An Improved Lindell-Waisbard Private Web Search Scheme

Zhengjun Cao[1], Lihua Liu[2], Zhenzhen Yan[1]
(Corresponding author: Zhengjun Cao)

Department of Mathematics, Shanghai University[1]
No. 99, Shangda Road, Shanghai, China.
Department of Mathematics, Shanghai Maritime University[2]
No. 1550, Haigang Ave, Pudong New District, Shanghai, China.
(Email: caozhj@shu.edu.cn)
(Received May 27, 2015; revised and accepted Aug. 10 & Sept. 6, 2015)

Abstract

In 2010, Lindell and Waisbard proposed a private web search scheme for malicious adversaries. At the end of the scheme, each party obtains one search word and queries the search engine with the word. We remark that a malicious party could query the search engine with a fake word instead of the word obtained. The malicious party can link the true word to its provider if the victim publicly complain for the false searching result. To fix this drawback, each party has to broadcast all shares so as to enable every party to recover all search words and query the search engine with all these words. We also remark that, from a user's perspective, there is a very simple method to achieve the same purpose of private shuffle. When a user wants to privately query the search engine with a word, he can pick another $n-1$ padding words to form a group of n words and permute these words randomly. Finally, he queries the search engine with all these words.

Keywords: ElGamal encryption, private shuffle, private web search

1 Introduction

As we see private web search (PWS) has become a serious problem. There are several tricks to deal with it. The anonymous routing system [6] can be used though it is somewhat inefficient. So do the private information retrieval [4, 21] and mix-net [3, 5, 11]. In 2009, Castellà-Roca et al. [2] suggested a new approach for the problem.

Their proposal is for a group of users to shuffle their search words amongst themselves. After the shuffle, each user has someone's search word (but doesn't know whose), and each party then query the search engine with the word obtained. Finally, the parties all broadcast the result. Their private shuffle protocol is secure only in the presence of semi-honest adversaries.

At PETS'2010, Lindell and Waisbard [17] pointed out that the scheme suggested by [2] is unrealistic because it is vulnerable to many attacks. They proposed a private shuffle protocol for malicious adversaries and proved its security according to their security definition. They also addressed some practical considerations. But we would like to stress that at the end of the Lindell-Waisbard scheme, like the previous work [2], each party obtains only one search word and query the search engine with the word.

In this paper, we remark that in the Lindell-Waisbard private web search scheme a malicious party could query the search engine with a fake word instead of the word obtained. Thus the party corresponding to the true word cannot obtain the proper searching result.

More worse, the malicious party can link the true word to its provider if the victim publicly complain for the false searching result. However, the victim himself can not find who is the malicious party. To fix this drawback, each party has to broadcast all shares so as to enable every party to recover all search words and query the search engine with all these words. We also remark that from a user's perspective there is a very simple method to achieve the same purpose of private shuffle. Besides, we shall correct some misunderstandings about "denial of service" and "malicious attacks in cryptography".

The primitive of mix network is introduced by Chaum [3], which can be used for e-voting, e-auction and private web search. Loosely speaking, a mix network shuffles a number of inputting ciphertexts (each from one user) to the same number of outputting plaintexts such that: 1) the outputs are a permutation of the plaintexts of the inputs; 2) the permutation between the inputs and the outputs is unknown so that the users cannot be linked to their outputs.

Since then, researchers have put forth many proposals on mix network [1, 3, 5, 9, 10, 11, 20] and its applications [8, 15, 22, 24, 25, 26]. In recent, Juarez and Torra [12, 23] have studied the technique of dissociating privacy agent, which is a browser extension that acts as a proxy between the user and the search engine and semantically dissociates queries on real time. Romero-Tris et al. [23] have pointed out the differences between single-party PWS model and multi-party PWS model.

The anonymous routing system [6, 16, 18, 19] can be also used for private web search but it is somewhat inefficient for multi-party PWS. Recently, Li and Hwang [13, 14] have designed a lightweight anonymous routing protocol without public key en/decryptions for wireless ad hoc networks.

2 Review of the Lindell-Waisbard Private Web Search Scheme

A shuffle functionality is a probabilistic function $f(x_1, \cdots, x_n) = (y_1, \cdots, y_n)$, such that for every $i, y_i = x_{\pi(i)}$ where π is a random permutation over $\{1, 2, \cdots, n\}$. A shuffle is private if an adversary cannot link between the inputs and the outputs of the protocol.

Assume that all parties hold a unique session identifier sid (e.g., this could be a timestamp). There is a group $\mathbb{G}$ of order q with generator g, to be used for the ElGamal encryption [7]. Let (E, D) denote a CCA2-secure public-key encryption scheme. At the beginning of the scheme, each party P_j has a search word w_j, $j = 1, \cdots, n$. At the end of the scheme, each P_j obtains an arbitrary search word $w'_j \in \{w_1, \cdots, w_n\}$. The Lindell-Waisbard private web search scheme [17] can be described as follows.

Initialization stage.

1) Each party P_j chooses a random $\alpha_j \in \mathbb{Z}_q^*$, computes $h_j = g^{\alpha_j}$, and chooses a pair of keys (sk_j, pk_j) for the CCA2-secure encryption. P_j sends (h_j, pk_j) to all the other parties and proves knowledge of α_j. P_j signs the message and sends together with the identifier sid using its certified private signing key.
2) Each party verifies the signatures on the messages and the proofs that it received and aborts unless all are correct.
3) Each party P_j encrypts its input w_j using ElGamal with the public key $h = \prod_{i=1}^{n} h_i$. That is, it chooses a random $\rho_j \in \mathbb{Z}_q^*$ and computes an encryption
$$m_j = (g^{\rho_j}, h^{\rho_j} w_j).$$
4) Each party P_j computes
$$c_j = E_{pk_1}(E_{pk_2}(\cdots(E_{pk_n}(m_j))\cdots))$$
and sends c_j to P_1.

The output of this phase is the list of the encrypted c_j's, denoted by $\mu_0 = \langle c_1^0, \cdots, c_n^0 \rangle$.

Shuffle stage.

For $j = 1, \cdots, n$, P_j receives μ_{j-1} and computes μ_j as follows:

1) P_j checks that there are no duplications in μ_{j-1}. If there are, it aborts.
2) P_j decrypts every c_i^{j-1} in μ_{j-1} by computing
$$c_i^j = D_{sk_j}(c_i^{j-1}).$$
3) P_j randomly permutes the list of values c_i^j computed above. The result is denoted by μ_j.
4) P_j sends μ_j to P_{j+1}. The last party P_n sends μ_n to all parties.

Verification stage.

1) Every party P_j checks its ElGamal ciphertext m_j appears in the vector μ_n. If yes it sends $(sid, P_j, \mathbf{true})$, signed with its private signing key, to all the other users. Otherwise it sends $(P_j, \mathbf{false})$.
2) If P_j sent false in the previous step, or did not receive a validly signed message $(sid, P_i, \mathbf{true})$ from all other parties P_i, then it aborts. Otherwise, it proceeds to the next step.

Reveal stage.

1) For every (u_i, v_i) in μ_n, P_j computes $s_i^j = u_i^{\alpha_j}$ and sends s_i^j to P_i.
2) Every party P_j computes
$$w'_j = \frac{v_j}{\prod_{k=1}^{n} s_j^k},$$
thereby decrypting the ElGamal ciphertext and recovering the search word w'_j (here j denotes the current index in μ_n and not the index of the party who had input w_j at the beginning of the protocol).

Query stage.

After the above shuffle, each party has someone's search word, and the parties then query the search engine with the word obtained. Finally, the parties all broadcast the result to all others.

We refer to Figure 1 for the basic idea behind the Lindell-Waisbard private web search scheme.

3 An Attack Launched by any Malicious Party in Query Stage

The Lindell-Waisbard private web search scheme is builded on the previous work [2]. They claim that the protocol is secure in the presence of malicious adversaries. We now remark that the scheme is vulnerable to an attack launched by any malicious party.

Table 1: Difference between the Lindell-Waisbard scheme and the modification

	The Lindell-Waisbard scheme	The modification
Reveal	For every (u_i, v_i) in μ_n, P_j computes $s_i^j = u_i^{\alpha_j}$ and sends s_i^j to P_i. Every party P_j computes w'_j.	For every (u_i, v_i) in μ_n, P_j computes $s_i^j = u_i^{\alpha_j}$ and broadcasts s_i^j and *the proof of* α_j to all others. Every party P_j checks the proofs and computes $w'_1, \cdots, w'_n$.
Query	Each party P_j queries the search engine with w'_j, and broadcasts the searching result.	Each party P_j queries the search engine with $w'_1, \cdots, w'_n$.

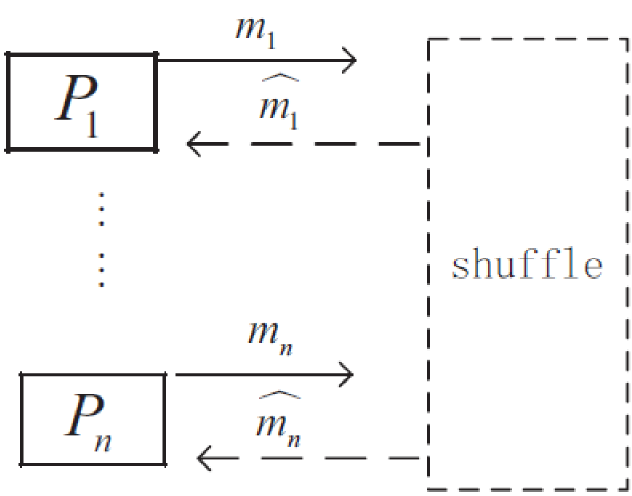

Figure 1: The Lindell-Waisbard shuffle

Suppose that P_k is a malicious party and the others are semi-honest. At the end of Reveal stage, P_k obtains a word w'_k which is in the set $\{w_1, \cdots, w_n\}$. In Query stage, P_k can query the search engine with an arbitrary word $\widehat{w_k}$ such that $\widehat{w_k} \neq w'_k$.

He broadcasts the searching result corresponding to the word $\widehat{w_k}$. Since the probability that $\widehat{w_k} \in \{w_1, \cdots, w_n\}$ is negligible, the party corresponding to the word w'_k shall not obtain the proper searching result. More worse, if the victim publicly complains for the false searching result, then P_k can link the true word w'_k to the victim. Note that the victim himself can not find who is the malicious party.

4 A Modification of the Lindell-Waisbard Scheme

The Lindell-Waisbard PWS scheme requires many broadcast channels. For example, each party P_j has to broadcast (h_j, pk_j) in Initialization stage, $(sid, P_j, \textbf{true})$ or $(P_j, \textbf{false})$ in Verification stage, and the searching result in Query stage. In view of that each party can access to these broadcast channels, in Reveal stage for every (u_i, v_i) in μ_n each party P_j can broadcast s_i^j to all others, instead of sending it to P_i in the mode of point-to-point. Hence, every party can recover all search words $w'_1, \cdots, w'_n$. Finally, every party can query the search engine with all these search words. See the following Table 1 for the differences between the original Lindell-Waisbard scheme and its modification.

The modification requires that P_j broadcasts the zero-knowledge proof of α_j with respect to u_i. The requirement cannot be removed. Otherwise, there exists a similar attack launched by any malicious party. Suppose that P_j is the malicious party and the others are semi-honest. In Reveal stage, P_j broadcasts $\widehat{s_i^j}$ such that $\widehat{s_i^j} \neq u_i^{\alpha_j}$ for some index i. Hence, the others shall obtain $w'_1, \cdots, \widehat{w'_i}, \cdots, w'_n$. If the party corresponding to w'_i complains for the false word $\widehat{w'_i}$, P_j can link the true word w'_i to the victim. However, the victim himself can not find who is the malicious party.

We refer to Figure 1 and Figure 2 for the essential differences between the original Lindell-Waisbard scheme and the modification.

5 A Simple Method for Single-party Private Web Search

The essence of a private shuffle protocol is to mix a user's search word with another $n-1$ words such that an adversary cannot know which is the user's true search word. In fact, from a user's perspective there is a very simple method to achieve the same purpose. Concretely, when a user wants to privately query the search engine with a word, he first chooses $n-1$ padding words to form a group of n words and then permutes these words. Finally, he queries the search engine with these words. Figure 3 is a simple private web search method.

It is easy to see that the simple scheme is secure be-

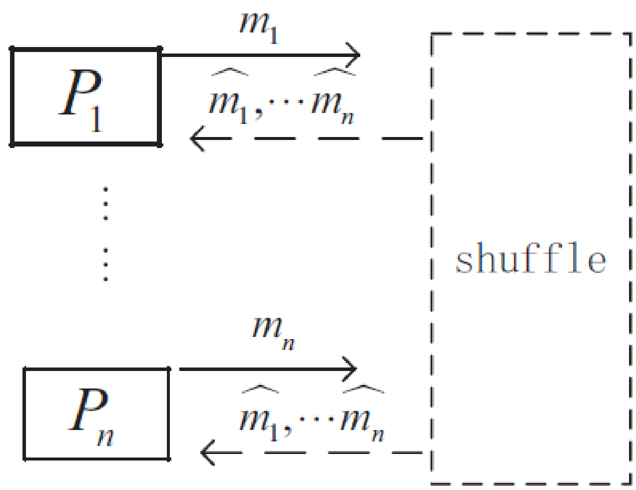

Figure 2: The modified Lindell-Waisbard shuffle

cause the adversary can know the true word with the probability of $1/n$. In comparison with the modified Lindell-Waisbard scheme, the simple method requires relatively little cost.

6 Further Discussions

We have received some comments on the manuscript. Somebody argues that

> *The attack proposed could be viewed as a type of denial of service where a malicious party always complains in the protocol, causing the whole session to abort. The last simple fix does not work because one will know all the words are from the same user, and as long as one of the words is sensitive, it is linked to that user. The correctness guarantee is not required for the Lindell-Waisbard scheme, and as such malicious parties are allowed to perform denial of service type attacks (the attack mentioned above is one such attack).*

We now want to point out that:

- Their security model has actually considered replacement attacks but they did not pay attentions to the proposed malicious attack in the paper. It points out in the introduction that [17]: "we still have to deal with 'replacement attacks' where the first party carrying out the mix replaces all of the encrypted search words with terms of its own, except for the one ciphertext belonging to the user under attack."

- In the Lindell-Waisbard scheme, it is very likely to happen that a malicious adversary changes the search word from others when submitting it to a search engine. This is because: 1) his malicious behavior cannot be detected by others so that he does not undertake any obligations; 2) the false searching result broadcasted could tempt the victim to complain.

- The last simple fix is helpful to explain the essence of Lindell-Waisbard Scheme. From each user's point of view, the Lindell-Waisbard shuffle scheme is just mixing his searching word with other $n-1$ words submitted by other $n-1$ users. We do not consider whether an adversary can find a "sensitive" word among these n words. Actually, it is difficult to define the term of "sensitive" in the scenario.

- The replacement attack cannot be falsely regarded as a type of "denial of service", because it takes place just at the end of the whole session. In such case, users can obtain proper searching results, except for the victim.

7 Conclusion

We show that there is a drawback in the Lindell-Waisbard private web search scheme. We also remark that from a user's perspective there is a very simple method to achieve the same purpose of the Lindell-Waisbard scheme. This paper, we think, is helpful to explain the gist of Lindell-Waisbard private shuffle and correct some misunderstandings about "denial of service" and "malicious attacks in cryptography".

Acknowledgments

This work is supported by the National Natural Science Foundation of China (Project 61303200, 61411146001). The authors gratefully acknowledge the reviewers for their valuable suggestions.

References

[1] M. Abe, "Mix-networks on permutation net-works," in *Proceedings of Advances in Cryptology (ASIACRYPT'98)*, pp. 258–273, Beijing, China, Oct. 1998.

[2] J. Castellà-Roca, A. Viejo, and J. Herrera-Joancomarti, "Preserving user's privacy in web search engines," *Computer Communications*, vol. 32, no. 13-14, pp. 1541–1551, 2009.

[3] D. Chaum, "Untraceable electronic mail, return addresses, and digital pseudonyms," *Communications of the ACM*, vol. 24, no. 2, pp. 84–88, 1981.

[4] B. Chor, O. Goldreich, E. Kushilevitz, and M. Sudan, "Private information retrieval," *Journal of the ACM*, vol. 45, no. 6, pp. 965–981, 1998.

[5] Y. Desmedt and K. Kurosawa, "How to break a practical mix and design a new one," in *Proceedings of Advances in Cryptology (EUROCRYPT'00)*, pp. 557–572, Bruges, Belgium, May 2000.

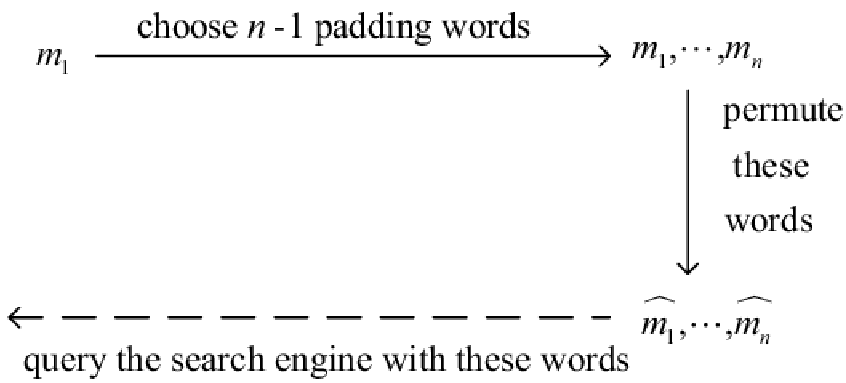

Figure 3: A simple private web search method

[6] R. Dingledine, N. Mathewson, and P. Syverson, "Tor: The second-generation onion router," in *Proceedings of the 13th USENIX Security Symposium*, pp. 303–320, San Diego, CA, USA, Aug. 2004.

[7] T. ElGamal, "A public key cryptosystem and a signature scheme based on discrete logarithms," in *Proceedings of Advances in Cryptology (CRYPTO'84)*, pp. 10–18, Santa Barbara, California, USA, Aug. 1984.

[8] T. H. Feng, W. T. Li, and M. S. Hwang, "A false data report filtering scheme in wireless sensor networks: A survey," *International Journal of Network Security*, vol. 17, no. 3, pp. 229–236, 2015.

[9] J. Furukawa and K. Sako, "An efficient scheme for proving a shuffle," in *Proceedings of Advances in Cryptology (CRYPTO'01)*, pp. 368–387, Santa Barbara, California, USA, Aug. 2001.

[10] J. Groth, "A verifiable secret shuffle of homomorphic encryptions," *Journal of Cryptology*, vol. 23, no. 4, pp. 546–579, 2010.

[11] M. Jakobsson, "A practical mix," in *Proceedings of Advances in Cryptology (EUROCRYPT'98)*, pp. 448–461, Espoo, Finland, June 1998.

[12] M. Juarez and V. Torra, "Dispa: An intelligent agent for private web search," *Studies in Computational Intelligence*, vol. 567, pp. 389–405, 2015.

[13] C. T. Li and M. S. Hwang, "A lightweight anonymous routing protocol without public key en/decryptions for wireless ad hoc networks," *Information Sciences*, no. 181, pp. 5333–5347, 2011.

[14] C. T. Li, C. C. Yang, and M. S. Hwang, "A secure routing protocol with node selfishness resistance in manets," *International Journal of Mobile Communications*, no. 10, pp. 103–118, 2012.

[15] W. T. Li, T. H. Feng, and M. S. Hwang, "Distributed detecting node replication attacks in wireless sensor networks: A survey," *International Journal of Network Security*, vol. 16, no. 5, pp. 323–330, 2014.

[16] X. D. Lin and et al., "Asrpake: An anonymous secure routing protocol with authenticated key exchange for wireless ad hoc networks," in *Proceedings of IEEE International Conference on Communications (ICC'07)*, pp. 5333–5347, Glasgow, Scotland, June 2007.

[17] Y. Lindell and E. Waisbard, "Private web search with malicious adversaries," in *Proceedings of Privacy Enhancing Technologies, 10th International Symposium (PETS'10)*, pp. 220–235, Berlin, Germany, July 2010.

[18] R. X. Lu, Z. F. Cao, L. C. Wang, and C. K. Sun, "A secure anonymous routing protocol with authenticated key exchange for ad hoc networks," *Computer Standards and Interfaces*, vol. 29, no. 5, pp. 521–527, 2007.

[19] R. Matam and S. Tripathy, "Provably secure routing protocol for wireless mesh networks," *International Journal of Network Security*, vol. 16, no. 3, pp. 182–192, 2014.

[20] C. Neff, "A verifiable secret shuffle and its application to e-voting," in *Proceedings of ACM Conference on Computer and Communications Security 2001*, pp. 116–125, Philadelphia, USA, Nov. 2001.

[21] R. Ostrovsky and W. Skeith, "A survey of single-database pir: Techniques and applications," in *Proceedings of 10th International Conference on Practice and Theory in Public-Key Cryptography (PKC'07)*, pp. 393–411, Beijing, China, Apr. 2007.

[22] S. Patel, D. Punjani, and D. Jinwala, "An efficient approach for privacy preserving distributed clustering in semi-honest model using elliptic curve cryptography," *International Journal of Network Security*, vol. 17, no. 3, pp. 328–339, 2015.

[23] C. Romero-Tris, A. Viejo, and J. Castella-Roca, "Multi-party methods for privacy-preserving web search: Survey and contributions," *Studies in Computational Intelligence*, vol. 567, pp. 367–387, 2015.

[24] S. Sarpong, C. X. Xu, and X. J. Zhang, "An authenticated privacy-preseving attribute matchmakng protocol for mobile social networks," *International Journal of Network Security*, vol. 17, no. 3, pp. 357–364, 2015.

[25] F. Wang, C. C. Chang, and Y. C. Chou, "Group authentication and group key distribution for ad hoc

networks," *International Journal of Network Security*, vol. 17, no. 2, pp. 199–207, 2015.

[26] T. J. Wei, "Communication efficient shuffle for mental poker protocols," *Information Sciences*, vol. 181, pp. 5053–5066, 2011.

Zhengjun Cao is an associate professor of department of Mathematics at Shanghai University. He received his Ph.D. degree in applied mathematics from Academy of Mathematics and Systems Science, Chinese Academy of Sciences. He served as a post-doctor in Department of Computer Science, University Libre de Bruxelles, from 2008 to 2010. His research interests include cryptography, discrete logarithms and quantum computation.

Lihua Liu is an associate professor of department of Mathematics at Shanghai Maritime University. She received her Ph.D. degree in applied mathematics from Shanghai Jiao Tong University. Her research interests include combinatorics, cryptography and information security.

Zhenzhen Yan is currently pursuing her M.S. degree from Department of Mathematics, Shanghai University. Her research interests include information security and cryptography.