IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 3 Issue 3, March 2016.
www.ijiset.com
ISSN 2348 – 7968

Generating Content Searchable Cipher Texts for Filtering the Spam

Nale Supriya[1], Khomane Sharmila[2], Yadav Ashwini[3], Pondkule Priyanka[4], Prof. Kadam P.R.[5]

[1 - 4] Student, Department of Computer Engineering, S.B. Patil College of Engineering, Maharashtra, India.

[5] Asst. Prof., Department of Computer Engineering, S.B. Patil College of Engineering, Maharashtra, India.

Abstract

Day by day securing communication over network problems are increased like unauthorized access, checking spam mails, access permissions. Whenever in the network communication is there, in this the sender sends file and server or firewall (gateway) may check contents of the file for security purpose. Public key encryption technique provides better solution in this case by creating searchable cipher texts for boosting the process of search. In the technique of searchable symmetric encryption sender can upload file to receiver's server in encrypted format with extracted keyword in encrypted format only and receiver can download file by searching the keyword. In this technique the receiver doesn't want to give permission to the server to decrypt the files. In this paper we are giving a new technique in which sender can upload files to the server with encrypted content and receiver is able to search by giving content searchable cipher texts. In this technique receiver is also able to report spam mails to the server after observing the contents of the file. As in previous techniques searching from large database is lengthy process but in this technique we are keeping the search time linear with the total number of the contents.

While implementing the creating content searchable cipher texts for reporting spam mails we are using Searchable Public-Key Cipher texts with generating Hidden Structures (SPCHS) algorithm in which semantic security of content doesn't reveal and we are creating different hidden star like structure for keeping track of same content and fast searching of content. Our concept is receiver is checks the file it may contains malicious things and it may harm the system in this case receivers system is under risk so he will report such files as spam files and server will have ability to block such files coming to the system. Finally we propose content searchable cipher texts generation for checking spam mails.

Keywords: *Hidden star structure, content searchable cipher texts.*

1. Introduction

Information security plays vital role in the development of the internet. For protecting the information one of the best techniques is public key encryption technique. The growing use of mobile phones, tabs and other wireless devices allow users to access their mails, data and files from any location.

For sake convenience user may store their data on to their remote servers rather on their machine and for the security and confidentiality purpose they are storing the data in the encrypted format. But it makes difficult to retrieve encrypted documents from server. Suppose an example if server is having bunch of mails for one receiver then it difficult for server to such data with specific keyword or subject in case of mails. In this situation public key encryption with keyword search works.

1.1 Public Key Encryption for Keyword Search

Among all security technic public key encryption is most effective one. Users are using different handheld devices to check their emails or data and according to keywords given in the mail server has to forward the mail to appropriate devices. With the email we except small numbers of keywords or mail id itself can work as keyword. This type of mobile email facility is provided in [1] and currently everybody using it also.

Now suppose sender sends the keyword and email in encrypted format by using receiver's public key to the server then server will face problems while routing the messages. In his privacy violation will happen in case of mobile project. PEKS's[ ] goal is that providing ability to the server to check the keyword in the email without knowing the rest of the things contained in the email. In this case receiver can give some keywords to search to the server and nothing about the rest of mail details. Let us discuss this system in detail. First sender sends message in encrypted format and then appends it with public key encryption with keyword search for each keyword. Encrypting message M with keywords K1,K2,….,Kn with receivers R public key then,

$$E_{Rpu}(M) \parallel PEKS(R_{pu},K1),\ldots\ldots,PEKS(R_{pu},Kn)$$

In this encryption receiver will give some sort of trapdoor Tw to the server with ability to check for certain keyword of his own choice is present or not. With the given $PEKS(R_{pu},K^{/})$ and the trapdoor Tw server can check the $K=K^{/}$. For this whole process there will be no any sort

ISSN 2348 – 7968

of communication between the sender and receiver. If $K \neq K^{/}$ then server will not learn anything about rest of message. In such way sender create only keyword searchable encrypted keywords with receiver's public key. In whole this process the search time is depends on the total numbers of cipher texts.to

1.2 Creating Hidden Relations

For getting higher results in the search performance without sacrificing the semantic security of data in the PEKS we can use hidden tree like structure provided in [2]. Usually keyword space without high minentropy. Semantic security is very important for keeping privacy of keywords in such type of applications. In previous schemes sever has to check each and every cipher text and requires more time for searching as every cipher text is alike to the server. To overcome from this in PEKS without sacrificing semantic security can organize these cipher texts in hidden tree like structures. With these hidden relations the search over the given cipher texts becomes fast.

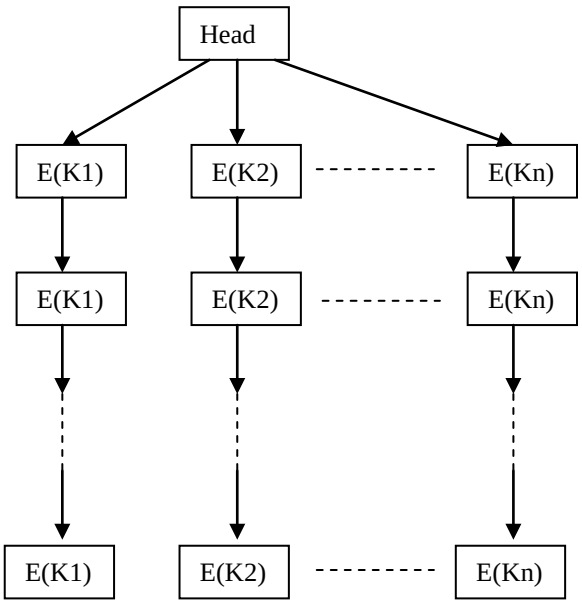

Fig -1: Hidden tree like structure for searchable cipher texts.

Especially from Fig-1 in this technique keywords forms a branches of tree by corresponding hidden relations and there is also link present between head and first cipher text. With the help of trapdoor and the head the server will find the matching keyword from the relation. Moving forward for the next keyword for search will be disclosed by the previously found word and will guide all searching process for finding next keyword. And for this technique it is clear that the search time is linear with the actual numbers of the cipher texts in the chain from the queried keyword.

For keeping the proper semantic security this tree like structure is not disclosing other relations to the server. Each sender of data is having ability to generate the keyword searchable cipher texts with these hidden relations with the receiver's public key and the trapdoor should keep the partial relations disclosed. With this we can keep the semantic security in the Public key encryption with keyword search.

1.3 Our work

We are working formally for the public key encryption with content search by keeping semantic security with the help of hidden relations to report spam mails. If any mail coming to the system contains the spam mail so filtering this spam mails is also equally important thing. In the previous techniques work focus on the keyword search we want to keep our track for the content search.in this new concept content searchable cipher texts are generated by the sender by using the receiver's public key and with the help of content searchable cipher texts trapdoor another partial relations can easily get find to guide further search. In this semantic security is for both content and hidden relations are defined.

Suppose for sender sends the message M with appending the encrypted content C(C1,C2….,Cn) with receiver's public key pu. We are calling it as public key encryption with content search.

$$E_{Rpu}(M) \,||\, PECS(R_{pu}, C1), \ldots\ldots, PEKS(R_{pu}, Cn)$$

The next step is to create the content searchable hidden relations like in the PEKS. This new concept receiver will give content searchable cipher texts to the trapdoor and if mail contains any encrypted malicious contains then can filter the spams mails. No doubt this concept will keep the semantic security of the content searchable cipher texts and any hidden structures.

The formation of the PECS with semantic security we have to keep secure ids of receiver. In the Identity-Based Key Encapsulation Mechanism (IBKEM) the sender is encapsulating the key with the ID of receiver and buy of course the this ID can be decapsulate to obtain actual key and now sender gets the key K. But if any third party comes with $ID^{/}$and tries to obtain this $K^{/}$. There are two views for this, 1. Receivers side is unknown about the K

ISSN 2348 – 7968

and $K^{/}$. 2. In the PBKEM technique the sender is having information about the $ID^{/}$ and $K^{/}$.

2. Related Work

The work for Symmetric key encryption technique with the keyword search is given by many researchers in the [4] and their refinements are given by [5], [6] with keeping search time linear with all cipher texts in the database. The next refinement by Curtmola et al.in [3] proposed the technique in which multiuser can send the queries to the server for searching the keyword and also proved that the semantic security with logarithmic search time. Bao F., Deng R. H., Ding X., Yang Y [7]. Their research discovers that in the current system content encryption is subject to the user privacy attacks. Then they proposed new technique in which we can privately broadcast encryption to protecting the user privacy. They developed the broadcast scheme in private manner with active attacker's privacy protection guarantee.

The next technique in [8] by Waters B. R built the audit logs in the encrypted form that can searchable. In the technique developed by the Chase et al. [9] proposed technique to encrypt the structured data and secure searching method on that data.

In the PEKS extension work Abdalla M [10] proposed a new methodology works for construction of verifiable functions with randomization. They are deriving these functions from identity based key encapsulation mechanism. In the next research of Park D. J., Kim K., Lee P. J [11] works on the public key encryption with keyword search and proposed a new technique of searching keyword with conjunction without leaking information to the server. The same sort of work with some extension is given in [12], [13]. The same technique with time search proposed in [14], [10], authority search [15], [16] and search based on subset in [17]. The similarity search is proposed by Cheung D. W. [18] is based on fuzzy identity based encryption. The proposed technique in Arriaga et al. [19] again works for the fuzzy keyword search by using trapdoor. They are providing the same fuzzy keyword search trapdoor and can be shared by two or more keyword and in this technique the search time of keyword is keeping liner with the number of cipher texts.

In the technique proposed in Chase M [9] considered the problem of encrypting structured data so that the query searching process is effective and efficient. They are also providing the chain like structure creation for fast searching. Extension to this work is provided in [20] works on protection of data which is being searched but the chain structure given in the paper is not totally hidden from the server and my get catches. Bellare et al. [21] proposed a technique for efficient keyword search with deterministic public key encryption (PKE) by formalizing strongest security mechanism. Same technique extended by Brakerski Z in [23] with some security enrichments but still facing problems of semantic security.

Peng Xu [2] proposed new methodology of PEKS with semantic security with creating hidden tress for fast searching of keywords. With this technique they are constructing the new algorithm based on the random oracle model.

2.1 Limitations with Existing Systems with Solution

In the existing technique the work has been done on the Keyword search with keeping semantic security. Work has been done on the fast searching of keyword by creating chain like structure also. But there is scope for implementing public key search for content by generating cipher texts. Such type of content searching is useful for filtering spam mails and preventing unauthorized access. With this scope in this paper we are proposing architecture for the generating content searchable cipher texts for filtering the spam. While working on this we are working with the same technique used in the [2] for content searching.

3. Proposed System

Public encryption technique with the cipher text works with the content search the architecture is shown in the Fig-2.In this architecture it is more suitable to work with multiple senders in the system. In this scenario work flow is as flows,

1. The various senders will produce the encrypted file by using the receiver's public key and upload this file to the server.

E_{Rpu} (Message)

ISSN 2348 – 7968

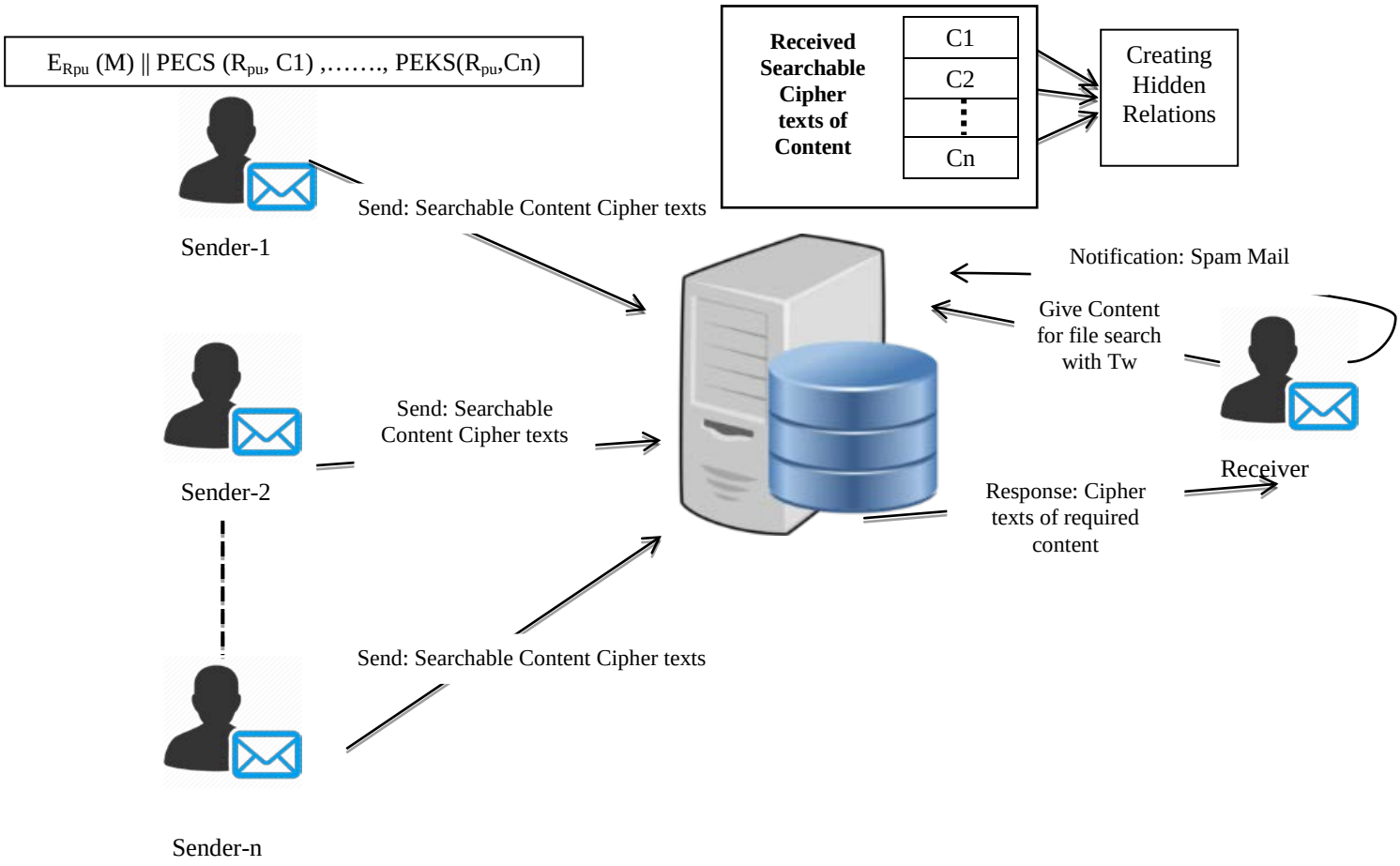

Fig -2: System Architecture

2. The Senders will also send some sequential keywords or content like with encryption to the server C1, C2…., Cn) for example any sequential keywords for example "Urgent Basis Requirements For Post" and this whole will not reveal anything to the server.

PECS (R_{pu}, C1) ,……., PEKS(R_{pu},Cn)

3. The receiver at other side will send a simple secrete key Tw to the server and this will give ability to the server to locate all messages with specific content but nothing else. Receiver will create this Trapdoor with his private key and server simply sends relevant emails according to search.

4. In next step the server is creating the hidden relations according to the content. Server will produce the one tree like structure for fast processing of contents.

5. If in the some cases the sender's security key gets revealed then forward security concept get applied over here. In this cipher texts on hidden relations will be confidential and server keeps it track for creating new cipher texts only.

6. With the backward security technique sender asks for the new structure initialization by using initialization algorithm and this structure is totally independent of the previous structure.

7. Next server will provide the generated cipher texts to the receiver according to his search and the searching time is liner with the actual size of content which is queried.

8. Receiver will check the content and if he/she thinks that the content of the mail is spam then the same go towards the server. Receiver will notify the server for spam mails and not to intenerating same mails in future. Server will keep track of such messages give restriction to the sender. In this whole process there will be no any

ISSN 2348 – 7968

interaction between the sender and receiver is there. Search time is linear with the queried content by the receiver.

Construction is as flows for the algorithm as follows. Given PECS has following setups keyGen, PECS, Trapdoor (Tw), Test

1. Setup: In this phase keyGen algorithm will generate

Rpu = Rpr. The master key will be Rpr.

2. PECS: first compute PECS (Rpu, C1) and output will be

PECS (Rpu, C1),……., PEKS(Rpu,Cn)

3. Trapdoor (Rpu, C1): Output TW for searching the keyword.

4. Test (Rpu, C1, Tw): let C1 = (A;B).

Test if H2(E(Tw,A)) = B. if this valid then Yes otherwise no.

4. Conclusions

This paper works on the public key encryption with the content search with semantic security provision. This technique allows creating content searchable cipher text with generating hidden tree like structure. The trapdoor with another side provides the facility to disclose another part of the hidden relations to guide the searching process. The PECH generates content searchable cipher texts with hidden relations. The search complexity is linear with the queried content can be maintained with this architecture. Such type of work is useful for the filtering the spam mail and authorizing the access. Thought this process ant thing will not get reveal to the server.

In the future scope one can work on the retrieval of complete verification with the PECH scheme by forming the hidden relations.

Acknowledgments

We thank colleagues of S.B. Patil College of Engineering Pune, for their constant feedback on the paper. We also thank the anonymous reviewers for their valuable comments.

References

[1] Boneh D., Crescenzo G. D., Ostrovsky R., Persiano G. “Public Key Encryption with Keyword Search”. In: Cachin C., Camenisch J. (eds.) EUROCRYPT 2004. LNCS, vol. 3027, pp. 506-522. Springer, Heidelberg (2004).

[2] Peng Xu, Qianhong Wu, Wei Wang, Willy Susilo, Josep Domingo-Ferrer, Hai Jin, “Generating Searchable Public-Key Ciphertexts with Hidden Structures for Fast Keyword Search”, IEEE Transactions on Information Forensics and Security Volume: PP Year: 2015.

[3] Curtmola R., Garay J., Kamara S., Ostrovsky R. “Searchable SymmetricEncryption: Improved Definitions and Efficient Constructions” .In: ACM CCS 2006, pp. 79-88. ACM (2006).

[4] Song D. X., Wagner D., Perrig A. “Practical techniques for searches on encrypted data”. In: IEEE S&P 2000, pp. 44-55. IEEE (2000).

[5] Goh E.-J. “Secure Indexes. Cryptography ePrint Archive”, Report 2003/216 (2003).

[6] Bellovin S. M., Cheswick W.R. “Privacy-Enhanced Searches Using Encrypted Bloom Filters”. Cryptography ePrint Archive, Report 2004/022 (2004).

[7] Bao F., Deng R. H., Ding X., Yang Y. “ Private Query on Encrypted Data in Multi-User Settings”. In: Chen L., Mu Y., Susilo W. (eds.) ISPEC 2008. LNCS, vol. 4991, pp. 71-85. Springer, Heidelberg (2008).

[8] Waters B. R., Balfanz D., Durfee G., Smetters D. K. “Building an Encrypted and Searchable Audit Log”. In: NDSS 2004 (2004).

[9] Chase M., Kamara S. “Structured Encryption and Controlled Disclosure”. In: M. Abe (ed.) Advances in Cryptology - ASIACRYPT 2010. LNCS, vol. 6477, pp. 577-594. Springer, Heidelberg (2010).

[10] Abdalla M., Bellare M., Catalano D., Kiltz E., Kohno T., Lange T., Malone-Lee J., Neven G., Paillier P., Shi H. “Searchable Encryption Revisited: Consistency Properties, Relation to Anonymous IBE, and Extensions”. In: Shoup V. (ed.) CRYPTO 2005. LNCS, vol. 3621, pp. 205-222. Springer, Heidelberg (2005).

[11] Park D. J., Kim K., Lee P. J. “Public Key Encryption with Conjunctive Field Keyword Search”. In: Lim C. H. and Yung M. (eds.) WISA 2004. LNCS, vol. 3325, pp. 73-86. Springer, Heidelberg (2004).

[12] Ballard L., Kamara S., Monrose F. “Achieving Efficient Conjunctive Keyword Searches over Encrypted Data”. In: Qing S. et al. (eds.) ICICS 2005. LNCS, vol. 3783, pp. 414-426. Springer, Heidelberg (2005).

[13] Hwang Y. H., Lee P. J. “Public Key Encryption with Conjunctive Keyword Search and Its Extension to a Multi-user System”. In: Takagi T., Okamoto T., Okamoto E. and Okamoto T. (eds.) Pairing 2007. LNCS, vol. 4575, pp. 2-22. Springer, Heidelberg (2007).

[14] Davis D., Monrose F., Reiter M. K. “ Time-Scoped Searching of Encrypted Audit Logs”. In: Lopez J., Qing S., Okamoto E. (eds.) ICICS 2004. LNCS, vol. 3269, pp. 532-545. Springer, Heidelberg (2004).

[15] Tang Q., Chen X. “Towards asymmetric searchable encryption with message recovery and flexible search authorization”. ASIACCS 2013, pp. 253-264 (2013).

[16] Ibraimi L., Nikova S., Hartel P. H., Jonker W. “Public-Key Encryption with Delegated Search”. In: Lopez J. and Tsudik G. (eds.) ACNS 2011. LNCS, vol. 6715, pp. 532-549. Springer, Heidelberg (2011).

ISSN 2348 – 7968

[17] Boneh D., Waters B. R. “Conjunctive, Subset, and Range Queries on Encrypted Data”. In: Vadhan S. P. (ed.) TCC 2007. LNCS, vol. 4392,pp. 535-554. Springer, Heidelberg (2007).
[18] Cheung D. W., Mamoulis N., Wong W. K., Yiu S. M., Zhang Y. “Anonymous Fuzzy Identity-based Encryption for Similarity Search”. In: Cheong O., Chwa K.-Y and Park K. (eds.) ISAAC 2010. LNCS,vol. 6505, pp. 61-72. Springer, Heidelberg (2010).
[19] Xu P., Jin H., Wu Q., Wang W. “Public-Key Encryption with Fuzzy Keyword Search: A Provably Secure Scheme under Keyword Guessing Attack”. IEEE Transactions on Computers, 62(11), pp. 2266- 2277 (2013).
[20] Camenisch J., Kohlweiss M., Rial A., Sheedy C.: Blind and Anonymous Identity-Based Encryption and Authorised Private Searches on Public Key Encrypted Data. In: Jarecki S. and Tsudik G. (eds.) PKC 2009. LNCS, vol. 5443, pp. 196-214. Springer, Heidelberg (2009).
[21] Bellare M., Boldyreva A., O’Neill A. “Deterministic and Efficiently Searchable Encryption”. In: Menezes A. (ed.) CRYPTO 2007. LNCS, vol. 4622, pp. 535-552. Springer, Heidelberg (2007).
[22] Camenisch J., Kohlweiss M., Rial A., Sheedy C.: Blind and Anonymous Identity-Based Encryption and Authorised Private Searches on Public Key Encrypted Data. In: Jarecki S. and Tsudik G. (eds.) PKC 2009. LNCS, vol. 5443, pp. 196-214. Springer, Heidelberg (2009).
[23] Brakerski Z., Segev G.: Better Security for Deterministic Public- Key Encryption: The Auxiliary-Input Setting. In: Rogaway P. (ed.) CRYPTO 2011. LNCS, vol. 6841, pp. 543-560. Springer, Heidelberg (2011).