the historical data, e.g., the request counter, the last request time, he cannot construct the correct authentication data. Therefore this illegal action will soon be detected by the administration server.
- ✓ Further, if the attacker has successfully eavesdropped all data of U_j , the attacker can correctly construct the authentication data and pretend himself to be U_j without being detected by the administration server. However, once the legal data user U_j performs his search, since the secret key on the administration server side has changed, there will be contradictory secret keys between the administration server and the legal data user. Therefore, the data user and administration server will soon detect this illegal action.

Search over Multi-owner

- ✓ The proposed scheme should allow multi-keyword search over encrypted files which would be encrypted with different keys for different data owners. It also needs to allow the cloud server to rank the search results among different data owners and return the top- k results. The cloud server stores all encrypted files and keywords of different data owners.
- ✓ The administration server will also store a secret data on the cloud server. Upon receiving a query request, the cloud will search over the data of all these data owners. The cloud processes the search request in two steps. First, the cloud matches the queried keywords from all keywords stored on it, and it gets a candidate file set. Second, the cloud ranks files in the candidate file set and finds the most top- k relevant files. Finally, we apply the proposed scheme to encode the relevance scores and obtain the top- k search results.

EXISTING SYSTEM:

- ❖ Secure search over encrypted data has recently attracted the interest of many researchers. Song et al. first define and solve the problem of secure search over encrypted data. They propose the conception of searchable encryption, which is a cryptographic primitive that enables users to perform a keyword-based search on an encrypted dataset, just as on a plaintext dataset. Searchable encryption is further developed.
- ❖ Secure search over encrypted cloud data is first defined by Wang et al. and further developed. These researches not only reduce the computation and storage cost for secure keyword search over encrypted

cloud data, but also enrich the category of search function, including secure ranked multi-keyword search, fuzzy keyword search, and similarity search.

DISADVANTAGES OF EXISTING SYSTEM:

- ❖ Existing schemes are concerned mostly with single or boolean keyword search.
- ❖ All the existing schemes are limited to the single-owner model. As a matter of fact, most cloud servers in practice do not just serve one data owner; instead, they often support multiple data owners to share the benefits brought by cloud computing.

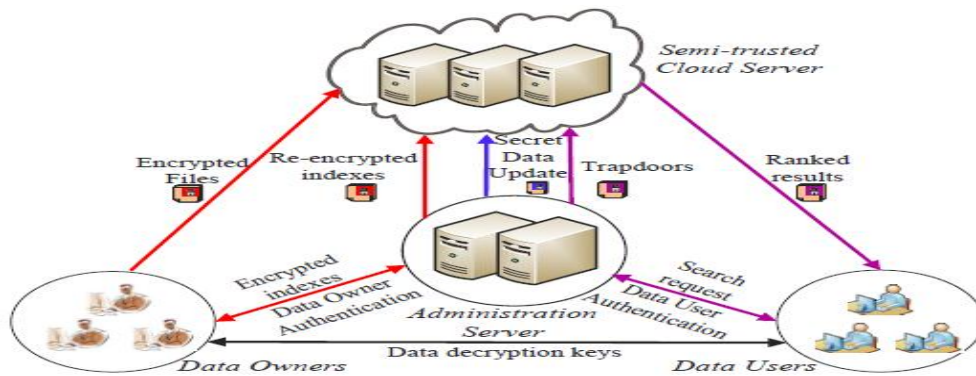
PROPOSED SYSTEM:

- ❖ In this paper, we propose PRMSM, a privacy preserving ranked multi-keyword search protocol in a multi-owner cloud model.
- ❖ We define a multi-owner model for privacy preserving keyword search over encrypted cloud data.
- ❖ We propose an efficient data user authentication protocol, which not only prevents attackers from eavesdropping secret keys and pretending to be illegal data users performing searches, but also enables data user authentication and revocation.
- ❖ We systematically construct a novel secure search protocol, which not only enables the cloud server to perform secure ranked keyword search without knowing the actual data of both keywords and trapdoors, but also allows data owners to encrypt keywords with self-chosen keys and allows authenticated data users to query without knowing these keys.
- ❖ We propose an Additive Order and Privacy Preserving Function family (AOPPF) which allows data owners to protect the privacy of relevance scores using different functions according to their preference, while still permitting the cloud server to rank the data files accurately.
- ❖ We conduct extensive experiments on real-world datasets to confirm the efficacy and efficiency of our proposed schemes.

ADVANTAGES OF PROPOSED SYSTEM:

- ❖ The proposed scheme allows multi-keyword search over encrypted files which would be encrypted with different keys for different data owners.
- ❖ The proposed scheme allows new data owners to enter this system without affecting other data owners or data users, i.e., the scheme support data owner scalability in a plug-and-play model.
- ❖ The proposed scheme ensures that only authenticated data users can perform correct searches. Moreover, once a data user is revoked, he can no longer perform correct searches over the encrypted cloud data.
- ❖ To enable cloud servers to perform secure search without knowing the actual value of both keywords and trapdoors, we systematically construct a novel secure search protocol. As a result, different data owners use different keys to encrypt their files and keywords. Authenticated data users can issue a query without knowing secret keys of these different data owners.
- ❖ To rank the search results and preserve the privacy of relevance scores between keywords and files, we propose a new additive order and privacy preserving function family, which helps the cloud server return the most relevant search results to data users without revealing any sensitive information.
- ❖ To prevent the attackers from eavesdropping secret keys and pretending to be legal data users submitting searches, we propose a novel dynamic secret key generation protocol and a new data user authentication protocol.

SYSTEM ARCHITECTURE



SCREEN SHOT



Fig: Home Page

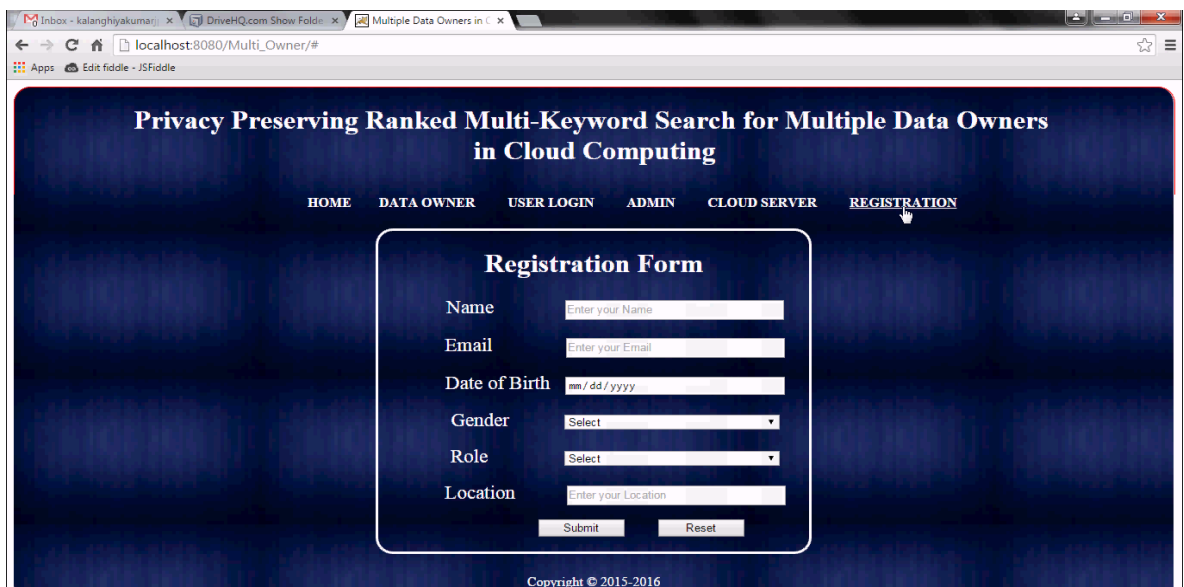


Fig: Registration Page

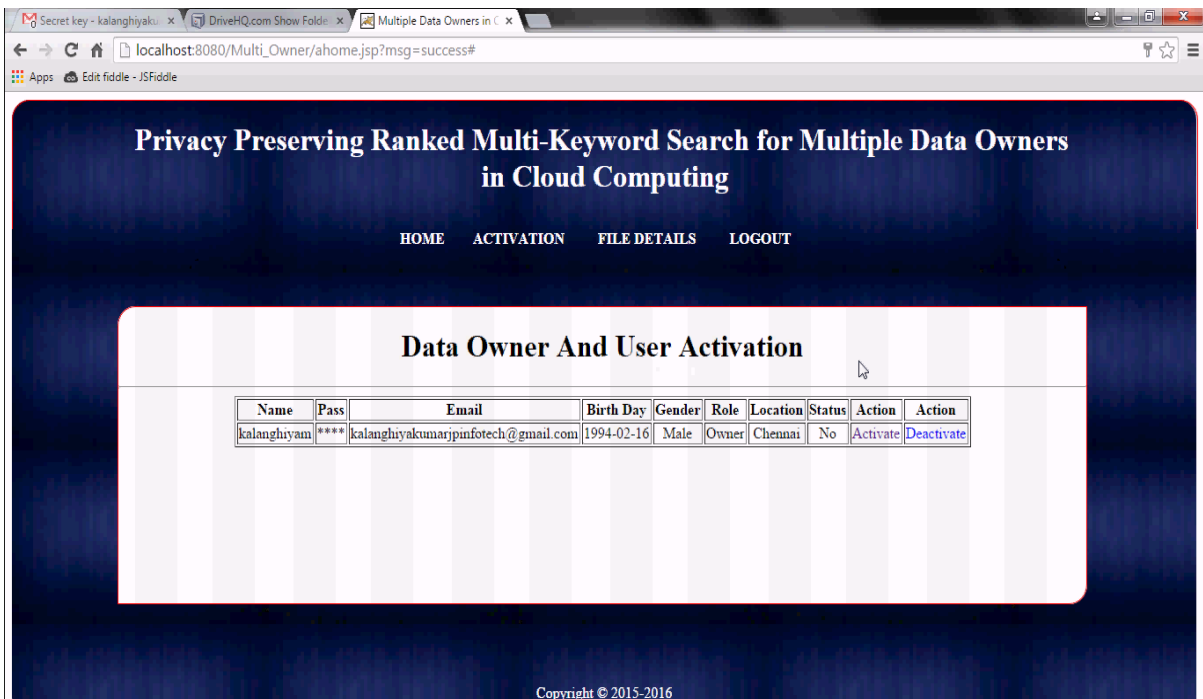


Fig: Data Owner and User Activation

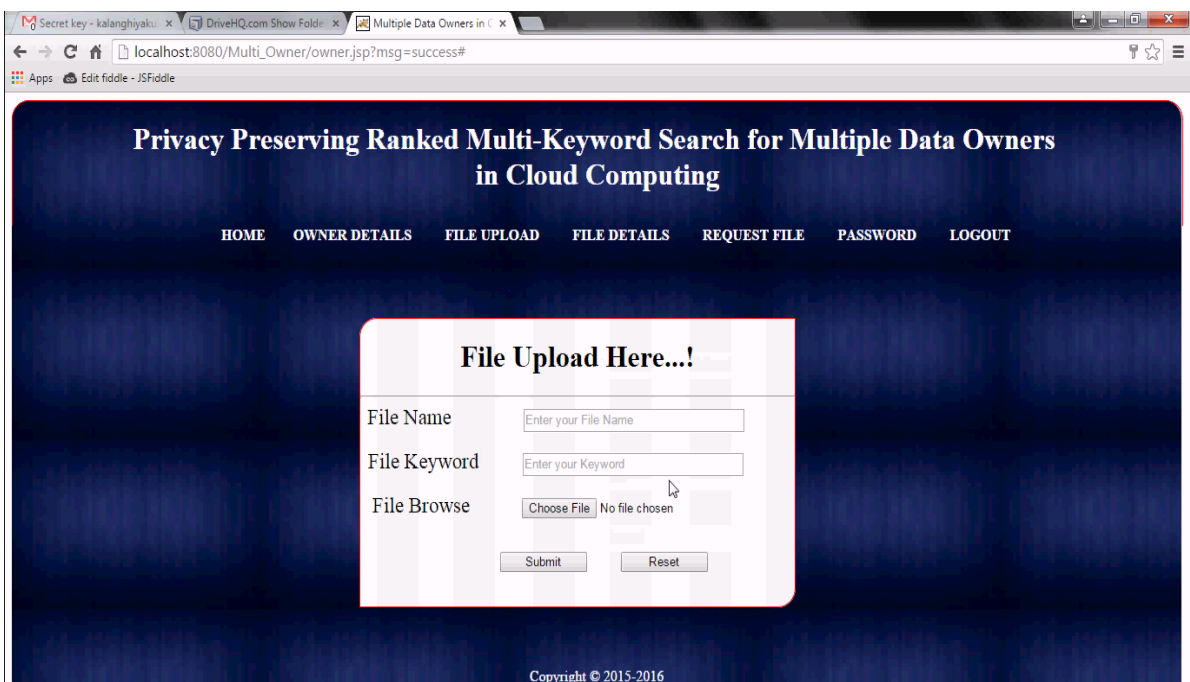


Fig: File Upload

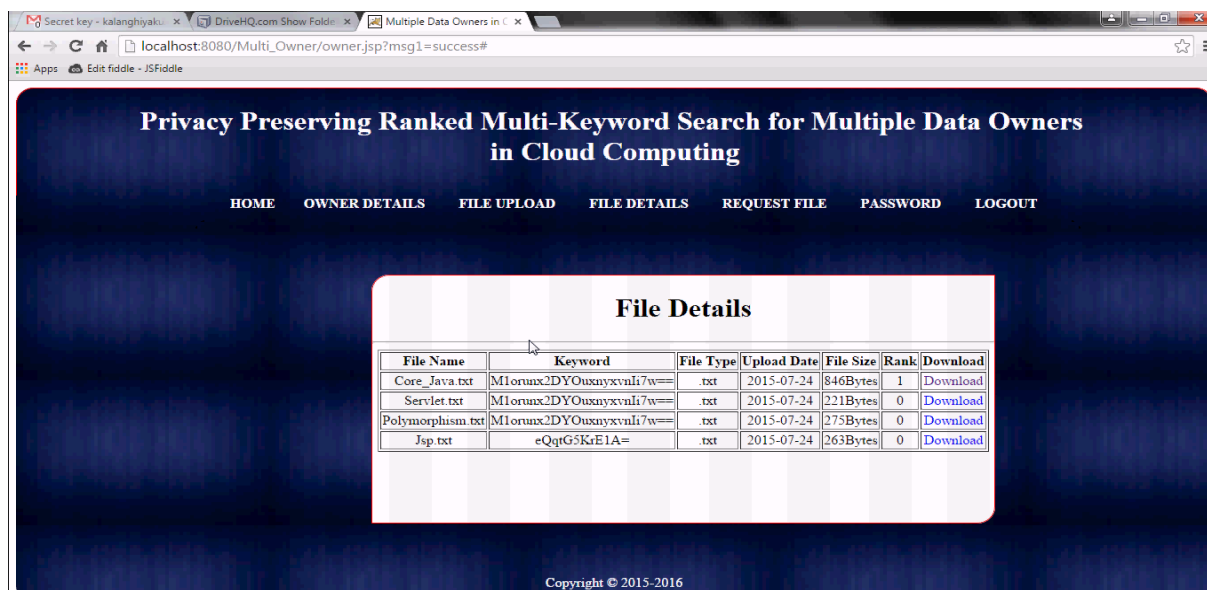


Fig: File Download

CONCLUSION

we explore the problem of secure multi-keyword search for multiple data owners and multiple data users in the cloud computing environment. Different from prior works, our schemes enable authenticated data users to achieve secure, convenient, and efficient searches over multiple data owners' data. To efficiently authenticate data users and detect attackers who steal the secret key and perform illegal searches, we propose a novel dynamic secret key generation protocol and a new data user authentication protocol. To enable the cloud server to perform secure search among multiple owners' data encrypted with different secret keys, we systematically construct a novel secure search protocol. To rank the search results and preserve the privacy of relevance scores between keywords and files, we propose a novel Additive Order and Privacy Preserving Function family. Moreover, we show that our approach is computationally efficient, even for large data and keyword sets. As our future work, on one hand, we will consider the problem of secure fuzzy keyword search in a multi-owner paradigm. On the other hand, we plan to implement our scheme on the commercial clouds.

REFERENCES

- [1] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "A view of cloud computing," *Communication of the ACM*, vol. 53, no. 4, pp. 50–58, 2010.
- [2] C. Wang, S. S. Chow, Q. Wang, K. Ren, and W. Lou, "Privacy-preserving public auditing for secure cloud storage," *Computers, IEEE Transactions on*, vol. 62, no. 2, pp. 362–375, 2013.
- [3] D. Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in *Proc. IEEE International Symposium on Security and Privacy (S&P'00)*, Nagoya, Japan, Jan. 2000, pp. 44–55.
- [4] E. Goh. (2003) Secure indexes. [Online]. Available: <http://eprint.iacr.org/>
- [5] R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky, "Searchable symmetric encryption: improved definitions and efficient constructions," in *Proc. ACM CCS'06*, VA, USA, Oct. 2006, pp. 79–88.
- [6] D. B. et al., "Public key encryption with keyword search secure against keyword guessing attacks without random oracle," *EUROCRYPT*, vol. 43, pp. 506–522, 2004.

- [7] P. Golle, J. Staddon, and B. Waters, "Secure conjunctive keyword search over encrypted data," in *Proc. Applied Cryptography and Network Security (ACNS'04)*, Yellow Mountain, China, Jun. 2004, pp. 31–45.
- [8] L. Ballard, S. Kamara, and F. Monrose, "Achieving efficient conjunctive keyword searches over encrypted data," in *Proc. Information and Communications Security (ICICS'05)*, Beijing, China, Dec. 2005, pp. 414–426.
- [9] C. Wang, N. Cao, J. Li, K. Ren, and W. Lou, "Secure ranked keyword search over encrypted cloud data," in *Proc. IEEE Distributed Computing Systems (ICDCS'10)*, Genoa, Italy, Jun. 2010, pp. 253–262.
- [10] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacypreserving multi-keyword ranked search over encrypted cloud data," in *Proc. IEEE INFOCOM'11*, Shanghai, China, Apr. 2011, pp. 829–837.