Helping Johnny to Search: Encrypted Search on Webmail System

Tatsuya Midorikawa
Akihiro Tachikawa
Akira Kanaoka
Toho University
Miyama 2-2-1, Funabashi,
Chiba, 274-8510, Japan
akira.kanaoka@is.sci.toho-u.ac.jp

Introduction

E-mail encryption has been a long-standing issue of usable security. After Whitten and Tygar has opened a new vista in usable security for E-mail encryption[1], there were many following studies which enhance the usability[2, 3, 4]. Finally, Ruoti et al. has achieved automatic e-mail encryption using Identity-based Encryption without S/MIME and PGP[5]. Such automatic or transparent encryption has become common in various messaging tools. As end users come to reach maturity, end-to-end encryption has come to be widely accepted.

In the e-mail application, there are some functions for messages sent and received, such as searching, sorting, threading and spam filtering. From the viewpoint of usability, it is desirable that these utility functions continue to be used even if e-mails are encrypted.

This paper focus on search function from such utility functions for e-mail. In the current search system, search index is prepared in order to realize a high-speed search. Even if e-mails are encrypted, information leaks will occur in its search index that have e-mail contents information in plain text. As a countermeasure, there are encrypted search (or searchable encryption) techniques. Encrypted search allows searching in encrypted manner without leaking any information about e-mail contents, such as encrypted index.

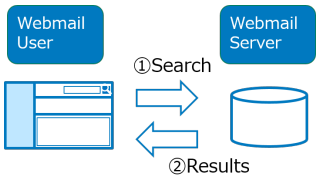

Figure 1: Normal Webmail System Model

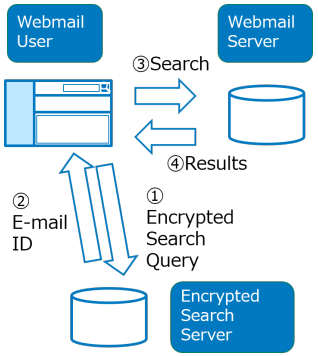

Figure 2: Proposed Model

Various studies have been actively studied in encrypted search field. Encrypted search is one of the most promising techniques. However, it has never been discussed in the usability point of view.

In this paper, a first look at usability of encrypted search is given. At first, a method that enables encrypted search transparently is proposed with the goal of achieving transparently available environment, as Ruoti et al has achieved transparently e-mail encryption. Then, the evaluation consists of two methods: Performance of encrypted search especially on encrypted query generation and encrypted search, and conducting user study using actual environment. In the user study environment, a new encrypted index server using Symmetric Searchable Encryption (SSE) technique is built as for server side. A Google chrome extension for SSE that generates encrypted query and communicate with the SSE server, is also prepared as for end user side. It enables SSE without changing current Google Gmail service. The user interface of end user side is almost identical in the Gmail service.

As a performance result, it shows acceptable levels of performance that takes about 1 msec for generating encrypted query and about 180 msec for one searching with index of 10,000 e-mails. The result of usability evaluation shows there are equivalent between normal Gmail and proposed method applied system.

On the other hand, some interesting reactions are given. For example, similar to the results of Ruoti et al.[5] and Fahl et al.[6], some users still expect that randomized string on the user interface as encrypted contents.

Our contribution is giving discussions in the usability point of view which is made for the first time in the field of encrypted search. And it shows that encrypted search can be applied without vitiating usability.

Encrypted Search to Existing Webmail Systems

In this section, basic approach to apply encrypted search to existing webmail systems and our proposal are described.

Basic Approach

Basically, there are two approaches to apply encrypted search to webmail system. First one is a method for improving the system itself. The other one is a method for add-on functionality without changing existing system. Since add-on function is applicable without the need to large improve existing system, it is applicable to webmail systems used already in many scenes. This approach has great practicality. In our paper, a method for applying encrypted search to existing webmail system by adopting this approach.

Proposed Method to Add Encrypted Search Function into Existing Webmail Service

In webmail systems, webmail users access to web mail server from browser, then send, receive, view, and search e-mails (Fig.1).

If we want to apply encrypted search here, it is difficult to use the APIs and data of existing webmail server directly. Encrypted search dedicated server is prepared to solve this problem. The search index is hold by the encrypted search server. Encrypted e-mails are stored in current webmail server. Each e-mail is assigned with unique ID, and the ID is included in title, header, or body in plain text, even e-mail contents itself are encrypted. In the encrypted index, the keyword with each of mail content is stored. Each keywords has binded to e-mails which has the keyword. Webmail user firstly create an encrypted search query, then send the query to encrypted search server. Encrypted search server searches using received encrypted search query, then returns obtained IDs to the webmail user. The webmail user

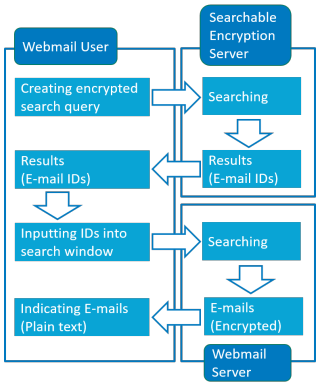

Figure 3: Flow of Proposed Model

Hypothesis

1. Generating encrypted query ($Trapdoor$) on browser and searching ($Search$) on SSE server are acceptable and unaware of additional action
2. There is no difference in user awareness between original Gmail and the proposed encrypted search system without e-mail contents encryption

then makes query of OR search using received e-mail IDs, and send the query to webmail server. As a search result from the webmail server, encrypted e-mails which include IDs on the titles are returned to webmail user. The webmail user decrypts and displays them. In this way, an e-mail system which enables encrypted e-mail and encrypted search, can be achieved. The relationship diagram for each entity is shown in Fig.2 , and the flow of search is also shown in Fig3.

Developed System for Evaluation
For use in the evaluation, a prototype system of proposed system is developed.

Google Chrome which can easily add functionality using Chrome Extension is selected for client side. And Gmail which is one of the most popular e-mail service and enables OR search, is selected for existing email service.

Encrypted Search Method
Curtmola's SSE[7] is implemented in the system. For server side, $Search$ function and web server function are developed using Java language. For client side, $Trapdoor$ function is developed using JavaScript.

Google Chrome Extension
A prototype chrome extension is developed to enable encrypted search on client side. It hook user input on the Gmail search window. When an user push search button, it obtain the value on the Gmail search windows. Then it calculates encrypted search query, then send the encrypted query to the SSE server. After receiving results (E-mail IDs) from SSE server, it generates query for Gmail using E-mail IDs, and send the OR query to Gmail Server. Finally, after receiving result from Gmail server, it decrypts all e-mails on the results then show to a user.

User interface of original Gmail and the proposed system which is Gmail adapted with developed extension, are almost identical . Magnifying glass icon was omitted from Gmail+SSE and a simple lock icon was shown next to the URL bar.

Evaluation Methodology

This section gives an overview of our experimentation for evaluating our proposed system. In the experimentation, performance evaluation and user study are conducted.

Hypothesis
Following the result by Ogata, et al.[8] which shows good performance of SSE, we hypothesize as shown on the side bar.

Performance Evaluation
The purpose of the performance evaluation is giving information to discuss usability between original Gmail and the proposed system (Gmail+SSE) whether the performance affect to the usability. Performance evaluation is conducted independent from following user study.

In Curmola's SSE, since the size of index is largely changed by the number of documents (E-mails), three types of e-mail nums (100, 10000, 10000) are prepared. E-mails are randomly chosen from Enron Email Dataset[9].

User Study
Study Setup
The study ran from July 18, 2017 to December 18, 2017 and included 13 participants that were randomly assigned to test either standard Gmail or the proposed system (Gmail+SSE) that applied encrypted search. Table 1 shows demography of participants.

Table 1: Participants

	Gender	Grade
p1	Female	4th, Bachelor
p2	Female	4th, Bachelor
p3	Male	4th, Bachelor
p4	Female	4th, Bachelor
p5	Male	2nd, Bachelor
p6	Male	2nd, Bachelor
p7	Male	2nd, Bachelor
p8	Male	1st, Master
p9	Male	1st, Master
p10	Male	2nd, Master
p11	Male	2nd, Master
p12	Male	4th, Bachelor
p13	Male	2nd, Master

Table 2: Average E-mail Size and Index Size of Curtmola's SSE

# of e-mails	Avg. E-mail size (KB)	Size (MB)
100	1.18	6.3
1000	1.87	88.1
10000	2.01	941.7

Table 3: Performance of $Trapdoor$ and $Search$

# of e-mails	$Trapdoor$ (ms)	$Search$ (ms)
100	1.30	65.93
1000	1.76	66.61
10000	1.57	179.66

Scenario and Task Design
We prepared e-mail data for the user study as natural as possible and designed realistic tasks. Participants were provided e-mail account during the study and conducted tasks using this e-mail account.

Task Searching for e-mails containing specific keywords from e-mails received in the past, and writing the result on the answer sheet. There are five keywords to find out, and all the keywords are described on answer sheet.

Study Questionnaire
After finishing tasks, semi-structured interview is conducted. Questionnaire of System Usability Scale (SUS) is is used to the interview, and addition to SUS questionnaire .

Usability Analysis
To analyze usability of proposed method, we conducted qualitative analysis using Grounded Theory Approach (GTA) from interview results and behavior during the study and prior orientation.

Evaluation Results

From performance evaluation and Usability evaluation, we found The result shows there are equivalent between normal original Gmail and a proposed system (Gmail+SSE).

Next we discuss these findings in detail based on performance result and categories derived from GTA.

Performance
Even index size is larger (Table 2) , performance of $Trapdoor$ and $Search$ (Table 3) seems acceptable level for end users. Each value is about less than 100 msec and it means that end users might feel instantaneous response. Its value is from "Response-Time Limits" by Jakob Nielsen [1]. This result supports the first hypothesis.

Usability
Categories Derived from GTA
Knowledge of End-to-end Encryption
As same as attitude of encryption, most participants did not know about message encryption in current popular messaging tools.

Perception of Search Keyword Encryption
Most participants did not perceive encryption of search keyword regardless of Gmail or Gmail+SSE.

Usability of the Proposed System
We asked participants about "stress", "easy to use", "usual", and obtain positive answers for the usability of the proposed system. It also supports the second hypothesis.

Attitude toward Message Encryption
Even though one participants think that user interface of encrypted message shows randomized string, most of participants did mention about message encryption and its user interface.

Answers for the question about information leakage tend to mention about e-mail address leakage. It seems that most of participants did not consider e-mail contents. It also reflect some attitude toward message encryption.

Conclusion

Our contribution is giving discussions in the usability point of view which is made for the first time in the field of encrypted search. And it shows that encrypted search can be applied without vitiating usability.

[1] https://www.nngroup.com/articles/website-response-times/

REFERENCES

1. Alma Whitten, J.D.Tyger: Why Johnny Encrypt: A Usability Evaluation of PGP 5.0, 8th USENIX Security Symposium(1999).
2. Simson L.Garfinkel, Robert C.Miller: Johnnny 2: A User Test of Key Continuity Management with S/MIME and Outlook Express, Symposium On Usable Privacy and Security(SOUPS)(2005).
3. Simson L.Garfinkel, David Margrave, Jeffrey I.Schiller, Erik Nordlander, Robert C.Miller: How to Make Secure Email Easier To Use, the SIGCHI Conference on Human Factors in Computing(CHI'05)(2005).
4. Steve Sheng, Levi Broderick, Colleen Alison Koranda: Why Johnny Still Can't Encrypt: Evaluating the Usability of Email Encryption Software, Symposium On Usable Privacy and Security(SOUPS)(2006).
5. Scott Ruoti, Nathan Kim, Ben Burgon, Timothy van der Horst, Kent Seamons: Confused Johnny: When Automatic Encryption Leads to Confusion and Mistakes, Symposium On Usable Privacy and Security(SOUPS)(2013).
6. Sascha Fahl, Marian Harbach, Thomas Muders, Matthew Smith, Uwe Sander: Helping Johnny 2.0 to Encrypt His Facebook Conversations, Symposium On Usable Privacy and Security(SOUP)(2012).
7. Reza Curtmola, Juan A.Garay, Seny Kamara, and RafailOstrovsky, Searchable Symmetric Encryption: Improved Definitions and Efficient Constructions. In Proceedings of the ACMSymposium on Information, Computer and Commucations Security (CCS'06) (2006)
8. Wakaha Ogata, Keita Koiwa, Akira Kanaoka, and Shin'ichiro Matsuo. 2013. Toward Practical Searchable Symmetric Encryption. In Proceedings of the International Workshop on Security (IWSEC'13) (LNCS), Kazuo Sakiyama and Masayuki Terada (Eds.), Vol. 8231. Springer, 151-167.
9. Enron Email Dataset, https://www.cs.cmu.edu/ enron/