Hiding Short Message Text In The Uzbek Language

Nodir Rasulovich Zaynalov
Head of the information security department,
Samarkand branch of Tashkent University of Information Technologies named after Muhammad al-Khwarizmi.
Samarkand, Uzbekistan
nodirz@mail.ru

Ulugbek Khayrullaevich Narzullaev
Dean of the faculty of Telecommunications technologies and professional education,
Samarkand branch of Tashkent University of Information Technologies named after Muhammad al-Khwarizmi.
Samarkand, Uzbekistan
ulug1956_56@mail.ru

Abdinabi Nuralievich Muhamadiev
Teacher of the information security department,
Samarkand branch of Tashkent University of Information Technologies named after Muhammad al-Khwarizmi.
Samarkand, Uzbekistan
nabi8888@bk.ru

Obid Nizomovich Mavlonov
Teacher of the information security department,
Samarkand branch of Tashkent University of Information Technologies named after Muhammad al-Khwarizmi.
Samarkand, Uzbekistan
mavlonov_obit@mail.ru

Jasur Utkirovich Kiyamov
Ph.D. student of Saint Petersburg State University, Russian Federation.
st080634@student.spbu.ru

Dusmurod Qilichev
Teacher of the information security department,
Samarkand branch of Tashkent University of Information Technologies named after Muhammad al-Khwarizmi.
Samarkand, Uzbekistan
sinfdosh1990@mail.ru

***Abstract*— Steganography is a method that is based on hiding or embedding additional information in digital objects while causing some distortion of these objects. In this case, an object or container can be used image, audio, video, network packets, etc. Recently, there have been a lot of publications in the field of hiding information in a text container. To embed a secret, steganographic methods rely on redundant information about the covering medium used or properties that the human perception system cannot distinguish. Since text documents are widely used in organizations, using a text document as a storage medium may be the preferred choice in such an environment, while having little excess information. Therefore, the choice of using a text document as a storage medium is the most difficult, since it contains less redundant information. In this article, we present a simple and new approach to steganography based on the Uzbek language. We analyze the characteristics of the Uzbek language and offer some effective methods for converting messages to bits. We find that a single Uzbek text message can be converted to multiple secret bits. The results show that the proposed method is very important for a successful steganographic technique. Besides, we find that embedded bits can be correctly and easily extracted by human observers without using a special decoder. Thus, the proposed scheme is practical and effective for hiding messages and is a new approach for the Uzbek language.**

Keywords*— *Information hiding, steganography, text steganography, Uzbek language

I. INTRODUCTION

The task of protecting information from unauthorized access has been solved at all times throughout the history of mankind. Already in the ancient world, there were two main directions for solving this problem that still exists today: cryptography and steganography.

Unlike cryptography, the purpose of classical steganography is to hide secret data in other open data sets or streams in a way that does not allow you to detect the presence of some hidden component and thus distinguish these messages from the rest.

Steganography is a field of knowledge that deals with the hidden transfer of information, i.e. the fact of information transfer is hidden. A common feature of steganographic methods and algorithms is that the hidden message is embedded in some harmless, non-attention-grabbing object, which is then openly transported to the recipient. If the object is text, it is text steganography, where character text is used to hide secret information. Storing text files requires less memory, and its wider use in communication makes it preferable to other types of steganographic methods. Because texts take up less memory, transmit more information, and require less printing costs, as well as some other advantages.

In this paper, the possibilities of the Uzbek language in text steganography are studied for the first time and an algorithm for embedding a sequence of bits in the generated text in Uzbek is proposed.

The rest of the article is organized as follows: Section 2 describes some of the existing approaches based on the features of languages. Section 3 describes the proposed approach. Section 4 provides an assessment of the proposed method and its further development. In section 5, we conclude and draw the appropriate conclusions.

II. EXISTING APPROACHES

It should be noted that the popularity of text steganography methods has led to a variety of approaches [1, 2]. after analyzing them, it was found that many works use grammatical features of the language to introduce information. You need to understand that the analysis of available methods allows you to identify different approaches based on the definition of features. In [1], which can be considered one of the classic works in this field, where punctuation marks are chosen as a feature.

So, depending on the type of implementation technique, these methods can be divided into the following categories: 1) syntactic, 2) semantic, and 3) arbitrary type.

Let's briefly consider some approaches to understanding these methods.

It should be noted that many syntactic methods are easily embedded in any text, regardless of its content, purpose, and language. Such systems are easy to develop and automate this process. But, unfortunately, syntactic methods are easily hacked, and secret information can be easily eliminated by

simple attacks. However, the appeal of these methods lies in their speed, which allows them to be applied in many areas.

Methods that can be applied in many languages are methods where punctuation marks are used, such as a dot (.), comma (,), etc., which are inserted in the appropriate places to hide bits 0 and 1 [1, 3, 4]. One of the variations of this method is the Punctuation Mark method [2]. What is used here is that there are some sentences in English in which the appearance of a punctuation mark, such as a comma, becomes optional?

The well-known White Steg method uses spaces to hide a secret message. For example, one space after a word represents bit 0, and two spaces after a word represent bit 1 [4, 5]. Depending on the container type, this method can be modified.

Features of grammatical constructions allow you to get interesting results. So, in the Telugu language, there are various types of consonant characters, as well as punctuation marks, which are successfully used as information carriers to hide information in text media [6].

In practice, there is a huge number of spelling errors in the text. And this effect is used in the Mistyping method, which embeds a secret message by intentionally creating some spelling errors or changing the position of characters in a text document [2]. Since this type of typo is very common in a text document, this method of hiding data may not attract third parties if they are unaware of the secret communication. Of course, it should be noted that this method is more suitable for SMS messages, where grammar rules are not observed at all.

In the article [7], a new approach to text steganography is proposed, which is called Text generation, and which is mainly used for English characters. But this approach generates a meaningless message, but it doesn't take much time to encrypt and decrypt the plain text. Thanks to fast processing, this approach can be used in cloud computing.

The rather interesting Null Cipher method is given in the review [2] generates sentences in such a way that a specific position of a letter from each word (say, the second letter from each word), or a sentence, or a paragraph, or a page carries a secret message.

Words are also generated using the Missing Letter Puzzle method [2], where a list of words from 6 to 15 in length is used to hide data.

An interesting method proposed in [8] and called SSCE (Secret Steganography Code for Embedding), involves the use of articles in English. Here, inserting articles with non-specific nouns in English tries to hide the message in the text. Specifically, a text-based quantum steganography technique was proposed based on the use of indefinite articles (a) or (an) in combination with non-specific or non-specific nouns in English and the quantum gate truth table.

In [9], a steganographic method is proposed for hiding data in Microsoft Word documents using change tracking technology. Data injection is disguised in such a way that the stegodocument written by joint efforts contains obvious errors. Decoding is performed by tracking changes, giving the impression that the author of the document is correcting their mistakes. The change tracking information contained in the stegocument allows you to restore the original cover, a degenerate document, and therefore a secret message. Of course, this algorithm can be applied to all text files.

A rather interesting method is proposed in [10], which is called the Cricket Match Scorecard. In this method, the data is hidden in the cricket match results table, where a meaningless zero is pre-added before the number to represent bit 1, and leaving the number unchanged, which will represent bit 0.

In the methods named " List of words" a word list approach is used in text steganography to hide a secret message. The approach uses ASCII character values and hides the secret message without changing the cover file [11]. Similar "Spelling of Words " method" the review [2] is based on the fact that the representation of a word using the spelling in the UK will encode "0", while the spelling in the US will encode "1". Thus, the stego work created by this method will contain a mixture of two different spellings of words, but the same meaning, for example, color and Color. It should be noted that these methods can be easily applied in all languages.

To minimize distortion when replacing words, the "Adaptive Synonym Method" method is proposed in [12]. This method offers adaptive text steganography to minimize the distortion caused by synonym substitution and offers a hybrid function to detect this distortion.

Of course, the weakness of many methods based on synonyms is the semantic load of the text. Based on this, in [13] the proposed method differs from other methods since a synonym that has the same attribute and semantics as the target word is selected from WordNet sets.

The SRLEM (Synonym run-length encoding Method) method offers new linguistic steganography based on encoding the lengths of synonym series, which allows not to disturb the balance of synonym frequencies when replacing them, and this improves statistical undetectability[14].

The" Set Synonym Method " of text steganography based on synonym substitution uses all the different types of synonyms from the synonym set, which makes it possible to increase the security of the steganographic system [15].

In the modern world, where we observe various technologies based on neural networks, it is a promising approach for hiding data. Thus, the new linguistic stegosystem proposed in [16] is based on a neural network with long short term memory (LSTM) [17]. This method combines high output quality (i.e., the stegotext is very similar to natural language) with the highest capacity (the number of bits encrypted per word). And in the RNN (Recurrent Neural Networks (RNN-Stega)) method, linguistic steganography based on recurrent neural networks is proposed, which can automatically generate high-quality text covers based on a secret bitstream that needs to be hidden [18].

A very large group of methods under the preliminary name "Language feature", where the method is based on the features of the written language. As an introduction, here are some of them:

Arabic Unicode: It uses features of the Arabic script and presents a Unicode-based steganographic algorithm. The proposed algorithm allows you to make about 180 bits per 1 kilobyte with minimal changes to the associated letters without any changes in the size and shape of the text [19].

Arabic Method: Most text steganography methods are applied to English texts. However, there are several methods of text steganography applied to other languages [1], including Arabic texts with various modifications of these methods [20-

23]. A new steganographic algorithm for the Arabic text based on the features of the Arabic text is proposed. The focus is on a more secure algorithm and high media throughput. The algorithm can resist traditional attack methods because it makes minimal changes to the media text. The following articles provide an overview of various steganographic methods for the Arabic text [24, 25].

One of the methods is Moving the Dot in Characters [2], which uses the presence of dots. So in Persian, out of 32 letters, 18 have dots, and in Arabic, out of 28 letters, 15 have dots. These points are moved up or left untouched for secret bit embedding. A similar approach is implemented in the method Reversing the Symbol in Characters [2].

Chinese text: the Use of Chinese characters was suggested in [26], but here the embedding and extraction processes are too complex. This paper uses double Chinese characters. The proposed scheme significantly simplifies the embedding and extraction procedures and increases the efficiency of visual steganographic images

In [27], a new lossless steganographic scheme for traditional Chinese text files is proposed, which relies on various encoding standards for traditional Chinese characters, namely the Big5 and GBK standards. The proposed scheme uses this advantage to perform data hiding by correcting the encoding of coherent characters, and incoherent characters are also used to increase the implementation speed [27]. And in [28] it is proposed to use the properties of Chinese texts and the introduction of an erroneous character to increase the volume of hidden information.

Thai Method: In this method, we propose a new steganographic scheme, which covertly sends a secret message to several recipients through a stream of running short text messages in the Thai language. It is specified that four secret message bits can be embedded in a single Thai short text message. In principle, the proposed conversion methods can be applied to short text messages in any language [29].

Bengali Method: This method suggests using the fact that for several characters in the Bengali alphabet, there are several ways to represent the character in its equivalent Latin form using a phonetic keyboard layout. This feature is used to hide information bits in the Bengali text [30].

And here, along the way, we note that in the Uzbek language, where writing systems based on Cyrillic and Latin still coexist in parallel, you can take advantage of some features. Namely, some Latin letters have their Cyrillic counterpart, and they have different binary codes in encoding. For example, the letters "Aa", “Cc”, “Ee”, “Kk”, “Oo”, “Pp”, "Xx". By combining these letters, you can create amazing algorithms for hiding data in the Uzbek language.

III. PROPOSED APPROACH

Some scientists are developing systems that allow two users to exchange encrypted messages so that a passive adversary who reads the messages cannot determine either the original content of the messages or the fact that the messages are encrypted. The emphasis here is on linguistic steganography, the science of encoding a secret piece of information ("payload") into a piece of text that looks like a natural language ("stegotext"). Many of the results show that such methods demonstrate high output quality (i.e. stegotext is similar to natural language) with the highest capacity (the number of bits encrypted per word).

Based on this, traditional linguistic stegosystems are based on a modification of the existing accompanying text, for example, using the replacement of synonyms and/or replacement of paraphrases. The idea is to encode secret information in the transformation of the accompanying text, ideally without affecting its meaning or grammatical correctness.

Modifying the cover may introduce a syntactic and semantic error. To solve this problem, we can offer an alternative stegosystem, in which a person generates stegotext manually, thereby improving the language's naturalness at the cost of human effort.

Linguistic stegosystems involve embedding information in text containers using any natural languages. The main requirements for such methods are that they should not arouse suspicion. In other words, the entire structure of the language (grammar, syntax, semantics) must be preserved.

There are two main types of linguistic stegosystems:

- with the specified text;
- with selectable text.

The basic principle of building stegosystems of the 1st type is that there are sections that are evenly distributed in a certain area and according to the rules of embedding secret information, they are replaced by others. The main method of constructing such systems is to use absolute or relative synonyms.

An absolute synonym is a word or phrase that can be replaced by another word or phrase in any context without changing its meaning. Examples of synonyms-phrases (including abbreviations): UNIVERSITY-higher education institution, former Minister — ex-Minister, etc.

Relative synonyms are words or phrases that can replace each other (or not) depending on the context (the environment of these words or phrases).

Examples of relative synonyms: castle(building) - Palace, lock (door) - latch.

Absolute and relative synonyms for each language are collected in special dictionaries, for example: Oxford Collocation Dictionary for Students of English.

A method of linguistic stegosystems based on changing the order of words in a sentence is known.

Paraphrasing is very useful for many applications that usually involve deep linguistic changes to a sentence, such as generalization, text entry, and question answers, and usually require complex external resources, preprocessing, and a semantic thesaurus.

The above techniques are usually used to automatically extract paraphrases, which are complex linguistic changes to the original sentence, requiring complex tools and resources that are not available for many languages. Paraphrases are limited in number, and the necessary tools are usually subject-oriented and not particularly reliable, meaning they have limited scope and can only be applied to text that is limited in structure and subject area.

In this regard, in many works, you can find processes for automatically generating small paraphrases, and then using

them to provide a steganographic connection between two parties who want to exchange secret information. Note that this is a small paraphrase since this is a rather complex task that requires artificial intelligence tools. In such tasks, the bits of the secret message is embedded in the accompanying text in an unremarkable way, which does not cause the suspect to suspect the existence of hidden information. Such applications define two main goals of the presented approach. The first goal is to prepare as many correct paraphrases as possible for the original sentence. Steganographic safety depends to a large extent on the number and grammaticality of paraphrases of each sentence in the text.

The second goal is to use limited language resources as possible. This will first allow you to transfer the proposed methodology to other languages that have certain syntactic properties and are not necessarily equipped with complex linguistic resources. Paraphrases should not be complex syntactic or semantic changes that involve high-level linguistic resources.

Trends in recent years clearly show the desire of Uzbek citizens to use Internet resources that require them to exchange personal data for various purposes, such as identification. A kind of boom in Internet services can potentially lead to the General disclosure of personal data of users of resources that do not meet the requirements of information security. In this regard, the need for scientific research in this area is an obvious necessity. It is necessary to clearly define the tasks of information technologies, clearly represent the area of responsibility of the personal data operator, make changes or additions to the administrative code that establish the type and degree of responsibility for non-compliance with the requirements for the protection and confidentiality of personal data, as well as the processing of personal data without the consent of their owner.

Modern information resources are largely engaged in processing personal data, which is stored, in many cases, in the cloud, and then there is a protection of this data, such as contact information, phone numbers, and phone numbers of friends and relatives.

And here the problem arises of transmitting this data in such a way as to preserve their confidentiality and secrecy. Since each user must feel responsible so that their data does not fall to hackers. who may use this personal data for their selfish purposes.

The study of steganography methods allows us to conclude that these methods demonstrate the unique capabilities of modern information technologies of prisoners in ensuring information security. Based on this, we will consider one of the possible methods based on the Uzbek language for hiding a secret message.

Consider the following table, where each phrase is mapped to a sequence of bits. In this case, take the text from the official portal uza.uz:

Sentence	**Bits**
Men ayol zotini muqaddas bildim,	0000
U Momo Havoning buyuk avlodi.	0001
Men ayol zotida bir jahon ko'rdim,	0010
U barcha insonning ma'sum bunyodi.	0011
Inson- Onaning farzandi.	0100
Hayotga daho farzand yetkazib bergan ona.	0101
Ayol oilaninggina emas, jamiyatning ham gultoji.	0110
Qiz bor uyning fayzi bo'lak, deydilar,	0111
Jannat yo'li-qiz bor yo'lak, deydilar	1000
Insoniyat xotin-qizlarga hurmat bilan e'tibor qaratgan.	1001
Qur'oni Karimda ham ayollar ulug'lanadi.	1010
Qaysi soha yoki qaysi jabhada bo'lmasin, ayol bor.	1011
Ayollarning o'rni va ta'sirini hech kim bosolmaydi.	1100
Insonlar ayollarga adolatli munosabatda bo'lishi kerak.	1101
Hayotga daho farzand yetkazib bergan ona-qahramon.	1110
Ayol jamiyatning ham gultoji.	1111

The peculiarity of this set of sentences in the Uzbek language is that these sentences have a single subject area. In our version, the main idea of the subject area is love and respect for the Mother.

The algorithm proposed here is that if you need to pass the sequence "0100 1110 1001 0011", the transmitted text message will look like this:

Inson- Onaning farzandi.
Hayotga daho farzand yetkazib bergan ona-qahramon.
Insoniyat xotin-qizlarga hurmat bilan e'tibor qaratgan.
U barcha insonning ma'sum bunyodi.

Since the subject area has a single meaning, changing the sequence of sentences does not change the content. Thus, an outside observer will not notice the fact that the message is hidden in the text.

From the resulting text, the decoder program restores the original data in a deterministic and simple way: it takes the generated stegotext as input, examines one sentence at a time, finds its location in the General table, and restores the original bit block. And accordingly, the reverse process is obvious. That is, the text will give us the codes "0100 1110 1001 0011". Thus, we have a one-to-one conversion of codes to text and Vice versa.

Based on this algorithm, a program in the PascalABC programming language was created. Of course, this algorithm can be implemented in other programming languages, but here the additional complexity does not make it easier to understand the stegoalgorithm.

IV. EVALUATION

The proposed method, which uses saving the subject area and changing the sequence of sentences, has a limited scope. Based on this, in the future, the authors hope to use suffixes to hide information.

A suffix is a significant part of a word that is located after the root and usually serves to form new words. The study of this process has been morphology. As we know, morphology is a section of grammar that studies the forms of words and their semantics. One of the main features of the grammatical structure of the Uzbek language is the mandatory change in the form of most of the so-called significant (independent) words when forming phrases and sentences. When constructing units of syntax, the forms of words must be adapted to each other, for example, in the text in Uzbek "Sinfdoshimning tug'ilgan kuni" (which in English means" It's my classmate's birthday") if you drop the endings, the resulting text "Sinfdosh tug'ilgan kun" will not have an unambiguous meaning. Since various forms of suffixes can be observed in the Uzbek language that forms the meaning of a sentence, the author considers it important to study the hiding of messages using suffixes in the Uzbek language.

CONCLUSION

Thus, the method of artificial text generation that is best suited to the Uzbek language is very useful for many applications that usually involve deep linguistic changes to the sentence, such as generalization, text entry, and question answers, and usually require complex external resources, preprocessing, and a semantic thesaurus. In this work, the main approaches of the linguistic approach to the steganographic text were revealed. A specific example from the portal was presented uza.uz to demonstrate the steganographic model using the example of the Uzbek language.

It is shown that many methods based on other languages for hiding information suggest that such approaches can be used in the Uzbek lexicon.

So digital steganography, which is inspired by ancient secret communication techniques, is the art of hiding a secret message inside a covert environment in an undetectable way. This has not lost its relevance in our time.

The simple scheme of embedding a set of bits in the generated text proposed in this article is easily implemented in the Uzbek language. At the same time, the basic text was taken from official sources, which has a single subject area without a specific reference, for example, to time. This, of course, narrows down the wide application of this method. However, this is the first text steganography in the Uzbek language. In the future, these methods will be improved and generalized.

REFERENCES

[1] M. Hassan Shirali-Shahreza, Mohammad Shirali-Shahreza (2006) A new approach to Persian/Arabic text steganography. In: 5th IEEE/ACIS international conference on computer and information science and 1st IEEE/ACIS international workshop on component-based software engineering, software architecture, and reuse, pp 310-315.

[2] R. Bala Krishnan, Prasanth Kumar Thandra, M. Sai Baba (2017) An overview of text steganography. 4th International Conference on Signal Processing, Communications, and Networking (ICSCN -2017), March 16 - 18, 2017, Chennai, INDIA.

[3] M. Hassan Shirali-Shahreza, Mohammad Shirali-Shahreza (2008) A new synonym text steganography. In: International conference on intelligent information hiding and multimedia signal processing, pp. 1524-1526.

[4] Bender W, Gruhl D, Morimoto N, Lu A (1996) Techniques for data hiding. IBM Syst J 3(3&4): pp.313-336.

[5] Por LY, Ang TF, Delina B (2008) WhiteSteg-a new scheme in information hiding using text steganography. WSEAS Trans Comput 7(6): pp.735-745.

[6] Prasad R.S.R., Alla K (2011) A new approach to Telugu text steganography. ISWTA 2011 - 2011 IEEE Symposium on Wireless Technology and Applications, pp. 60-65.

[7] Tarun Kumar, Abhinav Pareek, Jyoti Kirori, and Maninder Singh Nehra (2014) Development of Crossover and Encryption Based Text Steganography (CEBTS) Technique. Lecture Notes in Electrical Engineering 298, India 2014,pp.103-111.

[8] Banerjee I., Bhattacharyya S., Sanyal G. (2011) An approach of quantum steganography through special SSCE code.World Academy of Science, Engineering and Technology. Volume 80, August 2011, pp.939-946.

[9] Tsung-Yuan Liu, Wen-Hsiang Tsai (2007) A new steganographic method for data hiding in Microsoft word documents by a change tracking technique. IEEE Trans Inf Forensics Secur 2(1): pp.24-30.

[10] Md Khairullah (2011) A Novel Text Steganography System in Cricket Match Scorecard. International Journal of Computer Applications (0975 – 8887) Volume 21– No.9, May 2011. pp. 43-47.

[11] Monika Agarwal (2013) An Efficient Dual Text Steganographic Approach: Hiding Data in a List of Words. Lecture Notes in Electrical Engineering 131, pp. 477-488

[12] Hu Huanhuan, Zuo Xin, Zhang Weiming, Yu Nenghai (2017) Adaptive Text Steganography by Exploring Statistical and Linguistical Distortion. 2017 IEEE Second International Conference on Data Science in Cyberspace (DSC) Proceedings - 2017 IEEE 2nd International Conference on Data Science in Cyberspace, DSC 2017, pp. 145-150.

[13] Lin Huo, Yu-Chuan Xiao (2017) Synonym substitution-based steganographic algorithm with vector distance of two-gram dependency collocations. 2017 2016 2nd IEEE International Conference on Computer and Communications, ICCC 2016 – Proceedings, pp. 2776-2780

[14] Lingyuan Xiang, Xinhui Wang, Chunfang Yang, Peng Liu (2017) A novel linguistic steganography based on synonym run-length encoding. 2017 IEICE Transactions on Information and Systems E100D(2), pp. 313-322.

[15] Qi C., Sun X., Xiang L (2013) A secure text steganography based on synonym substitution 2014. 2013 IEEE Conference Anthology, ANTHOLOGY 2013.

[16] Tina Fang, Martin Jaggi, Katerina Argyraki (2017) Generating Steganographic Text with LSTMs Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics- Student Research Workshop, pages 100-106 Vancouver, Canada, July 30 - August 4, 2017.

[17] Ilya Sutskever, James Martens, and Geoffrey E Hinton (2011) Generating text with recurrent neural networks. In Proceedings of the 28th International Conference on Machine Learning (ICML-11). pp. 1017-1024.

[18] Xiaoqing Guo, Ziming Chen, Yongfeng Huang, Yu-Jin Zhang (2019) RNN-Stega: Linguistic Steganography Based on Recurrent Neural Networks. IEEE Transactions on Information Forensics and Security.14(5),2019, pp. 1280-1295.

[19] Obeidat A.A. (2017) Arabic text steganography using Unicode of non-joined to right side letters. 2017 Journal of Computer Science.13(6), pp. 184-191.

[20] Aabed, M.A., Awaideh, S.M., Elshafei, A-R.M., and Gutub, A.A. (2007) 'Arabic diacritics based steganography', IEEE International Conference on Signal Processing and Communications, 2007, ICSPC 2007, IEEE, pp.756-759.

[21] Ahmadoh, E.M., and Gutub, A.A-A. (2015) 'Utilization of two diacritics for Arabic text steganography to enhance performance. Lecture Notes on Information Theory, Vol. 3, No. 1, pp.42-47.

[22] Odeh, A., Elleithy, K. and Faezipour, M. (2013) 'Steganography in Arabic text using Kashida variation algorithm (KVA)', Systems, Applications, and Technology Conference (LISAT), IEEE Long Island, IEEE, pp.1-6.

[23] Al-Nofaie, S., Fattani, M. and Gutub, A. (2016) 'Capacity improved Arabic text steganography technique utilizing 'Kashida' with whitespaces', The 3rd International Conference on Mathematical Sciences and Computer Engineering (ICMSCE2016), pp.38-44.

[24] Mohamed A.A. An improved algorithm for information hiding based on features of Arabic text: A Unicode approach 2014 Egyptian Informatics Journal. 15(2), pp. 79-87.

[25] Allah Ditta, Cai Yongquan, Muhammad Azeem, Khurram Gulzar Rana, and Haiyang Yu, Muhammad Qasim Memon. (2018) Information hiding: Arabic text steganography by using Unicode characters to hide secret data. Int. J. Electronic Security and Digital Forensics, Vol. 10, No. 1, 2018.

[26] Bin Feng, Zhi-Hui Wang, Duo Wang, Ching-Yun Chang, Ming-Chu Li (2014) A novel, reversible, Chinese text information hiding scheme based on lookalike traditional and simplified Chinese characters. KSII Transactions on Internet and Information Systems. 2014, 8(1), pp. 269-281.

[27] Qin C., Chang C.-C., Wang S.-T.,Chang C.-C. (2014) A novel lossless steganographic scheme for data hiding in traditional Chinese text files. 2014 Journal of Information Hiding and Multimedia Signal Processing. 5(3), pp. 534-545.

[28] Chia-Chen Lin, Li-Cheng Yang, Yi-Hui Chen. (2011) Data hiding scheme based on features of Chinese text. Proceedings - 7th International Conference on Intelligent Information Hiding and Multimedia Signal Processing, IIHMSP 2011, pp. 161-164.

[29] Samphaiboon N. Steganography via running short text messages. (2011) Multimedia Tools and Applications. 52(2-3), pp. 569-596.

[30] Md Khairullah (2019) A novel steganography method using transliteration of Bengali text. Journal of King Saud University - Computer and Information Sciences. Volume 31, Issue 3, July 2019, Pages 348-366.