# A new technique using a shuffling method to protect confidential documents from shoulder surfers

Hyunsoo Kim[1], Hyoungshick Kim[2] and Ji Won Yoon[1]
[1] Center for Information Security Technologies (CIST), Korea University, Republic of Korea
[2] College of Information and Communication Engineeering, Sungkyunkwan University, Republic of Korea
aitch25@korea.ac.kr, hyoung@skku.edu, jiwon_yoon@korea.ac.kr

***Abstract*—In some environments (e.g., for government agencies or international corporations), it is challenging to protect and secure confidential information on a computer screen against shoulder surfers who want to access the confidential information by observing the victims computer screen. In this paper, we propose a simple and practical system named STM to mitigate shoulder surfers from reading computer screens by visually shuffling contents on an end users screen. To find an optimal setting for STM, we tested several configurations at character and word levels and showed that STM with a properly chosen configuration is effectively secure against using direct observation techniques.**

## I. Introduction

A shoulder surfing attack is one of the simple tricks that make information leak from a monitor. Many researchers have been studying how to protect such confidential information on the screen from the surfers. There already exist several commercial products for preventing the shoulder surfing attack. For example, there are Privacy Filter Film [1] and ePrivacy Filter Software [2] implemented by the commercial company, 3M. They are general methods that resist against adversary users trying to obtain information with the shoulder surfing attack.

However, Privacy Filter cannot protect attacker's gazing at the rear of the user, i.e., standing behind the user, but having the same angle on the vision. Moreover, ePrivacy may cause inconvenience during operation, especially when ePrivacy's protection mode is working. By the protection mode, the screen would be blurred as attacker's face is detected via the webcam. However, the authorized user cannot also read and understand the contents of the document in this condition. Therefore, we need a new approach to protect the user's confidential information from the shoulder surfing attack.

"Graphical method" [3] and "Gaze-based authentication method" [4] are other approaches to prevent the shoulder surfing attack. However, they work only at a condition with password inputs. In other words, they cannot be used for the various situations. It will be referred to more detail in next section.

In this paper, we propose a new approach to protect shoulder surfing attack. We name this as *Shuffling Texts Method* (STM) and it works for protecting information in confidential documents. When a user starts STM, he or she reads plain texts in the document. However, if unauthorized surfers are detected, the plain texts are shuffled and then the shuffled texts are reprinted on the screen. That is, this operation makes the attacker confuse which part they should read on the screen and finally not to obtain any confidential information from the documents. One of the key points of our proposed approach is that authorized users could still read and understand the document although STM is working.

## II. Related Work

Various researches have been conducted in order to maintain the secrecy and privacy of the confidential documents against shoulder surfers. Especially, 3M's Privacy filter film and ePrvacy are famous commercial products which can protect the documents from the malicious surfers. Privacy filter film is almost the same as the protection film attached on the computer screen, however, it has an advanced technology called Microlouver, which additionally protects the screen from a malicious user. In more detail, Microlouver changes the transmission rate of the screen according to the angle at which the attacker is standing from the screen. Therefore, the shoulder surfers on the side of the user cannot obtain any information since the contents on the screen is too blurred to read and understand. On the other hand, a product named ePrivacy is one of the approaches to prevent leaking information from the screen. The software perceives attackers through a webcam. When the attacker's face is detected, the software sets the screen to be blurred to protect information.

In addition, note that most prior researches have been focused on prevention of the secret passwords from shoulder surfing attacks rather than full documents. For example, Manu Kumar introduces a method to input the password in a secure way [4] utilizing a gaze tracking technique. The method obtains the password by gazing the characters of the password. Alternatively, Wiednbeck and Waters [3] suggested a scheme using a graphical method using users' memory. That is, the authorized users remember icons that they choose, and they input icons in regular sequence. In this way, users could securely input their password.

Human have the cognitive ability for recognizing words and sentences. Especially in English, although misspelled words are written, people often understand it either partly or fully. Kreiner demonstrated that a human being can understand the sentences even though they include spelling errors in terms of the human cognitive ability [5].

## III. Proposed approach

Let us assume that an authorizing person Alice is reading a confidential document and she does not want to open the

document to other people. However, a malicious attacker Trudy wants to see what Alice is reading and he is attempting shoulder surfing attack to Alice in order to obtain the secret information.

Given this situation, we can propose a new system to protect the information which satisfies the following conditions:

- Alice can continue to read the confidential document while Trudy cannot understand the document.
- The system should automatically detect the existence of the Trudy.
- The system can work in a software with the help of a few peripheral devices such as webcams.

In this paper, we propose an approach which is based on shuffling technique in order to satisfy the above conditions for protecting documents. In our system, texts in the conditional documents are shuffled with different complexity for Alice and Trudy. Therefore, Alice can read the document with less error or errorless while Trudy cannot read the document due to a myriad of error from the shuffling operation. We name this approach as *Shuffling Text Method* (STM) and it is designed for protecting contents in a document from Trudy's shoulder surfing attack. According to the second and third condition, STM uses webcams in order to detect Trudy's appearance.

Initially, STM first tokenizes and shuffles the texts after a document is loaded. The shuffled texts are stored in a memory. On the screen, STM works in "plain mode" like a conventional document viewer without any text shuffling. In the plain mode, we can read only plain text which is an actual document. However, as soon as STM detects Trudy's shoulder surfing, then STM changes its mode to "shuffling mode". In the shuffling mode, the stored shuffling texts are randomly displayed on the screen. The operating procedure is shown in Fig. 1.

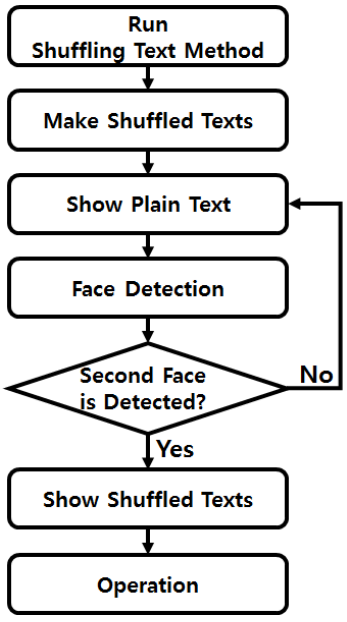



Fig. 1: Shuffling Text Method procedure

### A. *Scheme for shuffling texts*

In the first step, STM starts with a webcam and is loading documents. Afterward, STM extracts texts of the document as a string. Then, shuffles and rearranges them to make shuffling texts. This shuffling scheme is used to protect a document on a screen from Trudy's shoulder surfing attack. The detailed scheme of the shuffling operation is followed in Fig. 2, 3 and TABLE I. Fig. 2 shows a tokenizer which decomposes the texts of the document into word and character levels. After tokenizing the texts, we can freely make various shuffled texts as shown in Fig. 3. We name this process as 'Character Shuffler' although even the characters of the word are shuffled.

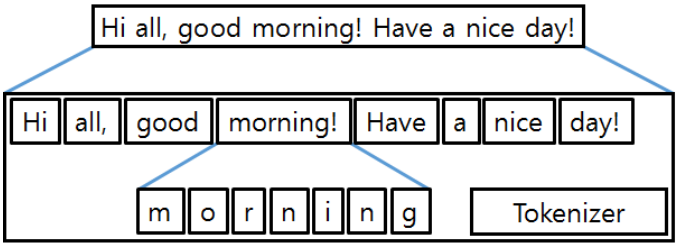



Fig. 2: Tokenizer

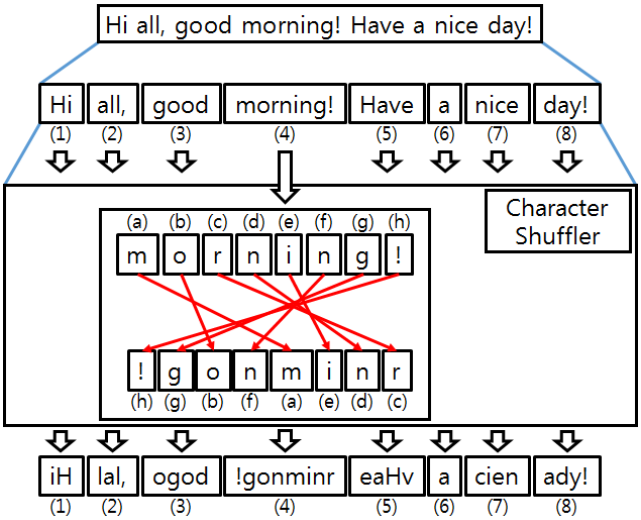



Fig. 3: Character shuffler

The output of the Character Shuffler can be various as plotted in TABLE I. We can easily design the shuffled texts with different shuffling complexity. With level 0, there is no misspelt error in the word so the texts are identical to plain texts. However, as the shuffling level increases, the misspelt errors are increased. The main idea of our proposed approach is coming from this various shuffling level. Since level 0 does not have any misspelt error, STM provides it to Alice. Conversely, it is highly likely for STM to provide shuffled texts with higher shuffling level to Trudy on the screen.

Therefore, in the 'plain mode' where only Alice is reading the document, she read the document with shuffling level 0. However, as soon as Trudy appears in the screen as a shoulder surfer, STM provides shuffled texts with higher shuffling level

to the surfer Trudy while Alice can still read the document with relatively lower shuffling level which is close to 0. From this point of view, Trudy cannot succeed to obtain the secret information from the document.

Even though Alice and Trudy gaze at the same screen at the same angle, they would each see a different part of the screen. Moreover, each word, shuffled in the different level will be printed out at each position. It provides possibility that only Alice can obtain meaningful information from the screen. How it is possible is that Alice possesses one more piece of information which Trudy does not know. The information is a *region of interest* (ROI). In this context, the ROI represents the region where Alice are gazing. The ROI is moved by the cursor, and the cursor is controlled by the user. STM sets the ROI in the screen with the concept of the cursor, and print low-level shuffled words at that region. In this way, Alice can obtain meaningful information, for she knows where the ROI is in the user screen. Conversely, Trudy could not obtain any information while Alice reads the document. The ROI information is not opened to Trudy; it makes her understand the document difficulty. Trudy would read high-level shuffled words spreaded out in the entire screen; it is almost impossible that she obtains significant information.

TABLE I: Shuffling level and shuffled words

| Shuffling level | Shuffled words | Shuffling level | Shuffled words |
|---|---|---|---|
| 0 | Shuffling | 3 | hfufiSlng |
| 1 | Shuffilng | 4 | fuSnlihgf |
| 2 | Sfhfnulig | 5 | gfIfSunhi |

For an advanced shuffler, STM can further embed a model that each shuffled words are optionally switched their position each other in a paragraph as described in Fig. 4. In this procedure, texts are shuffled not in character level, but in words level. The words are shuffled once again among the words which are the same length. Yet, this process might make Alice confuse to read sentences through STM although this advance shuffling operation obviously improves the security level against Trudy. Therefore, STM provides this function optionally.

The reason why different levels of shuffling exist is for increasing readability. As the procedure of Fig. 4 is not working due to the user setting, word-level shuffling would not be conducted. In this condition, readability of the document in STM is influenced by shuffling level. Words near the cursor would be shuffled in low level, that is around a level of 0 $\sim$ 2. In the shuffling level 0 $\sim$ 2, users can read the words continuously even though the words are slightly misspelled. For example, if the cursor is on the first word in a sentence, the shuffle level of that word is 0. And the word next to the first word has level 1 of shuffling. In this rule, the word on the third position might be shuffled in level 2. Through this method, users can read documents without delay, and it would be a good point as a secure document viewer program.

### B. Face detection algorithm to recognize Trudy

As it is referred to before, the procedure of "Face Detection" is used to detect attacks from a malicious user Trudy.

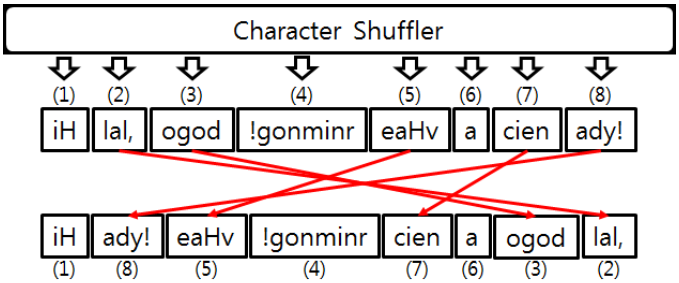
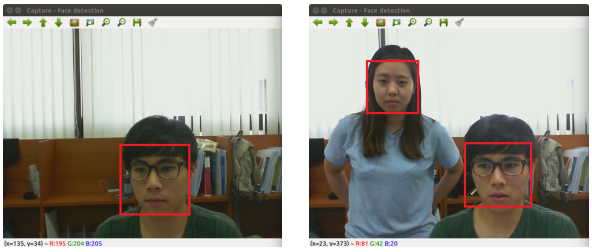



Fig. 4: One more shuffle for increasing safty



(a) A user is perceived (b) An attacker is perceived

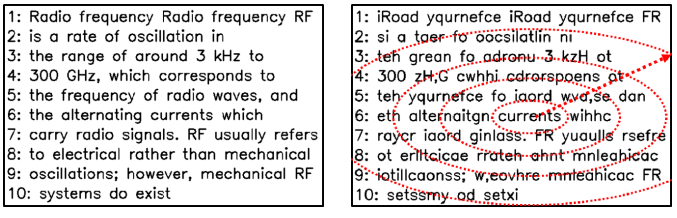
Fig. 5: Webcam perceives attacker with face

This face detection module is used to recognize the shoulder surfer Trudy who positions behind the normal user to steal information from the users screen as shown in Fig. 5.

Fig. 5a represents the condition that the normal user Alice gazes a monitor screen. And the condition, which is attacked by a malicious user Trudy, is captured in Fig. 5b. As can be seen in Fig. 5b, if the second face is detected through a webcam during operation, STM works to treat the threat such as a shoulder surfing attack by printing shuffled text out on the screen, because the documents on the screen are assumed to be exposed by the threat of stealing information.

We used Haar [6] [7] algorithm in STM in order to detect attacker's face since Haar algorithm is one of the famous face detection algorithms. It is not only powerful for detecting faces but also fast as much as it can operate on a smartphone without load.

### C. Displaying shuffled texts in a shuffling mode



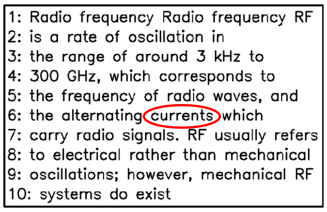

(a) Plain texts (b) Shuffled texts

Fig. 6: General and protection mode

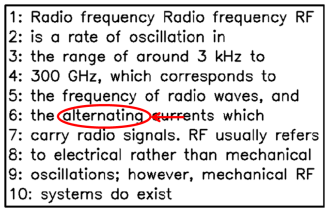
This step is one of the main steps in a protection mode. In this step, the shuffled texts would be printed out on the screen. The shuffled texts, which are created in the previous steps, will be used in this step. By utilizing it, a normal user Alice could preserve the information on the screen while the attacker Trudy would not be able to understand the sentences on the screen that had been shuffled in enough.

Fig. 6a represents the condition that plain texts are printed out on the screen. If there are no malicious surfers, STM prints plain texts as shown in the left picture. On the other hand, the shuffled texts which Fig. 6b describes are displayed on the screen as soon as the attacker is detected by STM. While the shuffling mode is operating, attackers must be felt hard to understand sentences of the document on the screen. Therefore, even though the attacker attempts a shoulder surfing attack, the normal user could protect his or her own information through this way.

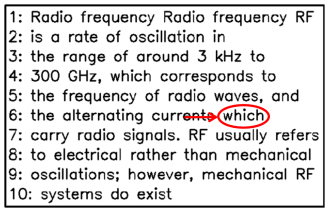
### *D. The concept of region of interest and cursor*




(a) Cursor is at the center of the screen




(b) Cursor moves to left




(c) Cursor moves to right

Fig. 7: The concept of cursor

In this section, we explain the concept of the cursor detailedly. As previously mentioned, ROI is the region in the screen where a normal user is reading at the moment. Words near the ROI might be shuffled in low level, and the shuffling level would increase as the words become far from the cursor position. The concept of the ROI is used to distinguish between a normal user and a malicious user. The normal user who is reading a document knows where the ROI is. Therefore, they can read and understand the document while protection mode is working. However, the malicious user who does not know which part the normal user is reading in the document, cannot steal any information even though the attacker stares at the screen. The cursor is utilized for designating ROI in the screen. The cursor is controlled by the normal user, they can move the position of ROI by controlling the cursor. In order to operate t he STM with the concept of the cursor and ROI, it is necessary to use control device. We used a keyboard as a control device. That is, the cursor was controlled by the keyboard. Fig. 7 is describing cursor operation. If the normal user presses the left or right button, the cursor moves to left or right as it is shown in Fig. 7b, 7c. Additionally, there is an index number for finding cursor. Whenever the user lose where the cursor is during operation, they can find cursor via index number. The cursor will come to next to the index number as the user press the index number on the keyboard.

## IV. Experimental results

We evaluate the usability of STM and performance of face detection in this section. The STM is a kind of security-adapted model for a document viewer. Accordingly, we need to consider both security and usability since the shuffled texts are harder to read and understand than normal document viewer in terms of usability even for the normal user Alice. Therefore, we need to evaluate this program to find a new way to make usability become higher while enough security levels are maintained.

### *A. Setting for usability test*

The procedure for the usability test as follows: 1) A participant sits on a chair. 2) A moderator explains about the experiment to obtain the participant's agreement. If the participant agrees about the experiment, we pay 5$. 3) We do a demographic survey. 4) The moderator explains to Participant about experiment methods. 5) The participant reads the contents with STM while an attacker tries a shoulder surfing attack. 6) Participant and attacker write on the paper what they understand. If the attacker successes to steal information, the moderator pays 1$ to the attacker as an incentive. 7) Participants evaluate tool's usability

TABLE II plots the detailed description about how we set options in this experiments. Twelve people are participated in the experiment. We distinguish them with age, academic background, and country. We also choose two skilled attackers for the experiment. Additionally, we choose 30 paragraphs for a test, paragraphs are referred to from English reading book for middle school second-grade. Paragraphs were composed by around 100 words.

### *B. Language dependency*

We also tested how much STM is dependent on the language and countries. The most affected to understand the document is related with the country as expected. Participants who are using English as mother language answered "Very Easy" 2 times and another participant who is using English as official language answered "Easy" for the question of "Was the text easy to understand?" whereas participants from non-English-speaking countries answered "Normal" averagely: "Easy" 2 times, "Normal" 3 times and "Hard" 4 times . However, one level difference could not be construed as they have a big gap in their English ability. In the experiment, we reduce the gap of English ability by choosing easy English paragraphs. On the other hand, factors such as gender, age, the academic background could not influence on understanding the paragraph, as we choose an easy level paragraph that we mentioned before. However, only one person was influenced by this factor. The person from non-English-speaking country could not understand the sentence properly.

### C. Understanding rate to read

With the question "what do you think the good in the STM", participants answered "it is really easy to use", "make me focus on the text", "good for protecting English information." However, several participants answered that "it is bit hard to use for reading a document." While they spent around 1 minute per a paragraph for a normal user to read and understand the document without STM, the elapsed time was increased around 1 minutes averagely with STM. [1] Reading speeds are different depending on each people and it makes hard to steal information. One of the subjects said that they would not follow the cursor if they had missed the position of the cursor.

TABLE II: Experiment demographics and result

| | | Participants | Attackers |
|---|---|---|---|
| Number of participants | | 12 | 2 |
| Gender | Male | 10 | 1 |
| | Female | 2 | 1 |
| Age | Less than 20 | 2 | - |
| | 21-25 | 2 | 2 |
| | 26-30 | 8 | - |
| Acacemic background | High School | 2 | - |
| | Under Graduate | 2 | - |
| | Graduate School | 8 | 2 |
| Mother language | Korean | 8 | 2 |
| | English | 2 | - |
| | Chinese | 1 | - |
| | Hindi | 1 | - |
| Understanding Rate | Average | 87.5% | 20.8% |
| Elapsed Time | With STM | 2min 15sec | - |
| | Without STM | - | 45sec |
| | Finding Cursor | - | 2min 2sec |

Another factor that influences the result was English ability. An attacker could follow the cursor and understand what was written on the screen if the participants English ability is worse than the attacker's one by following cursor well. But when the participants are good at English or if participants try to make attacker confuse consciously, an attacker could not understand the contents almost. As it is mentioned in the TABLE II, participants of the experiment was understand 87.5% of a paragraph per person on average. However, attackers could understand only 20.8% of the paragraph. This means that our proposed approach is influential to prevent the shoulder surfing attack. Attackers feel hard not only to understand the contents of the paragraph but also guess what the paragraph is indicated. To measure understanding rate, we required participants to answer what they understand from the document by writing after the experiment. Then we compared the answer and document which they read via STM. In this experiment, the result was difficult to be expressed quantitatively. Therefore, we set points depend on whether they understand or not. More specifically, if they understand the key point of the texts, we gave them 1 point. Otherwise, we gave 0 points. Additionally, we scored about the attacker in the same way. As a result, we measured score in this way, extract understanding rate by calculating the average.

[1]Note that, of course, if the reader is adapted to STM, they could read it for a second.

### D. Elapsed time

It is revealed that STM badly affects the elapsed time. Participants who use STM to read a document take double or more the time than users who do not use it. In fact, users reading the paragraph with STM spent around 2 minutes while the users who do not use STM spent about 45 seconds. It is related with the usability factor because the meaning of the result appeals that the program has a weakness for reading paragraph briefly in usability.

We conducted another experiment to estimate how much attackers spend until finding a cursor on the screen. For this experiment, we randomly set the initial cursor position at first and then move it sequentially. In this way, we make attackers confuse to find the cursor where is ROI. We confirm that attackers could find the cursor around 2 minutes. This result is meaningful because users spent 2min 15sec to understand all texts with STM. However, attackers spent 2min 2sec to find the cursor. The result represents that attackers would not be able to steal any information from the screen, if they could not anticipate the cursor positon.

### E. Difficulty of STM

The difficulty of using STM had been reputed as normal or hard level. After the experiments, we asked participants with question "Was the tool easy to use?" and they answered as described in TABLE III. Some people evaluated the difficulty of this tool as "Very easy" or "Easy" but most people felt hard to use it. This implies that STM is not good to use in terms of usability. Therefore, we need to improve its usability for the further work.

TABLE III: Difficulty of STM

| | Korean | English | Chinese | Hindi |
|---|---|---|---|---|
| Very easy | 1 | - | - | - |
| Easy | 5 | - | - | - |
| Normal | - | 1 | - | - |
| Hard | 2 | 1 | 1 | 1 |
| Very hard | - | - | - | - |

According to the answers from participants who felt uncomfortable to use STM, it is revealed that this tool is not intuitive. Especially, some participants tend to lose the cursor position frequently during operation. This implies that the tool does not offer intuitiveness for finding a cursor. Consequently, we concluded STM has limitation about operating method using the keyboard. Users feel inconvenient to operate the tool due to the gap of the reading speed. Reading the document with their eyes is faster than the keyboard operation, hence users had frequently lost the position where the cursor is. Therefore, more advanced approaches such as gaze-tracking are needed to make it convenient to use.

### F. Performance of the webcam

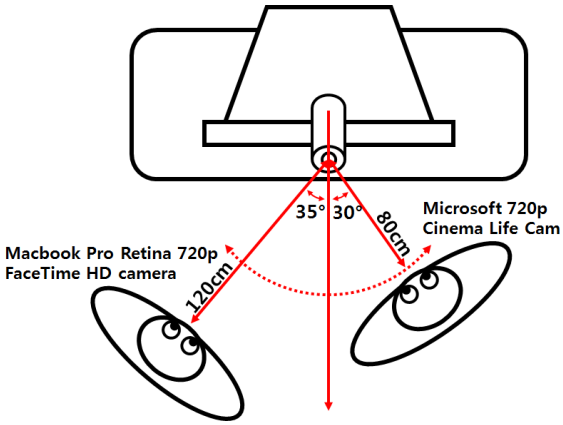



Fig. 8: Webcam conditions for shoulder surfing attacks

Fig. 8 explains the performance of webcams and a face detection algorithm. We chose two webcams. The first one is the Life Cam Cinema made by Microsoft, and the another is FaceTime HD camera in MacBook Pro Retina 13' mid-2014. These support 720p HD video. For face detection, we utilized a famous face detection algorithm based on Haar in OpenCV. In order to improve its performance in real time conditions, we implemented a simple averaging filter in the algorithm. We experimented with this environment, we can obtain the result that is represented in Fig. 8. The Microsoft webcam detects a face within a distance of 80cm and an angle of 30 degrees. The MacBook camera could detect a face within a distance of 120cm and angle of 35 degrees maximally. The result implies that the perceiving procedure might be influenced by the performance of the webcam.

## V. Future work

In this paper, we introduced a new method named STM which for preventing shoulder surfing attacks. However, our proposed approach has some problems yet. The problems are STM could still be exposed by the threat, especially from the well-trained attacker. We find that the attacker could be learned tricks to steal information over the STM as the experiments are repeating. The attacker who repeats attacking more than 10 times in the same environment could steal information with a good success rate in comparison with an attacker who is not trained. It means that an attacker has adapted to the environment of STM during the attack. On the other hands, the necessity of research for increasing convenience about utilizing STM has emerged through the experiment. Several participants have pointed out the flaw that STM's operation is not intuitive.

In order to complement the flaws, we considered applying gaze-tracking technique to STM. [8] [9] [10] Moreover, STM can be used for increasing both security and usability by utilizing suggested gaze-tracking. Gaze-tracking does not move the cursor with keyboard or mouse explicitly. Instead, users could directly move the cursor to a position just by gazing where they read continuously. In this way, it might perform for not only increasing usability but also improving the security of STM. It is because the cursor would not be moved in a uniform pattern in different with the condition that utilizing mouse and keyboard. Hence, an attacker must feel hard to steal meaningful information from the screen.

## VI. Conclusion

As a method for preventing a shoulder surfing attack on a screen, we can effectively use *shuffling texts method* (STM) which displays shuffled texts to the malicious shoulder surfers. Even though a few participants feel uncomfortable to use it, they could understand the meaning of paragraph without any problem. Furthermore, the result is enough to satisfy the requirements from a special organization dealing with highly confidential documents. Unfortunately, today's standard methods for protecting shoulder surfing attack are mainly focused on the password input condition only. The method is a fully new concept and approach to preventing the shoulder surfing attack in that there is no approach performed for a document before. We have demonstrated how much it could prevent the shoulder surfing attack with various people. Therefore, we conclude that STM is enough to be a method for protecting information on the screen.

## Acknowledgment

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and Future Planning (NRF-2013R1A1A1s012797). This work was also supported in part by the National Research Foundation of Korea (No. 2014R1A1A1003707), the ITRC (IITP-2015-H8501-15-1008), and the NIPA (NIPA-2014-H0301-14-1010).

## References

[1] R. Austin, "Privacy filter for a display device," Jun. 18 1996, uS Patent 5,528,319. [Online]. Available: https://www.google.com/patents/US5528319

[2] 3M, "eprivacy filter software." [Online]. Available: http://solutions.3m.com/wps/portal/3M/en_US/3MScreens_NA/Protectors/Tips-Resources/Product-Catalog/~/3M-ePrivacy-Filter-Software-Professional-Version?N=8703164+3293735472+3294857497&rt=d

[3] S. Wiedenbeck, J. Waters, L. Sobrado, and J.-C. Birget, "Design and evaluation of a shoulder-surfing resistant graphical password scheme," in *Proceedings of the working conference on Advanced visual interfaces*. ACM, 2006, pp. 177–184.

[4] M. Kumar, T. Garfinkel, D. Boneh, and T. Winograd, "Reducing shoulder-surfing by using gaze-based password entry," in *Proceedings of the 3rd symposium on Usable privacy and security*. ACM, 2007, pp. 13–19.

[5] D. S. Kreiner, S. D. Schnakenberg, A. G. Green, M. J. Costello, and A. F. McClin, "Effects of spelling errors on the perception of writers," *The Journal of general psychology*, vol. 129, no. 1, pp. 5–17, 2002.

[6] R. Lienhart and J. Maydt, "An extended set of haar-like features for rapid object detection," in *Image Processing. 2002. Proceedings. 2002 International Conference on*, vol. 1. IEEE, 2002, pp. I–900.

[7] P. Viola and M. J. Jones, "Robust real-time face detection," *International journal of computer vision*, vol. 57, no. 2, pp. 137–154, 2004.

[8] A. Duchowski, *Eye tracking methodology: Theory and practice*. Springer Science & Business Media, 2007, vol. 373.

[9] T. Ohno, N. Mukawa, and A. Yoshikawa, "Freegaze: a gaze tracking system for everyday gaze interaction," in *Proceedings of the 2002 symposium on Eye tracking research & applications*. ACM, 2002, pp. 125–132.

[10] K. Talmi and J. Liu, "Eye and gaze tracking for visually controlled interactive stereoscopic displays," *Signal Processing: Image Communication*, vol. 14, no. 10, pp. 799–810, 1999.