Indonesian Journal of Electrical Engineering and Computer Science
Vol. 17, No. 2, February 2020, pp. 950~956
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v17.i2.pp950-956

Analysis review on linguistic steganalysis

Syiham Mohd Lokman[1], Aida Mustapha[2], Azizan Ismail[3], Roshidi Din[4]
[1,2,3]Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia, Malaysia
[4]School of Computing, College Arts and Sciences, Universiti Utara Malaysia, Malaysia

Article Info

Article history:

Received Jun 2, 2019
Revised Aug 5, 2019
Accepted Aug 23, 2019

Keywords:

Data hiding
Information security
Linguistic
Steganography

ABSTRACT

Steganography and steganalysis are essential topics for hiding information. Steganography is a technique of conceal secret messages by transmitting data through different domains. Its objective is to avoid discovery of secret messages. Steganalysis, meanwhile, is a method for locating the secret messages contained in the stego text. The objective of steganalysis is to find concealed data and to break the security of its domains. Steganalysis can be categorized into two types: targeted steganalysis and blind steganalysis. Steganography and steganalysis both have domains that are split into natural, also known as linguistic and digital media. There are three kinds of digital media which are picture, video and audio. The aim of this paper is to provide a survey on different linguistic steganalysis techniques used to find secret messages. This paper also highlighted two type of steganalysis method that are used in research and real practice. The discussion include findings on the most recent work on linguistic steganalysis techniques. This review hoped to help future research for improving and enhancing steganalytic capabilities.

Corresponding Author:

Aida Mustapha,
Faculty of Computer Science and Information Technology,
Universiti Tun Hussein Onn Malaysia,
86400 Parit Raja, Batu Pahat, Johor, Malaysia,
Email: aidam@uthm.edu.my

1. INTRODUCTION

Steganalysis is science's art which the objective of steganalysis is to differentiate possibly hidden data from sample information with slight or no knowledge about the steganography algorithm [1]. Nowadays, people utilized steganography in both legitimately and illegitimately methods [2]. For instance, residents may probably practice it to look after their privacy while terrorists or criminal utilize it for transferring ferocity information [3]. There are two kinds of steganalysis, known as targeted steganalysis and steganalysis that is blind or universal [4]. A specific steganographic embedding algorithm is intended for targeted steganalysis, while blind steganalysis is a universal method capable of detecting different kinds of steganography [5]. This is because blind steganalysis can identify a broader class of steganographic methods, but in comparison with the targeted steganalysis it is mostly less accurate [4]. According to USA TODAY, terrorists use steganography to communicate over the Internet and to manage cyber-crimes; which is deemed profitable to the attackers [6].

Secret information is concealed in pictures, videos, audio, and text, including in any digital media and even in a simpler form such as HyperText Markup Language (HTML), executable documents, and Extensible Markup Language (XML) [7]. However, this paper will focus on linguistic steganalysis, which is a method that uses simple text to discover possibly concealed data [8]. This paper is considering the implementation of linguistic steganalysis that is used to detect secret message within the text medium. The objective of this paper is to review different types of steganalysis method that consist views from targeted steganalysis and blind steganalysis. A particular steganographic embedding algorithm is intended for

targeted stegan analysis, while blind steganalysis is a universal method capable of detecting different kinds of steganography [9].

2. TARGETED STEGANALYSIS

To attack an exact type of steganography algorithm, targeted steganalysis, also known as specific steganalysis, is introduced. If integrated with a recognized algorithm, the steganalyst is aware of the embedding techniques and statistical trends of the stego text. This method of attack is very effective when the known embedding techniques are tested on text, but it may fail exponentially if the steganalyst does not understand the algorithm used to embed the information. The basic model of linguistic steganography and steganalysis model is shown in Figure 1.

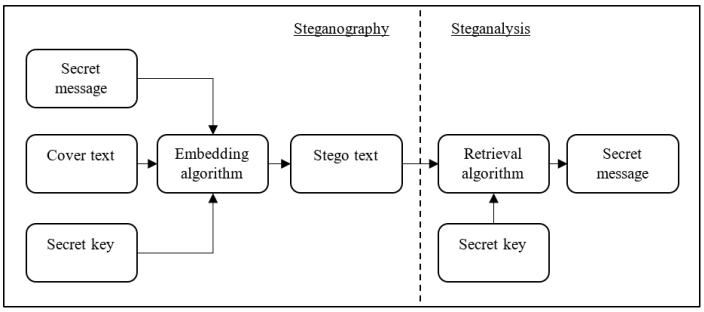

Figure 1. Basic model of linguistic steganography and linguistic steganalysis [10]

Based on the figure, if a steganalyst is alerted of the steganographic method, then the steganalyst must outline a steganalysis method that is capable of distinguishing concealed messages inserted with their steganography method only. The targeted or specific steganalysis plans utilize full information of a specific (targeted) steganographic algorithm and are created particularly to discover such a plan [11]. They are more consistent with enhanced execution in discovery than the universal plans [12]. Thus, numerous targeted steganalysis methods are considered as active as they evaluate the inserted message size.

There are a number of methods commonly used for targeted steganalysis in textual domain, such as distribution of space characters, distribution of first letters of words, font format, context information, source feature and immune mechanism, natural language watermarking, character substitution, evolution algorithm, and synonym frequency. Table 1 shows review categories in steganalysis.

Table 1. Categories of Steganalysis

Type	Schemes
Targeted Steganalysis	Distribution of space characters [13]
	Distribution of first letters of words [14]
	Font format [15]
	Context information [16]
	Natural language watermarking [17]
	Character substitution [18]
	Synonym frequency [12]
	Evolution algorithm [19]
	Source feature and immune mechanism [20]
Blind Steganalysis	[21]

3. TARGETED STEGANALYSIS METHOD

3.1 Distribution of Space Characters

In 2006, Sui Xi and Luo Hui suggested a technique of steganalysis based on space character distribution [13, 22, 23]. They are studying the space-steganography used in text papers. Steganalysis method detects stego-texts by varying statistical features between stego-texts and natural texts [24]. In the experiment, 10 natural texts are chosen and 5 stego-texts embedded by open space. Then their probabilities of p1 and p2 is calculated. From experiment, it indicates that p1 of every 10 natural texts is smaller than $\Delta 1$ and p2 smaller than $\Delta 2$. Meanwhile for 5 stego-texts, the p1 is larger than $\Delta 1$ and p2 larger than $\Delta 2$ [25, 26].

3.2 Distribution of First Letters of Words

Introduced a technique of steganalysis based on the allocation from English texts of first letters of phrases [14]. The method of steganalysis separates stego-texts and natural texts by statistical features. According this research, the assessment arbitrarily took 24 natural texts and then calculated the similarity parameters between the allocation of probabilities of first letters of words from the texts and the allocation of references [27, 28]. Next, 8 texts documents from natural text were arbitrarily picked and implanted with certain messages by utilizing the semantic algorithm [29, 30]. The similarity coefficients are then computed. From the test, natural text similarity coefficients are shown to be really small, with most of them being lower than 0.7, and the average value being only 0.6738. However, the similarity coefficients of stego-texts are very big and their average value is 0.8095, significantly higher than natural texts [14, 31].

3.3 Font Format

New algorithm for font format in text steganalysis was defined by [15]. It utilizes SVM to interpret the vector trait of font features to detect the presence of hidden information and then evaluate the length of the hidden information as stated by the font attribute value variants [32, 33]. The algorithm acquires a high level of accuracy in determining whether or not concealed information are available in font format steganography in ordered papers [18, 19, 31]. It is shown in the result of the experiment.

3.4 Context Information

This paper [16], evaluate one type of text steganography that uses synonymous substitution. Using context data, they attempt to differentiate between altered articles and original articles. The final decision taken by a classifier of the SVM (support vector machine) leads by assessment of suitability of words for their context, and then the suitability sequence of words. In the following to balance prevalent and rare phrases, IDF (inverse document frequency) is used to measure the suitability of phrases. With the assistance of Google, this system is assessed on the web rather than in a specific corpus. Experimental findings indicate 90.0 percent of classification precision.

3.5 Natural Language Watermarking

In 2008, [17] suggested synonymous replacement steganalysis based on watermarking in the natural language. The suggested method attempts to distinguish using context information between watermarked articles and unwatermarked articles [34]. For final determination of an SVM (support vector machine) classifier, it is shown the assessment of the reasonableness of phrases of the context and then the appropriate word arrangement. Inverse document frequency (IDF) is utilized to adjust periodic and unusual phrases to weigh the suitability of phrases with particular end goals. Using Google, this evaluation plan is on the internet rather than in a particular corpus. Experimental findings indicate accuracy of 90.0%, precision of 86.8% and recall rate of 82.5% [17].

3.6 Character Substitution

In order to detect concealed data using character substitution in texts, a modern technique of steganalysis is suggested. This is achieved by utilizing Support Vector Machine (SVM) to classify the distinctive input of the vector into SVM. Building a correct characteristic vector is the most significant phase of this detection algorithm. Under the precondition of uniform distribution of hidden bits to be encoded, the allocation of the characters used to hide data after the steganographic process has changed, so that the ratio of unusual characters to ordinary characters is dissimilar in cover texts and stego texts [18].

3.7 Synonym Frequency

A linguistic steganalysis method is suggested to identify synonymous steganography based on replacement that embeds concealed message into a text by replacing words with synonyms First, attribute pairs of synonyms are provided in order to reflect their position in an ordered set of synonyms in descending frequency order and the quantity of their synonyms. As a result of synonymous replacements, the number of pairs of high frequency attributes can be reduced while the number of pairs of low frequency attributes would be increased [12].

3.8 Evolution Algorithm

In 2012, [19] presented a steganalysis technique utilizing evolution algorithm approach. The study introduces another method of steganalysis, based on the Java Genetic Algorithms Package (JGAP) evolution algorithm approach, to recognize concealed texts in text steganalysis called the Evolution Detection Steganalysis System (EDSS). The result of the EDSS can be split into two groups in terms of fitness values that are great fitness and bad fitness [19].

3.9 Source Feature and Immune Mechanism

Another natural language detection algorithm was suggested by [20] relying on evaluating text features. The computer program has different characteristics of elements, including good couriers of text characteristics, and then, by searching for feature modifications, the hidden text is subdivided into Success-Stego-Text (SST) and False-Stego Text (FST) correspondingly, after the feature extraction has been arranged, the respective detector, so that the identification algorithm in the text is usually brief. Tests indicate that when the length of the text film reaches 3 K, both SR and NR exceed 95%, and consistency is greater [20].

4. BLIND STEGANALYSIS

Nobody knows the steganographic algorithm in blind steganalysis. Therefore, the designing of detector not relies upon on steganographic algorithm [21, 35]. Blind steganalysis is broadly used compared to the specific one because it is not relied upon steganographic algorithm[36]. Differentiate cover text from stego text is the main step in analyzing steganography. Analyzing text characteristics and searching abnormalities proof also included. The inserting operation will change the content of text and from normal text characteristics, it produces deviations. The step is therefore sensible and the most fundamental level of blind steganalysis is this evaluation [11].

Blind steganalysis stage can be extended to a better range called multi-class steganalysis. Multi-class steganalysis is essentially the same as the basic level if evaluated from a practical perspective. The distinction, however, is that multi-class steganalysis can differentiate the text into more groups of various kinds of stego text generated by various inserting methods. Consequently, the task of multi-class steganalysis is to trace the embedding algorithm used to generate a given stego text or interpret it as a cover text if no insertion is applied to it. There are two phases in blind steganalysis which known as feature extraction and classification [11], as shown in Figure 2.

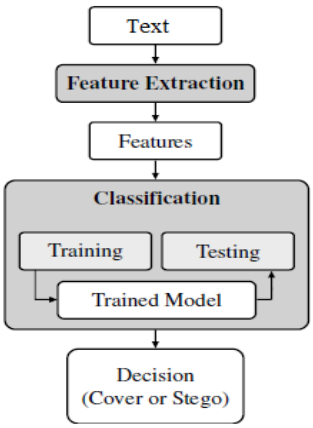

Figure 2. Standard blind steganalysis phases [37]

In paper [38] provided a technique capable of blindly distinguishing natural texts, machine translated texts and stego texts produced by steganography based on translation. This technique examined the features of the allocation of words in the same natural frequency area (NFZ).

5. COMPARATIVE ANALYSIS

Linguistic steganalysis has its own implementation performance especially in statistical technique. Table 2 identified the advantages and disadvantages of linguistic steganalysis in the last decade. Based on Table 2, the advantages of techniques are seen through the high recall in first letters of words distribution technique [14] and anatomize variance of statistical features of stego text and natural text in distribution of space character techniques [25]. Next, the semantic technique which is introduced by [16] is simple and effective variants. Meanwhile, the syntactic technique proposed by [15] has lower noise in text and lower missed detection ratio. Finally, the machine learning technique by [19] can support the text based document while the technique by Licai Zhu in 2017 is more specific on source feature.

Table 2. Comparative Analysis

No.	Years	Methods	Advantages	Disadvantages
1	2006	Distribution of Space Characters	Anatomize variance of statistical features of stego text and natural text	Complex algorithm
2	2006	First letters of words distribution	High recall and low error	Require much time
3	2007	Font format	Less hidden information, lower noise in text, lower missed detection ratio	Require big dataset
4	2008	Context Information	Simple and effective variants	Lack of vocabulary
5	2009	Natural Language Watermarking	High accuracy, recall rate and precision	Hard to gain similar synonym
6	2009	Character substitution using support vector machine (SVM)	Can implement in hard copy and soft copy	Only based on assumption
7	2012	Synonym frequency	High speed in detecting	Complex
8	2012	Evolution algorithm	Use EA approach to support the text based document	Complexity of computation
9	2017	Source features and immune mechanism	More specific on source features	Complex algorithm

The Figure 3 shows the disadvantages such as complex algorithm or complex computation that burden the linguistic steganalysis approach especially in distribution of space characters [25], evolution algorithm [19], as well as source features and immune mechanism [20].

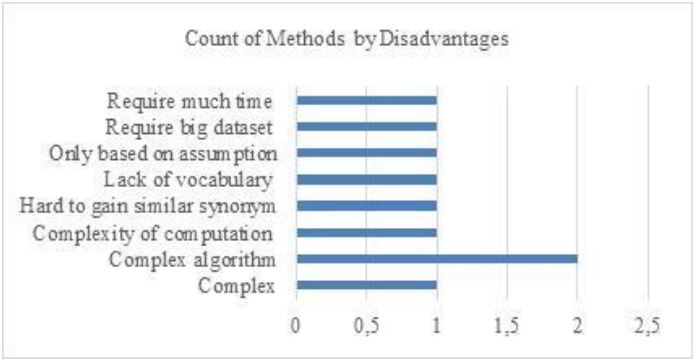

Figure 3. Count of methods by disadvantages

The obvious concern in implementing linguistic steganalysis is the requirement of big dataset and the high processing time. This issue was caught in font format technique proposed by [15] and first letters of words distribution by [14]. Finally, lack of vocabulary also posed a disadvantage in linguistic steganalysis like context information technique in [21].

6. CONCLUSION

This paper interpreted linguistic steganalysis used in recent years between 2006 until 2017. It review some methods of linguistic steganalysis that consists of two categories that are targeted and blind steganalysis. In general, only one or two techniques proposed each year. However, there is no technique proposed in linguistic steganalysis between 2013 until 2016. Most researchers focused on image steganalysis but there is only a few doing research for linguistic steganalysis. Furthermore, the study demonstrates that the steganalysis method have their own strengths and weakness. The most noticeable concern for linguistic steganalysis is complex algorithm. Hopefully, this study might empower for future research and to grow better steganalysis devices that can add to better execution.

ACKNOWLEDGEMENTS

This research is supported by the Postgraduate Research Grant Scheme (GPPS) Vot H335 from Universiti Tun Hussein Onn Malaysia.

REFERENCES

[1] S. Xin, L. Hui, L. Z. Zhong, "*Steganalysis method based on the distribution of first letters of words*," International Conference on Intelligent Information Hiding and Multimedia Signal, 2006.

[2] L. Xiang, X. Sun, G. Luo, C. Gan, "*Research on steganalysis for text steganography based on font format*," IEEE Third International Symposium on Information Assurance and Security, China, pp. 490-495, 2007.

[3] S. Pramanik, R. Singh, R. Ghosh, "A new encrypted method in image steganography," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 14(3), pp. 1412, 2019.

[4] M. Warkentin, E. Bekkering. M. Schmidt, "Steganography: Forensic, security, and legal issues," *Journal of Digital Forensics, Security and Law*, 2008.

[5] A. Hudson, "The sociology and psychology of terrorism: Who becomes a terrorist and why?," *Federal Research Division Library of Congress*, Available: https://fas.org/irp/threat/frd.html, 2009.

[6] R. Chhikara, L. Singh, "A review on digital image steganalysis techniques categorised by features extracted," *Journal of Engineering and Innovative Technology (IJEIT)*, vol. 3(4), pp. 203-213, 2013.

[7] X. Luo, D. Wang, P. Wang, F. Liu, "A review on blind detection for image steganography," *Signal Processing*, vol. 88(9), pp. 2138-2157, 2008.

[8] B. Li, J. He, J. Huang, Q. S. Yun, "A survey on image steganography and steganalysis," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 2(22), pp. 142-172, 2011.

[9] F. Z. Mansor, A. Mustapha, R. Din, A. Abas, S. Utama, "An antonym substitution-based model on linguistic steganography method," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 12(1), pp. 225, 2018.

[10] C. Chang, S. Clark, "Practical linguistic steganography using contextual synonym substitution and a novel vertex coding method," *Computational Linguistics*, vol. 40(2), pp. 403-448, 2014.

[11] N. Meghanathan, L. Nayak, "Steganalysis algorithms for detecting the hidden information in image, audio and video cover media," *International Journal of Network Security & Its Application*, vol. 2(1), pp. 44-55, 2010.

[12] H. Ge, M. Huang, Q. Wang, "*Steganography and steganalysis based on digital image*," 4th IEEE International Congress on Image and Signal Processing, China, 2011.

[13] I. Cox, M. Miller, J. Bloom, J. Fridrich, T. Kalker, "Digital watermarking and steganography," *Morgan Kaufmann Publishers*, 2008.

[14] L. Xiang, S. Sun, G. Luo, B. Xia, "Linguistic steganalysis using the features derived from synonym frequency," *Multimedia Tools and Applications*, vol. 71(3), pp. 1893-1911, 2012.

[15] K. Sakthisudhan, P. Prabhu, P. Thangaraj, C. Marimuthu, "*Dual steganography approach for secure data communication*," Procedia Engineering, vol. 38, pp. 412-417, 2012.

[16] Z. Yu, L. Huang, Z. Chen, L. Li, X. Zhao, Y. Zhu, "*Detection of synonym-substitution modified articles using context information*," IEEE Second International Conference on Future Generation Communication and Networking. China, pp. 134-139, 2008.

[17] Z. Yu, L. Huang, Z. Chen, L. Li, X. Zhao, Y. Zhu, "Steganalysis of synonym-substitution based natural language watermarking," *International Journal of Multimedia and Ubiquitous Engineering*, vol. 4(2), pp. 21-34, 2009.

[18] X. Zhao, L. Huang, L. Li, W. Yang, Z. Chen, Z. Yu, "Steganalysis on character substitution using support vector machine," *IEEE Second International Workshop on Knowledge Discovery and Data Mining*, Retrieved from https://doi:10.1109/wkdd.2009.105, 2009.

[19] R. Din, M. A. Samsudin, T. Muda, P. Lertkrai, A. Amphawan, "Fitness value based evolution algorithm approach for text steganalysis model," *International Journal of Mathematical Models and Methods in Applied Sciences*, vol. 7(5), pp. 551-558, 2013.

[20] L. Zhu, "Linguistic steganalysis approach base on source features of text and immune mechanism," *Computer and Information Science*, vol. 10(4), pp. 60, 2017.

[21] J. Lenti, "Steganographic methods," *Periodica Polytechnica Ser El Eng*, vol. 44(3-4), pp. 249-258, 2000.

[22] J. Aparna, S. Ayyappan, "*Image watermarking using diffie hellman key exchange algorithm*," Procedia Computer Science, vol. 46, pp. 1684-1691, 2015.

[23] D. Walia, P. Jain, Navdeep, "An analysis of LSB & DCT based steganography," *Global Journal of Computer Science and Technology*, vol. 10(1), pp. 4-8, 2010.

[24] C. Jensen, "Fingerprinting text in logical markup languages," *Springer*, Berlin, Heidelberg, Retrieved from https://link.springer.com/chapter/10.1007/3-540-45439-X_30, 2001.

[25] S. Guang, "*Steganalysis method based on the distribution of space characters*," IEEE International Conference on Communications, Circuits and Systems, pp. 54-56, 2006.

[26] E. El-Kwae, "New approach for hiding multimedia information in text. electronic imaging," *Society of Photo-Optical Instrumentation Engineers (SPIE)*, Retrieved from https://doi.org/10.1117/12.465269, 2002.

[27] J. Zhou, Z. Yang, X. Niu, Y. Yang, "Research on the detecting algorithm of text document information hiding," *Journal on Communications*, vol. 25(12), pp. 97-101, 2004.

[28] G. Luo, X. Sun, Y. Liu, "Research on steganalysis of stegotext based on noise detecting," *Journal of Hunan University (Natural Sciences)*, vol. 32(6), pp. 181-184, 2005.

[29] S. Xin, L. Hui, L. Z. Zhong, "*Steganalysis method based on the distribution of first letters of words*," International Conference on Intelligent Information Hiding and Multimedia Signal, 2006.

[30] J. Brassil, S. Low, N. Maxemchuk, "*Copyright protection for the electronic distribution of text documents*," Proceedings of the IEEE, vol. 87(7), pp. 1181-1196, 1999.

[31] M. Taskiran, U. Topkara, M. J. Topkara, E. Delp, "Attacks on lexical natural language steganography systems," *SPIE, Security, Steganography, and Watermarking of Multimedia Contents VIII*, Indiana, pp. 97-105, 2006.

[32] C. Yang, F. Liu, X. Luo, Y. Zeng, "Pixel group trace model-based quantitative steganalysis for multiple least-significant bits steganography," *IEEE Transactions on Information Forensics and Security*, vol. 8(1), pp. 216-228, 2013.
[33] J. Fridrich, J. Kodovsky, "Rich models for steganalysis of digital images," *IEEE Transactions on Information Forensics and Security*, vol. 7(3), pp. 868-882, 2012.
[34] M. H. Shirali-Shahreza, M. Shirali-Shahreza, "*A new approach to Persian/Arabic text steganography*," 5th IEEE/ACIS International Conference on Computer and Information Science and 1st IEEE/ACIS, USA, pp. 310-315, 2006.
[35] F. Petitcolas, R. Anderson, M. Kuhn, "*Information hiding-a survey*," Proceedings of the IEEE, vol. 87(7), pp. 1062-1078, 1999.
[36] C. Cachin, "An information-theoretic model for steganography," *Information and Computation*, vol. 192(1), pp. 41-56, 2004.
[37] M. P´erez, "Blind steganalysis method for detection of hidden information in images," *National Institute for Astrophysics, Optics and Electronics*, 2013.
[38] Z. Chen, L Huang, P. Meng, W. Yang, H. Miao, "*Blind linguistic steganalysis against translation based steganography*," 9th International Workshop IWDW: International Workshop on Digital Watermarking, Springer; 2010.

BIOGRAPHIES OF AUTHORS

Syiham Mohd Lokman is a student at the Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia. She received her B.Comp Sc from Universiti Tun Hussein Onn Malaysia in 2016. Currently, she is currently pursuing for Master degree in Steganalysis.

Aida Mustapha received the B.Sc. degree in Computer Science from Michigan Technological University and the M.IT degree in Computer Science from UKM, Malaysia in 1998 and 2004, respectively. She received her Ph.D. in Artificial Intelligence focusing on dialogue systems. She is currently an active researcher in the area of Computational Linguistics, Soft Computing, Data Mining, and Agent-based Systems.

Azizan Ismail is a lecturer at the Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia. His current research interests lie in Information, Computer Security and Communications Technology.

Roshidi Din is an Associate Professor at the School of Computing (SoC), UUM College of Arts and Sciences (CAS), Universiti Utara Malaysia (UUM). His current research interests lie in Information Security, Steganography and Steganalysis.