On the Security of a Privacy-Preserving Ranked Multi-Keyword Search Scheme

Ziba Eslami[1]*, Mahnaz Noroozi[1], and Joonsang Baek[2]
[1]*Department of Computer and Data Science, Shahid Beheshti University, G.C., Tehran, Iran*
[2]*Institutue of Cybersecurity and Cryptology, University of Wollongong, Australia*

Received: January 10, 2019; Accepted: March 5, 2019; Published: March 31, 2019

Abstract

Ranked keyword search over encrypted cloud data is a hot research topic with huge practical implications. A complex variant of this problem is to allow searching among ciphertexts belonging to multiple data owners. Unfortunately, there exist only a few papers in the literature which address this important setting. Recently, Zhang et al. proposed a solution to this problem and claimed that their scheme provides two fundamental security requirements: ciphertext and trapdoor indistinguishability. We prove, however, that in their scheme neither the ciphertexts nor the trapdoors achieve indistinguishability. Our result shows that their scheme is insecure to be used in practice.

Keywords: Searchable encryption, Ranked keyword search, Security, Privacy.

1 Introduction

Nowadays, outsourcing data to third party cloud servers has brought more convenience by making data access possible from anywhere. However, since uploaded data may now be easily exposed to a variety of threats such as unauthorized access or information leakage, security and privacy of sensitive documents stored on the servers are important challenges that organizations as well as individuals must consider. The most naive approach to surmount the challenge is to use some sort of encryption to ensure confidentiality. Nevertheless, this technique makes searching a formidable task in turn. To overcome this problem, a cryptographic notion called *searchable encryption* (SE) is introduced by which users can perform search on encrypted data and at the same time maintain information privacy.

SE techniques usually provide search ability over encrypted documents by extracting the keywords of those documents and generating searchable ciphertexts corresponding to these keywords. In the basic setup, as shown by Fig. 1, there are three entities in SE [1]: (1) a data owner whose task is to generate searchable ciphertexts C_{w_i} corresponding to keywords w_i and upload the results (along with the main encrypted document) on the cloud, (2) a data receiver who can search for the keyword w within uploaded ciphertexts through generating a trapdoor T_w, and (3) a cloud server which by using a received trapdoor T_w runs an algorithm called Test to find the ciphertexts containing w and returns the corresponding encrypted documents.

There exist different types of SE schemes in the literature [2, 3, 4, 5]; however, the minimum essential security requirements for any variation of searchable encryption schemes are "ciphertext indistinguishability" and "trapdoor indistinguishability". Informally, these security notions capture the idea that neither the ciphertexts nor the trapdoors should reveal any information about the corresponding keywords to a

Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, 10:1 (March 2019), pp. 75-85
DOI: 10.22667/JOWUA.2019.03.31.075

*Corresponding Author: Department of Computer and Data Science, Shahid Beheshti University, G.C., Tehran, Iran, Email: z_eslami@sbu.ac.ir, Web: http://facultymembers.sbu.ac.ir/eslami/, Tel: +98-2129903007, Fax: +98-2122431655

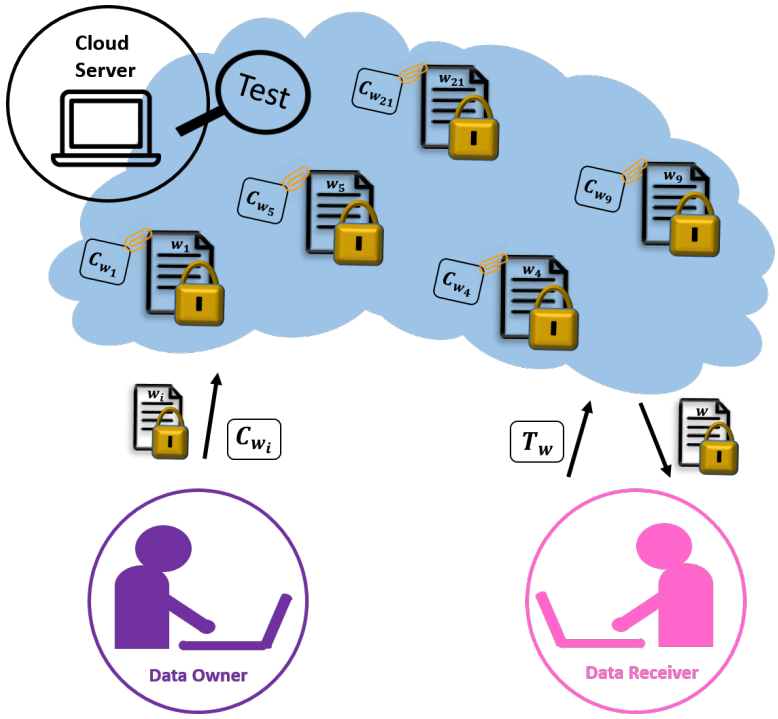

Figure 1: Entities involved in SE

polynomially-bounded active adversary. In other words, such adversaries should not be able to distinguish between the searchable ciphertexts or trapdoors corresponding to two different keywords. We will formally define these security concepts in Section 3.

1.1 Literature Review

First introduced by Song et al. in 2000 [6], searchable encryption has since been the subject of extensive research [7, 8, 9, 10, 11]. A diverse collection of schemes are proposed in the literature under different threat models to achieve various search functionalities, such as single keyword search [12], similarity search [13], fuzzy keyword search [14], multi-keyword boolean search [15], ranked keyword search [16] and etc.

In particular, ranked keyword search, which allows a cloud server to permute the search results according to certain relevance criteria, is the focus of attention due to its practical applications. In this setting, the server processes the search requests in two steps: first, matching the queried keywords from all keywords stored on the cloud and getting a candidate file set, and next ranking the candidate files to find the most relevant ones. Therefore, instead of going through every matched document in the search result, data receivers can quickly find relevant information. This feature is really desirable in the pay-as-you-go cloud computing paradigm.

In [17], Wang et al. proposed the first solution for ranked keyword search over encrypted cloud data. They explored the statistical measure approach from IR and text-mining to provide the relevance score of each document before saving their encrypted form on the cloud. Moreover, the authors integrated the cryptographic primitive of *order-preserving symmetric encryption* (OPSE) [18] to protect the frequency information to be revealed through relevance scores. The authors later improved the scheme of [17] and adopted the statistical measure approach to build a secure searchable ciphertext in [19]. However, both schemes can be used only in single-keyword search setting. In many scenarios, there are large data sets

and therefore, a search query with a single keyword may result in large amount of matched documents due to lack of search precision in defining target documents. Nevertheless, usually only a few of them are relevant and a user has to apply several queries to take the intersection of the results. Thereupon, a serious burden is imposed on the user in terms of computation and communication overhead. In 2011, Cao et al. realized the first privacy-preserving multi-keyword ranked search scheme [20] in which a conjunction of several keywords is considered in a single query by representing documents and queries as vectors. The authors later published an improved version of their scheme in [20]. In 2013, Sun et al. presented a multi-keyword search scheme [21] that supports similarity-based ranking. Their approach employs a tree-based index structure and ranking capability is achieved based on cosine similarity measure in vector space model. In [22], Orencik et al. utilized local sensitive hash functions to cluster similar documents in multi-keyword search setting. In [23], the authors defined and solved the problem of privacy-preserving ranked fuzzy keyword search over encrypted cloud data. The notion of fuzzy keyword search was first introduced by Li et al. [24] in which minor typos and format inconsistencies are allowed. In 2015, a searchable encryption scheme with multi-keyword ranked search was proposed in which the blind storage system is adopted to conceal the access pattern [25]. Later in 2017, Jiang et al. presented a multi-keyword ranked search scheme over encrypted cloud data which also supports search results verification [26]. The authors stated that in real world applications, search results may contain corrupted data caused by inevitable human errors, failures of the underlying hardware/software or the attacks from hackers. So it would be beneficial to provide users with a verifiable mechanism to assure the correctness and the completeness of search results.

The above-mentioned schemes are all in single-data-owner setting where searching is performed among encrypted documents which belong to a specific owner. Clearly, this is a rather restrictive scenario while the more practical one is multiple-data-owner model which is common in a data sharing system. In many situations, we need to provide search ability for authorized users amongst encrypted documents contributed by multiple data owners. In 2014, Zhang et al. [27] firstly defined a multi-owner model for secure ranked keyword search over encrypted data. In this scheme, data owners encrypt keywords in such a way that data receivers can generate trapdoors without knowing the keys used during encryption. Then in 2016, the same authors published an improved version [28] in which a trusted administrator is introduced in the system model to further protect the secret keys of users and enable user authentication and revocation. Their aim was to prevent attackers from eavesdropping secret keys and pretending to be legal data users who perform searches. In this paper, we are primarily concerned with the material presented in [28].

1.2 Contribution and Organization

We recall that regardless of the variation of searchable encryption used, at minimum, two basic security requirements should be satisfied: ciphertext indistinguishability and trapdoor indistinguishability. In particular, ranked keyword search (over encrypted data) schemes are no exception and failure to achieve these properties renders them completely insecure.

In [28], Zhang et al. proposed a ranked keyword search scheme in multi-owner setting. The aim of this paper is to show how any attacker can distinguish both the ciphertexts and the trapdoors generated by their scheme. In other words, we prove that, contrary to the claims made in [28], their scheme achieves neither ciphertext indistinguishability nor trapdoor indistinguishability. Therefore, this scheme can not be used in practice.

The rest of this paper is organized as follows. A brief review of Zhang et al.'s scheme [28] is covered in Section 2. Then, in Section 3, we demonstrate the essential security requirements of a searchable encryption scheme. In Section 4, we prove that the scheme of Zhang et al. is insecure in the sense that ciphertexts and trapdoors are distinguishable. Finally, Section 5 concludes the paper.

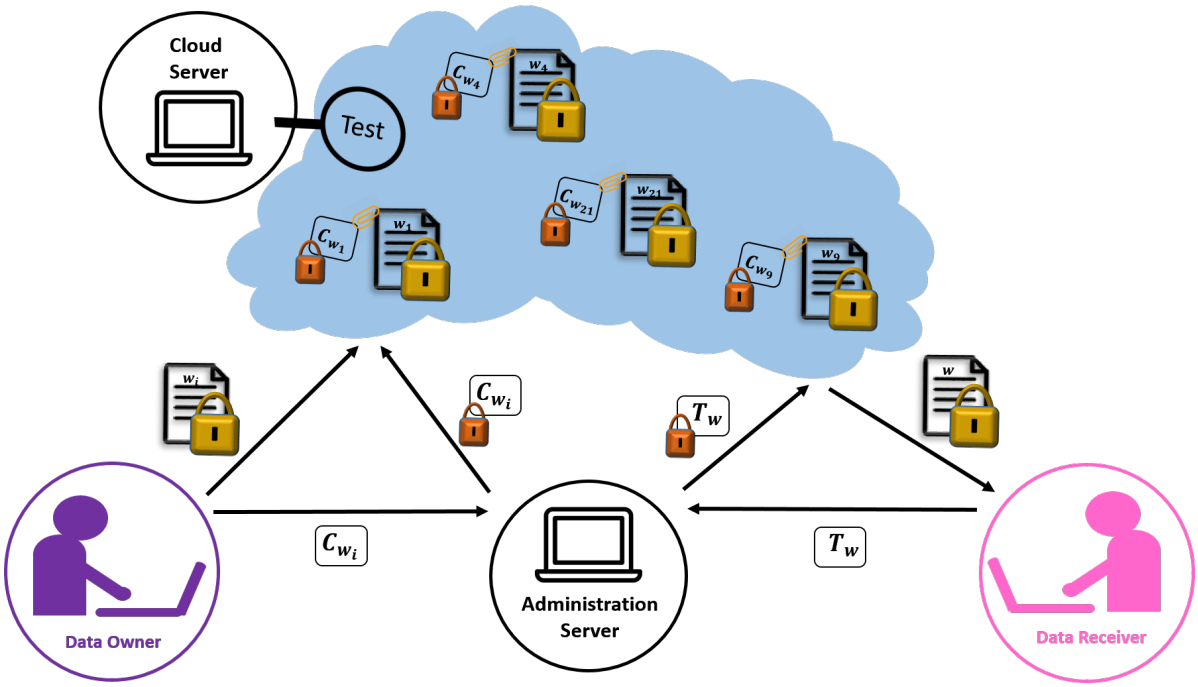

Figure 2: Entities involved in the scheme of [28]

2 Review of Zhang et al.'s scheme

In this section, we briefly review the SE scheme of Zhang et al. [28] in which the search requests are processed in two separated steps:

- finding correspondence between the queried encrypted keywords and all encrypted keywords stored on the cloud and getting a candidate file set of matching results,
- ranking the candidate files using a metric to find the most relevant ones and returning them.

In [28], the authors use 'sum of relevance scores' as the metric for ranking the search results. This metric helps the cloud server to return relevant search results without revealing any sensitive information. They proposed an *additive order and privacy preserving function* (AOPPF) family 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. However, since our reported security flaws regarding Zhang et al.'s scheme are not related to the ranking step, in this section, we only cover the algorithms corresponding to the first step of their protocol.

As Fig. 2 shows, in the system model of Zhang et al.'s scheme, there also exists another entity called administration server who receives the encrypted keyword from the data owner and after re-encrypting it, sends the resulting ciphertext to the cloud server. Furthermore, the generated trapdoors by the data receivers are first delivered to the administration server who re-encrypts the trapdoors and submits them to the cloud server. The searchable encryption scheme of Zhang et al. is composed of the following algorithms.

2.1 Setup

- Two cyclic groups G (with generator g) and G_1 (with generator g_1) of order p are chosen.

- A bilinear map $\hat{e} : G \times G \longrightarrow G_1$ and a hash function $H : \{0,1\}^* \rightarrow \mathbb{Z}_p^*$ are chosen.
- Using a randomized key generation algorithm, two keys k_{a_1}, k_{a_2} are produced which are the private keys of the administration server. Also, corresponding to each data owner O_i, the private key k_i is generated.

2.2 Keyword Encryption

- Data owner O_i who wants to encrypt w computes $E_a^{'} = g^{k_i.r_o.H(w)}$ and $E_o = g^{k_i.r_o}$, where r_o is a randomly generated number. Then, $C_w^{'} = (E_a^{'}, E_o)$ is sent to the administration server.
- The administration server computes $E_a = (E_a^{'}.g^{k_{a_1}})^{k_{a_2}}$ and sends $C_w = (E_a, E_o)$ to the cloud server.

2.3 Trapdoor Generation

- The data receiver who wants to search for a keyword w computes $T_w^{'} = (g^{H(w).r_u}, g^{r_u})$ (where r_u is a randomly generated number) and sends it to the administration server.
- The administration server upon receiving $T_w^{'} = (X,Y)$ computes $T_1 = X^{k_{a_1}.k_{a_2}.r_a}$, $T_2 = Y^{k_{a_1}}$ and $T_3 = Y^{k_{a_1}.r_a}$ where r_a is a randomly generated number. Then, it computes $S_a = g^{k_{a_1}.k_{a_2}.r_a}$ and sends $T_w = (T_1, T_2, T_3)$ together with S_a to the cloud server.

2.4 Keyword Matching

- On input $C_w = (E_a, E_o)$, $T_{w'} = (T_1, T_2, T_3)$ and S_a, the cloud server returns 'True' if the following equation holds and 'False' otherwise.
  $\hat{e}(E_a, T_3) =^? \hat{e}(E_o, T_1).\hat{e}(S_a, T_2)$

3 Essential security requirements of a searchable encryption scheme

Any secure searchable encryption scheme should provide two essential properties: ciphertext indistinguishability and trapdoor indistinguishability. These notions informally state that the adversary cannot distinguish between ciphertexts or trapdoors corresponding to two different keywords with a probability better than $1/2$ [5]. In order to formalize these notions, the following security games are defined in the literature which are played between a challenger and an adversary who wants to break the indistinguishability of the scheme. During the games, appropriate oracles accesses are provided for the adversary to simulate the disclosure of information in the real world.

Note that there also exist some other security requirements regarding the ranking step of a ranked keyword search scheme. However, since our reported security flaws are not related to this step, we do not cover the corresponding security notions here and deliberately omit ranking-related material.

Consider the following two games defined for any searchable encryption scheme Π =(Setup,Keyword Encryption,Trapdoor Generation,Keyword Matching) and a probabilistic polynomial-time adversary $\mathscr{A}$.

Ciphertext indistinguishability game ($Game_{\mathscr{A},\Pi}^{C-IND}$):

- **Initialization phase:** Setup is run to generate the required parameters.
- **Query phase 1:** $\mathscr{A}$ can adaptively query the following oracles:
  - $\mathscr{O}^{Ciphertext}(w)$: This query returns $C_w^{'}$ and C_w.

  - $\mathcal{O}^{Trapdoor}(w)$: This query returns $T_w^{'}$ and T_w.
  - $\mathcal{O}^{AdminCipher}(C_w^{'})$: This query returns C_w.
  - $\mathcal{O}^{AdminTrap}(T_w^{'})$: This query returns T_w.[1]

- **Challenge phase:** $\mathscr{A}$ outputs two keywords w_0^* and w_1^* which have not been queried before. Then, $\mu \in \{0,1\}$ is randomly chosen and $C_{w_\mu^*}^{'}$ and $C_{w_\mu^*}$ are computed according to the Keyword Encryption algorithm. These values are returned to $\mathscr{A}$.

- **Query phase 2:** $\mathscr{A}$ can continue to adaptively access the oracles of Phase 1 (for keywords other than w_0 and w_1).

- **Response phase:** $\mathscr{A}$ outputs his guess $\mu^{'}$ and wins the game if and only if $\mu^{'} = \mu$.

We define $Adv_{\mathscr{A},\Pi}^{C-IND} := |Pr[\mu = \mu^{'}] - \frac{1}{2}|$.

Definition 1. *A searchable encryption scheme satisfies ciphertext indistinguishability against an adaptive chosen keyword attack if for any polynomial-time adversary $\mathscr{A}$, $Adv_{\mathscr{A},\Pi}^{C-IND}$ is negligible.*

Trapdoor indistinguishability game ($Game_{\mathscr{A},\Pi}^{T-IND}$):

- **Initialization:** Setup is run to generate the required parameters.

- **Phase 1:** $\mathscr{A}$ can adaptively query $\mathcal{O}^{Ciphertext}$, $\mathcal{O}^{Trapdoor}$, $\mathcal{O}^{AdminCipher}$ and $\mathcal{O}^{AdminTrap}$ (defined in previous game).

- **Challenge:** $\mathscr{A}$ outputs two keywords w_0^* and w_1^* which are not queried before. Then, $\mu \in \{0,1\}$ is randomly chosen and $T_{w_\mu^*}^{'}$ and $T_{w_\mu^*}$ are computed according to the Trapdoor Generation algorithm. These values are returned to $\mathscr{A}$.

- **Phase 2:** $\mathscr{A}$ can continue to adaptively access the oracles of phase 1 (for keywords other than w_0 and w_1).

- **Response:** $\mathscr{A}$ outputs his guess $\mu^{'}$ and wins the game if and only if $\mu^{'} = \mu$.

We define $Adv_{\mathscr{A},\Pi}^{T-IND} := |Pr[\mu = \mu^{'}] - \frac{1}{2}|$.

Definition 2. *A searchable encryption scheme satisfies trapdoor indistinguishability against an adaptive chosen keyword attack if for any polynomial-time adversary $\mathscr{A}$, $Adv_{\mathscr{A},\Pi}^{T-IND}$ is negligible.*

In [28], Zhang et al. formally defined ciphertext indistinguishability and claimed that their proposed scheme achieves this requirement (Theorem 1 of [28])[2]. However, as for trapdoor indistinguishability, they did not present any formal security definition. They only mentioned it in the security analysis of their proposed scheme, where they claimed (without proof) that their scheme achieves trapdoor indistinguishability (Section 7.3 of [28]). In the following, we will show that Zhang et al.'s scheme fails to achieve these two essential security requirements.

[1] Note that additional oracles are due to the existence of the extra entity (i.e. administration server) in Zhang et al.'s scheme.
[2] Note that in [28], the authors use the term 'keyword semantic security' instead of the standard 'ciphertext indistinguishability' (see Definition 1 of [28]).

4 Cryptanalysis of Zhang et al.'s scheme

In [28], Zhang et al. claimed that their scheme provides ciphertext and trapdoor indistinguishability. In this section, we prove that their claim is not true. More precisely, we demonstrate that:

- The ciphertexts generated by the data owners (which are publicly sent to the administration server) are distinguishable by any attacker.
- The trapdoors generated by the data receivers (which are publicly sent to the administration server) are distinguishable by any attacker.
- The trapdoors generated by the administration server (which are publicly sent to the cloud server) are distinguishable by any attacker.

In the following, we provide three theorems to prove our claims.

Theorem 1. *The scheme of Zhang et al. does not satisfy ciphertext indistinguishability.*

proof: Let $\mathscr{A}$ be an adversary in $Game^{C-IND}_{\mathscr{A},\Pi}$, where Π is Zhang et al.'s scheme. To prove the theorem, we show that $\mathscr{A}$ can win the game with probability 1 and therefore, $Adv^{C-IND}_{\mathscr{A},\Pi}$ is not negligible.

Without using any oracles of query phase 1, $\mathscr{A}$ outputs two different keywords (w^*_0, w^*_1). Then, a random bit $\mu \in \{0,1\}$ is picked and the ciphertext $C'_{w^*_\mu}$ (along with $C_{w^*_\mu}$) is produced by running the Keyword Encryption algorithm. $C'_{w^*_\mu} = (E'_a, E_o)$ (along with $C_{w^*_\mu}$) is returned to $\mathscr{A}$. Then, $\mathscr{A}$ checks the following equation:

$$e(E'_a, g) \stackrel{?}{=} e(E_o, g^{H(w^*_0)})$$

If it holds, $\mathscr{A}$ outputs $\mu' = 0$ and otherwise, $\mu' = 1$. Note that by properties of bilinear map e, we have:

$$\begin{aligned} e(E_a{}', g) &= e(g^{k_i.r_o.H(w^*_\mu)}, g) \\ &= e(g,g)^{k_i.r_o.H(w^*_\mu)} \\ &= e(g^{k_i.r_o}, g^{H(w^*_\mu)}) = e(E_o, g^{H(w^*_\mu)}) \end{aligned}$$

Therefore, $\mu' = \mu$ with probability 1 and $\mathscr{A}$ wins the game (without knowing k_i or r_o). In other words, we showed that the ciphertexts generated by the data owners are distinguishable by any attacker. ■

Theorem 2. *The scheme of Zhang et al. does not satisfy trapdoor indistinguishability.*

proof: Let $\mathscr{A}$ be an adversary in $Game^{T-IND}_{\mathscr{A},\Pi}$ where Π is Zhang et al.'s scheme. To prove the theorem, we show that $\mathscr{A}$ can win the game with probability 1 and therefore, $Adv^{T-IND}_{\mathscr{A},\Pi}$ is not negligible.

Without using any oracles of query phase 1, $\mathscr{A}$ outputs two different keywords (w^*_0, w^*_1). Then, a random bit $\mu \in \{0,1\}$ is picked and the trapdoor $T'_{w^*_\mu}$ (along with $T_{w^*_\mu}$) is produced by running the Trapdoor Generation algorithm. $T'_{w^*_\mu} = (X, Y)$ (along with $T_{w^*_\mu}$) is returned to $\mathscr{A}$. Then, $\mathscr{A}$ checks the following equation:

$$e(X, g) \stackrel{?}{=} e(g^{H(w^*_0)}, Y)$$

If it holds, $\mathscr{A}$ outputs $\mu' = 0$ and otherwise, $\mu' = 1$. Note that by properties of bilinear map e, we have:

$$\begin{aligned} e(X, g) &= e(g^{H(w^*_\mu).r_u}, g) \\ &= e(g,g)^{H(w^*_\mu).r_u} \\ &= e(g^{H(w^*_\mu)}, g^{r_u}) = e(g^{H(w^*_\mu)}, Y) \end{aligned}$$

Therefore, $\mu^{'} = \mu$ with probability 1 and $\mathscr{A}$ wins the game (without knowing r_u). In other words, we showed that the trapdoors generated by the data receivers are distinguishable by any attacker. ■

Theorem 3. *The trapdoors generated by the administration server of Zhang et al.'s scheme are distinguishable and therefore, their scheme does not satisfy trapdoor indistinguishability.*

Proof: Let $\mathscr{A}$ be an adversary in $Game_{\Pi}^{T-IND}$ where Π is Zhang et al.'s scheme. We show that $\mathscr{A}$ can distinguish between the two trapdoors generated by the administration server and win the game with probability 1. Therefore, $Adv_{\mathscr{A},\Pi}^{T-IND}$ is not negligible and Π does not achieve trapdoor indistinguishability.

At first, $\mathscr{A}$ randomly selects r_u and computes $T_{w_2}^{'} = (g^{H(w_2).r_u}, g^{r_u})$ for an arbitrary chosen keyword w_2. Then, $\mathscr{A}$ requests for $\mathscr{O}^{AdminTrap}(T_{w_2}^{'})$. As a result, $\mathscr{A}$ receives:

$$T_{w_2} = (T_1, T_2, T_3) = (g^{H(w_2).r_u.k_{a_1}.k_{a_2}.r_a}, g^{r_u.k_{a_1}}, g^{r_u.k_{a_1}.r_a})$$

Now, $\mathscr{A}$ can simply calculate the value of $g^{k_{a_1}}$ by using $T_2 = g^{r_u.k_{a_1}}$ and r_u.

Afterwards, $\mathscr{A}$ sets $C_{w_3}^{'} = e(g^{k_{a_1}}, E_o)$ (for an unknown keyword w_3 and an arbitrary value E_o). Then, $\mathscr{A}$ requests for $\mathscr{O}^{AdminCipher}(C_{w_3}^{'})$. As a result, $\mathscr{A}$ receives $C_{w_3} = (E_a, E_o) = (g^{2k_{a_1}k_{a_2}}, E_o)$. Now, $\mathscr{A}$ can calculate the value of $g^{k_{a_1}k_{a_2}}$ by computing ${E_a}^{1/2}$.

At this time, $\mathscr{A}$ outputs two different keywords (w_0^*, w_1^*). Then, a random bit $\mu \in \{0,1\}$ is chosen and the trapdoor $T_{w_\mu^*}$ (along with $T_{w_\mu^*}^{'}$) is computed by running the Trapdoor Generation algorithm. Then, $T_{w_\mu^*} = (T_1, T_2, T_3)$ (along with $T_{w_\mu^*}^{'}$) is returned to $\mathscr{A}$. Now, $\mathscr{A}$ checks the following equation:

$$e(g^{k_{a_1}k_{a_2}}, T_3)^{H(w_0^*)} \stackrel{?}{=} e(g^{k_{a_1}}, T_1)$$

If it holds, $\mathscr{A}$ outputs $\mu^{'} = 0$ and otherwise, $\mu^{'} = 1$. Note that by properties of bilinear map e, we have:

$$\begin{aligned} e(g^{k_{a_1}k_{a_2}}, T_3)^{H(w_\mu^*)} &= e(g^{k_{a_1}k_{a_2}}, g^{k_{a_1}r_a^*r_u^*})^{H(w_\mu^*)} \\ &= e(g,g)^{k_{a_1}k_{a_2}r_u^*k_{a_1}r_a^*H(w_\mu^*)} \\ &= e(g^{k_{a_1}}, g^{k_{a_1}k_{a_2}r_u^*r_a^*H(w_\mu^*)}) = e(g^{k_{a_1}}, T_1) \end{aligned}$$

Therefore, $\mu^{'} = \mu$ and $\mathscr{A}$ wins the game (without knowing k_{a_1}, k_{a_2}, r_a^* or r_u^*) with a non-negligible advantage $Adv_{\mathscr{A},\Pi}^{C-IND}$. ■

5 Conclusion

This paper considers the security of the proposed scheme of Zhang et al. [28] where the authors target privacy preserving ranked multi-keyword search in a multi-owner model. They claimed that the scheme achieves ciphertext and trapdoor indistinguishability which informally states that an adversary should not be able to distinguish between ciphertexts and trapdoors corresponding to two different known keywords. Unfortunately, we found that their claim is not true and the scheme fails to guarantee these fundamental security requirements.

However, to be applicable in practice, these issues should somehow be addressed. One approach is to utilize the idea of using authenticated encryption with keyword search which has recently been applied in the literature [29, 30]. In such schemes, besides resolving the above mentioned concerns, by incorporating data sender's private key in creating searchable ciphertexts, no adversaries (including the server itself) can launch an important attack called keyword guessing attack. Therefore, the resulting scheme would be secure enough to be deployed in practice.

References

[1] Q. Huang and H. Li, "An efficient public-key searchable encryption scheme secure against inside keyword guessing attacks," *Information Sciences*, vol. 403–404, pp. 1–14, September 2017.

[2] H. S. Rhee, J. H. Park, W. Susilo, and D. H. Lee, "Trapdoor security in a searchable public-key encryption scheme with a designated tester," *Journal of Systems and Software*, vol. 83, no. 5, pp. 763–771, May 2010.

[3] L. Fang, W. Susilo, C. Ge, and J. Wang, "Public key encryption with keyword search secure against keyword guessing attacks without random oracle," *Information Sciences*, vol. 238, pp. 221–241, July 2013.

[4] N. Pakniat, "Public key encryption with keyword search and keyword guessing attack: a survey," in *Proc. of the 13th International Iranian Society of Cryptology Conference on Information Security and Cryptology (ISCISC'16), Tehran, Iran*. IEEE, September 2016, pp. 1–4.

[5] M. Noroozi, I. Karoubi, and Z. Eslami, "Designing a secure designated server identity-based encryption with keyword search scheme: still unsolved," *Annals of Telecommunications*, vol. 73, no. 11–12, pp. 769–776, December 2018.

[6] D. X. Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in *Proc. of the 2000 IEEE Symposium on Security and Privacy (SSP'00), Berkeley, Carifornia, USA*. IEEE, May 2000, pp. 44–55.

[7] R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky, "Searchable symmetric encryption: Improved definitions and efficient constructions," in *Proc. of the 13th ACM conference on Computer and communications security (CCS'06), Alexandria, Virginia, USA*. ACM, October-November 2006, pp. 79–88.

[8] W. TY, T. TT, and T. YM, "Efficient searchable id-based encryption with a designated server," *Annals of Telecommunations*, vol. 69, no. 7–8, pp. 391–402, August 2014.

[9] L. Guo and W.-C. Yau, "Efficient secure-channel free public key encryption with keyword search for EMRs in cloud storage," *Journal of Medical Systems*, vol. 39, no. 2, pp. 1–11, February 2015.

[10] G. Asharov, M. Naor, G. Segev, and I. Shahaf, "Searchable symmetric encryption: Optimal locality in linear space via two-dimensional balanced allocations," in *Proc. of the 2016 ACM forty-eighth annual ACM symposium on Theory of Computing (STOC'16), Cambridge, Morocco, USA*. ACM, June 2016, pp. 1101–1114.

[11] H. Li, F. Zhang, and C.-I. Fan, "Deniable searchable symmetric encryption," *Information Sciences*, vol. 402, pp. 233–243, September 2017.

[12] D. Boneh, G. D. Crescenzo, R. Ostrovsky, , and G. Persiano, "Public key encryption with keyword search," in *Proc. of the 2004 International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT'04), Interlaken, Switzerland*, ser. Lecture Notes in Computer Science, vol. 3027. Springer, May 2004, pp. 506–522.

[13] M. Kuzu, M. Islam, and M. Kantarcioglu, "Efficient similarity search over encrypted data," in *Proc. of the 28th IEEE International Conference on Data Engineering (ICDE'12), Washington, D.C., USA*. IEEE, April 2012, pp. 1156–1167.

[14] P. Xu, H. Jin, Q. Wu, and W. Wang, "Public-key encryption with fuzzy keyword search: A provably secure scheme under keyword guessing attack," *IEEE Transactions on Computers*, vol. 62, no. 11, pp. 2266–2277, September 2012.

[15] B. Zhang and F. Zhang, "An efficient public key encryption with conjunctive-subset keywords search," *Journal of Network and Computer Applications*, vol. 34, no. 1, pp. 262–267, January 2011.

[16] Z. Xia, X. Wang, X. Sun, and Q. Wang, "A secure and dynamic multi-keyword ranked search scheme over encrypted cloud data," *IEEE Transactions on Parallel and Distributed Systems*, vol. 27, no. 2, pp. 340–352, February 2015.

[17] C. Wang, N. Cao, J. Li, K. Ren, and W. Lou, "Secure ranked keyword search over encrypted cloud data," in *Proc. of the 30th IEEE International Conference on Distributed Computing Systems (ICDCS'10), Genova, Italy*. IEEE, June 2010, pp. 253–262.

[18] A. Boldyreva, N. Chenette, Y. Lee, and A. ONeill, "Order-preserving symmetric encryption," in *Proc. of the 28th International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT'09), Cologne, Germany*, ser. Lecture Notes in Computer Science, vol. 5479. Springer-Verlag, April 2009, pp. 224–241.

[19] C. Wang, N. Cao, K. Ren, and W. Lou, "Enabling secure and efficient ranked keyword search over outsourced cloud data," *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 8, pp. 1467–1479, December 2011.

[20] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy-preserving multi-keyword ranked search over encrypted cloud data," in *Proc. of the 2014 IEEE Transactions on Parallel and Distributed Systems (TPDS'14), Shanghai, China*. IEEE, November 2013, pp. 222–233.

[21] W. Sun, B. Wang, N. Cao, M. Li, W. Lou, Y. Hou, and H. Li, "Privacy-preserving multi-keyword text search in the cloud supporting similarity-based ranking," in *Proc. of the 8th ACM SIGSAC symposium on Information, computer and communications security (ASIA CCS'16), Hangzhou, China*. ACM, May 2013, pp. 71–82.

[22] C. Orencik, M. Kantarcioglu, and E. Savas, "A practical and secure multi-keyword search method over encrypted cloud data," in *Proc. of the 6th International Conference on Cloud Computing (ICCC'13), Santa Clara, California, USA*. IEEE, June–July 2013, pp. 390–397.

[23] Q. Xu, H. Shen, Y. Sang, and H. Tian, "Privacy-preserving ranked fuzzy keyword search over encrypted cloud data," in *Proc. of the 2013 IEEE International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT'13), Taipei, Taiwan*. IEEE, December 2013, pp. 239–245.

[24] J. Li, Q. Wang, C. Wang, N. Cao, K. Ren, and W. Lou, "Fuzzy keyword search over encrypted data in cloud computing," in *Proc. of the 2010 IEEE INFOCOM, San Diego, California, USA*. IEEE, March 2010, pp. 1–5.

[25] H. Li, D. Liu, Y. Dai, T. H. Luan, and X. S. Shen, "Enabling efficient multi-keyword ranked search over encrypted mobile cloud data through blind storage," *IEEE Transactions on Emerging Topics in Computing*, vol. 3, no. 1, pp. 127–138, December 2014.

[26] X. Jiang, J. Yu, J. Yan, and R. Hao, "Enabling efficient and verifiable multi-keyword ranked search over encrypted cloud data," *Information Sciences*, vol. 403–404, pp. 22–41, September 2017.

[27] W. Zhang, S. Xiao, Y. Lin, T. Zhou, and S. Zhou, "Secure ranked multi-keyword search for multiple data owners in cloud computing," in *Proc. of the 44th IEEE International Conference on Dependable Systems and Networks (DSN'14), Atlanta, Gabon, USA*. IEEE, June 2014, pp. 276–286.

[28] W. Zhang, Y. Lin, S. Xiao, J. Wu, and S. Zhou, "Privacy preserving ranked multi-keyword search for multiple data owners in cloud computing," *IEEE Transactions on Computers*, vol. 65, no. 5, pp. 1566–1577, June 2015.

[29] M. Noroozi and Z. Eslami, "Public key authenticated encryption with keyword search: revisited," *IET Information Security*, October 2018.

[30] H. Li, Q. Huang, J. Shen, G. Yang, and W. Susilo, "Designated-server identity-based authenticated encryption with keyword search for encrypted emails," *Information Sciences*, vol. 481, pp. 330–343, May 2019.

Author Biography

Mahnaz Noroozi received her B.S. degree in Computer Sciences in 2010 from Sharif University of Technology, Tehran, Iran. In 2012, she received her M.S. degree in Computer Science from Shahid Beheshti University, Tehran, Iran. She is currently a Ph.D. candidate in Computer Science at Shahid Beheshti University, Tehran, Iran. Her research interests include cryptographic protocols and their security.

Ziba Eslami received her B.S., M.S., and Ph.D. in Applied Mathematics from Tehran University in Iran. From 1991 to 2000, she was a resident researcher in the Institute for Studies in Theoretical Physics and Mathematics (IPM), Iran. During the academic years 2000-2003, she was a Post Doctoral Fellow in IPM. She served as a non-resident researcher at IPM during 2003-2005. Currently, she is an associate professor at Shahid Beheshti University in Iran. Her research interests include design theory, combinatorial algorithms, cryptographic protocols, and steganography.

Joonsang Baek received the Ph.D. degree from Monash University, Australia, in 2004. His Ph.D. thesis was on security analysis of signcryption, and has received great attention from the research community. He was a Research Scientist in the Institute for Infocomm Research, Singapore, and an Assistant Professor in the Khalifa University of Science and Technology, United Arab Emirates. He is currently a Senior Lecturer in the School of Computer Science and Information Technology, University of Wollongong, Australia. He has published his work in numerous reputable journals and conference proceedings. His current research interests are in the field of applied cryptography and cybersecurity. He has also served as a Program Committee Member and the Chair for a number of renowned conferences on information security and cryptography.