Pak. j. eng. technol. sci.
Volume 1, No 2, 2011, 106-113
ISSN: 2222-9930 print
ISSN: 2224-2333 online

A Novel Text Steganography Technique to Arabic Language Using Reverse Fatha (الفتحة)

Mujtaba S. Memon[1] and Asadullah Shah[2]

ABSTRACT

This paper presents a new technique for information hiding in Arabic language. The technique uses text steganography to hide the information. Considering the existence of Harakat (Fatha, Kasra and Damma) in the languages a new approach of reversing the Fatha has been developed for message hiding. The technique has been also evaluated to ensure its quality. This technique can be applied on other languages like Persian, Sindhi and other Regional languages.

KEYWORDS: Steganography, Arabic Text, Text Steganography, Information Security, Fatha, Reverse Fatha and Harakat.

1. INTRODUCTION

The security of information has gained special significance due to heavy penetration of information exchange devices like computers, mobile phones in our daily life routine. One of the major concerns of information security is to securely exchange information between the communicators. For this purpose various techniques like cryptography, steganography, coding and digital watermarking (Cox I. J., et al. 2008) are being used. Steganography is derived from two Greek words "*Steganos*" and "*Graphia*" (Cox I. J., et al. 2008) means "covered" or "hidden" and "written" respectively. The main purpose of this technique is to hide data in a cover media so that data is undetectable by anyone else. The simple implementation method of this technique is to hide text data in cover image file. The cover media for a steganography technique can text (Shirali-Shahreza M. H. and Shirali-Shahreza M., 2006) (Gutab A. and Fatani M. 2007) (Bennett K., 2004), audio (Gopalan K., 2003), image (Doerr G. and Dugelay J. L., 2003) (Doerr G. and Dugelay J. L., 2004) or video

[1] Engr. Mujtaba S. Memon, Lecturer, Telecommunication Department, College of Engineering, Institute of Business Management, Karachi, Pakistan. Email: mujtaba.memon@iobm.edu.pk.

[2] Dr. Asadullah Shah, Professor, Kulliyyah of Information and Communication Technology, International Islamic University, Kuala Lumpur, Malaysia. Email: asadullah@kict.iium.edu.my

(Chandramouli R. and Memon N., 2001) (Shirali-Shahreza M., 2005). Mostly text steganography is not used as it contributes in increasing the difficulty level for detection of hidden bits while text cover data offers smaller memory occupation and is simpler to communicate (Shirali-Shahreza M. H. and Shirali-Shahreza M., 2006). Steganography add another layer of secrecy *undetectability* over cryptography's *confidentiality* (Bohme R. 2010). The individuals can not notice the existence of secret information if steganography is used while they can observe the encoded secret information if cryptography is used.

The prime objective of this research is to develop a new steganography technique for Arabic language. This work covers on applying steganography in such a way that although the language text remains accessible to everyone but only the concerned receiver is able to extract the crucial information. The next phase is to evaluate the quality of technique which is defined in (Bohme R. 2010) steganography as how difficult is to detect the presence of hidden information or to break the technique.

2. METHODOLOGY

2.1. Proposed Technique

Languages like Arabic use Harakat or short vowel marks (Fatha, Kasra and Damma) for the correct pronunciation of a word. A single word in these languages has multiple meanings depending on the pronunciation handled by Harakat. Harakat can be used on a single alphabet of these languages. Few changes have been made in these Harakat to achieve steganography in this research like reversing the Fatha.

The technique reversal of Fatha; is used in this research to hide secret messages within the text. To achieve reversal of Fatha one needs to change the display manner of an original Fatha. Figure 1 and 2 shows this technique being implemented on a single alphabet "NOON" of Arabic language. The original display manner of a Fatha is a small line inclining from left to right as shown in Figure 1. Figure 2 shows the display manner of a Fatha that has been changed to a small line declining from left to right called as Reverse Fatha.

Figure 1 - NOON with Original Fatha

Figure 2 - NOON with Reverse Fatha

In this technique alphabets in a text are selected in a manner that hidden message can be extracted by simply placing them on side by side manner. Reverse Fatha is used on these selected alphabets instead of an original Fatha. These selected alphabets with reverse Fatha are treated as a secret alphabet within the text.

Figure 3 shows information has been hidden in cover text with original Fatha (as shown in figure 4) with the help of reverse Fatha technique. The information hidden in the cover text is shown in figure 5.

Figure 3 – Information hidden in cover text with Reverse Fatha

Figure 4 – Cover text with original Fatha

Figure 5 – Hidden information

The algorithm designed to hide the secret message in this technique is as follows

- Break each word of secret message into alphabets.
- Find the above alphabets having Fatha in the cover article.
- Reverse the Fatḥ a of these alphabets with the help of new Font family created as mention earlier in a word processing software.

The algorithm designed to detect the secret message on receiving end involves the following steps

- Mark the alphabets in the article with reverse Fatha.
- Extract and simultaneously place the alphabets starting from the first letter of the article.
- Combine these extracted alphabets to achieve the secret word.

2.2. Implementation

The implementation process involves following phases

- Creation/Editing of font that is Reverse Fatha.
- Installation of the fonts in the Windows Fonts.
- Using the font in word processing softwares like MS Word 2007, WordPad to implement Reverse Fatḥ a technique.

Software named Font Creator version 6.2 and predefined fonts for the Arabic language are used in this research. The process begins by editing a glyph Fatha in Arabic Transparent font family having Unicode $064E to glyph Reverse Fatḥ a having same Unicode in the same font family. To achieve this just change the font style of Fatha with the help of software Font Creator version 6.2 by using following steps

- Initially open the font family (Arabic Transparent in this case) in Font Creator software.
- Open the desired glyph of font (Fatha in this case).
- Change the direction of Fatha by using the mirror button available in the drawing toolbar.
- Finally save the glyph further save the font family with a new name.

The next phase in the process is to install the font which can be done simply by copying newly made font file into Font folder of the Windows. The last phase involves usage of any word processing softwares (MS Word in this case) for hiding the secret message with the help of algorithm earlier mentioned. It should be insured that Arabic fonts should be installed before using the fonts.

3. EVALUATION

3.1. Method of Evaluation

The purpose to evaluate this technique was to observe how strong or efficient this technique is in the real world. This technique was developed to ensure securely hide the secret message but if someone would easily break the technique and view the secret message then there was no point to use the technique. Initially some secret message was hidden in an article of Arabic language using this technique, and then a questionnaire was developed to ask the readers if they were able to find the secret message. The article and questionnaire were distributed among different types of targeted audiences like some ordinary readers and some having familiarity with steganography or cryptography. It was obvious that negligible amount of ordinary readers would be suspicious about the article. Hence a hint for presence of secret message was given to the readers with the help of a question in the questionnaire. It was expected that this hint would be able to raise the level of doubt in the readers and correspondingly increase their focus level to detect the secret message.

3.2. Statistics of Evaluation

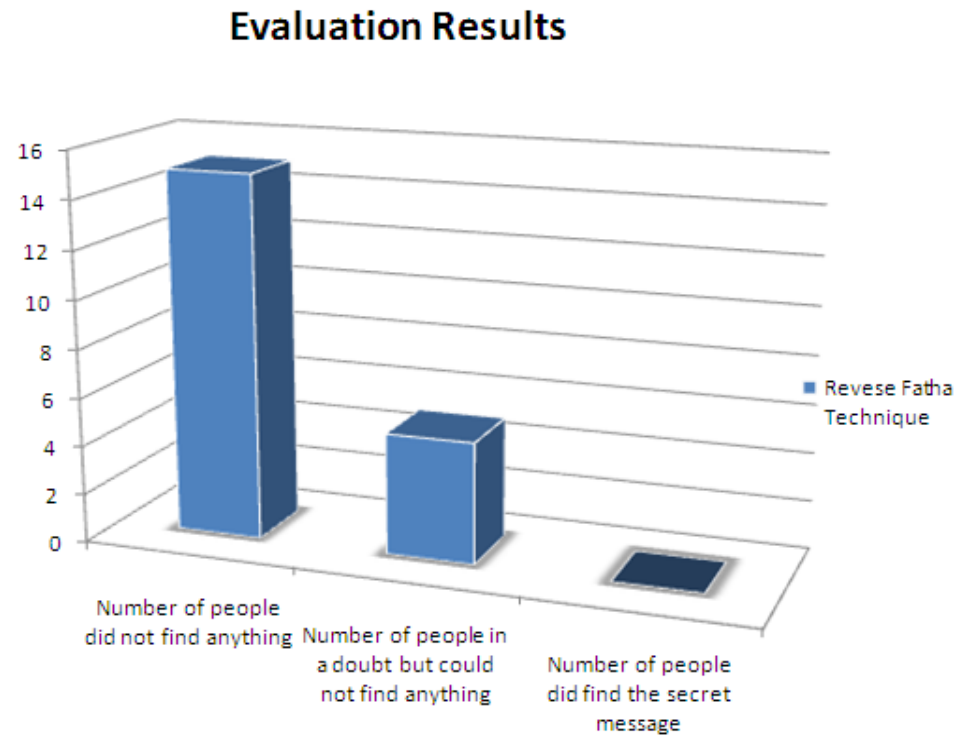

Figure 6 - Bar chart for results obtained in evaluation process

Figure 6 shows the statistics that have been calculated from the questionnaire asked in evaluation process. The article with questionnaire was distributed amongst 20 people, out of whom only 5 people were in a doubt while the rest could not find anything suspicious. These statistics show that even by giving the readers a hint for presence of secret message, about 75% of the readers neither were able to detect anything nor had any doubt. The remaining 25% of the readers had some kind of doubt about the presence of a secret message in the article but were unable to detect the secret message. It was observed that none of the readers was able to detect the secret message or the technique used for the steganography.

4. CONCLUSION

In this research a new approach for steganography of information in Arabic text is introduced. This technique is based on the existence of Harakat in majority of Arabic alphabets. The information was hidden in the text by placing the reverse Fatha. This technique can be used in hidden exchange of information through text documents and text watermarking. In addition to establishing secret communication, this technique can be used for preventing illegal duplication and distribution of text especially electronic text (Westfield A., 2001) (Westfield A. and Pfitzmann A., 2000). This technique can be applied on hard copy

documents, an OCR (Optical Character Recognition) and software is needed to detect the hidden information on receiver end. This technique can be used in languages like Persian, Urdu, Sindhi and other regional languages considering the similarity of these languages with Arabic. The number of hidden bits of information can be increased by combining reverse Fatha technique with size or width variation of other Harakat, and (or) with line shifting and word shifting techniques. A combination of different text steganography techniques with this technique can be applied in the same article.

5. RECOMMENDATION

Further work can be carried out for the following

- To design software that extracts the hidden alphabets from hard copied articles using this technique.
- To implement this technique for other languages like Persian, Urdu, Sindhi and other regional languages.
- To implement watermarking by using this technique. The original extra information is sent to attend and authenticate the fact that message is indeed sent from the particular source in watermarking.
- To increase the level of security by incorporating steganographic techniques with proposed information security techniques like cryptography, watermarking or both.
- To design techniques for Reverse Kasra expansion of Harakat.
- To design a technique to further increase the level of hidden bits by combining techniques like Reverse Fatha with Reverse Kasra, Line Shifting and Word Shifting techniques.

REFERENCES

[1] Cox I. J., et al. (2008) “Digital Watermarking and Steganography”, 2nd Edition, Morgan Kaufmann Publishers, Elsevier Inc.

[2] Shirali-Shahreza M. H. and Shirali-Shahreza M., (2006) “A New Approach to Persian/Arabic Text Steganography”, 5th IEEE/ACIS International Conference on Computer and Information Science, 310- 315.

[3] Gutab A. and Fatani M. (2007) “A Novel Arabic Text Steganography Method Using Letter Points and Extensions”, World Academy of Science, Engineering and Technology, 27, 28-31.

[4] Bennett K. (2004) “Linguistic Steganography: Survey, Analysis and Robustness Concerns for Hiding Information in Text”, Purdue University, CERIAS Tech.

[5] Gopalan K., (2003) “Audio Steganography using bit modification”, Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, 2, 421-424.

[6] Doerr G. and Dugelay J. L., (2003) “Ä Guide Tour of Video Watermarking”, Signal Processing: Image Communication, 18, 263-282.

[7] Doerr G. and Dugelay J. L., (2004) “Security Pitfalls of Frame by Frame Approaches to Video Watermarking”, IEEE Transactions on Signal Processing, Supplement on Secure Media, 52, 2955-2964.

[8] Chandramouli R. and Memon N., (2001) “Analysis of LSB based image steganography techniques”, Proceedings of International Conference on Image Processing, 3, 1019-1022.

[9] Shirali-Shahreza M., (2005) “An Improved Method for Steganography on Mobile Phone”, WSEAS Transactions on Systems, 4, 955-957.

[10] Bohme R. (2010) “Advanced Statistical Steganalysis”, Springer-Verlag, Berlin.

[11] Westfield A., (2001) “A Steganographic Algorithm High Capacity Despite Better Steganalysis”, LNCS 2137, Springer-Verlang, 2137, 289-302, 2001.

[12] Westfield A. and Pfitzmann A., (2000) “Attacks on Steganographic Systems”, Proceedings of 3rd international Workshop on Information Hiding, Lecture Notes on Computer Science 1768, Berlin: Springer-Verlang, 61-75.