Generative Linguistic Steganography: A Comprehensive Review

Lingyun Xiang[1,2*], Rong Wang[1], Zhongliang Yang[3], and Yuling Liu[4]
[1] School of Computer and Communication Engineering, Changsha University of Science and Technology, Changsha, 410114 China
[e-mail: xiangly210@163.com, wr@stu.csust.edu.cn]
[2] Changsha Hunan Provincial Key Laboratory of Intelligent Processing of Big Data on Transportation, Changsha University of Science and Technology, Changsha, 410114, China
[3] Department of Electronic Engineering, Tsinghua University, Beijing, 100084 China
[e-mail: yangzl15@tsinghua.org.cn]
[4]College of Computer Science and Electronic Engineering, Hunan University, Changsha, 410082 China
[e-mail: yuling_liu@126.com]
*Corresponding author: Lingyun Xiang

Received October 9, 2021; accepted March 14, 2022; published March 31, 2022

Abstract

Text steganography is one of the most imminent and promising research interests in the information security field. With the unprecedented success of the neural network and natural language processing (NLP), the last years have seen a surge of research on generative linguistic steganography (GLS). This paper provides a thorough and comprehensive review to summarize the existing key contributions, and creates a novel taxonomy for GLS according to NLP techniques and steganographic encoding algorithm, then summarizes the characteristics of generative linguistic steganographic methods properly to analyze the relationship and difference between each type of them. Meanwhile, this paper also comprehensively introduces and analyzes several evaluation metrics to evaluate the performance of GLS from diverse perspective. Finally, this paper concludes the future research work, which is more conducive to the follow-up research and innovation of researchers.

Keywords: Steganography, steganalysis, text steganography, generative linguistic steganography, text generation.

This project is supported by National Natural Science Foundation of China under Grant 61972057, 62172059 and 61872134, Hunan Provincial Natural Science Foundation of China under Grant 2019JJ50655 and 2020JJ4624, Scientific Research Fund of Hunan Provincial Education Department of China under Grant 21A0211 and 19A020.

http://doi.org/10.3837/tiis.2022.03.013 **ISSN : 1976-7277**

1. Introduction

The massive amount of multimedia data in the Internet provides rich hiding carriers for information hiding, such as image [1], audio [2], video [3], text [4] and so on, therefore, steganography has become a hot research subject in the information security field.

As one of the most extensive carriers of information transmission, text has great research value and practical significance, therefore, a large number of text steganographic methods have been proposed [5]. Text steganography is a kind of covert communication means by hiding secret information in text without being aware of its existence, which can be mainly divided into three categories: text-modification-based, the text-selection-based and the text-generation-based steganography. The text-modification-based steganography (TMS) mainly hides secret information by modifying text characteristics or content, such as changing tracking [6], synonym substitution [7], morphosyntactic alterations [8], etc. The text-selection-based steganography, also called coverless linguistic steganography (CLS) [9], constructs large-scale text corpus and then makes great efforts to select suitable steganographic text (stego text) as carrier based on keywords and labels. The text-modification-based steganography is not secure enough, it is easy to be discovered the existence of the secret information by using steganalysis methods [10]. Moreover, the hidden capacity of the text-modification-based and CLS is extremely limited. The text-generation-based steganography, also known as generative linguistic steganography (GLS) [11], does not require cover in advance, but then generates stego text based on the secret information by NLP techniques. This type of the steganography is no upper limit to the length of the generated stego text, since there are more locations available for embedding secret information to achieve larger hidden capacity, however, the quality of the long generated stego text is poor early. With the development of deep learning-based text-generation techniques, researchers have successively proposed many generative linguistic steganographic methods, which gradually alleviate the bottleneck of GLS.

This paper focuses on summarizing the development of GLS, and the rest of this paper is organized as following sections: Section 2 briefly introduces the GLS. And we conclude evaluation metrics in section 3. Then existing generative linguistic steganographic methods are analyzed and discussed from the perspectives of text generation, encoding algorithm and evaluation metrics in section 4. Finally, this paper is discussed with several directions worthy of further development in section 5.

2. Background of GLS

The steganography system can be described vividly by the Simmons' "Prisoners Problem" [12]: Prisoners *Alice* and *Bob* attempt to stealthily discuss an "escape plan" which cannot be perceived by warden *Eve*, so they intend to hide the true information in normal message. We can mathematically model the scenario as Equation (1) [13]:

$$\begin{aligned} &Emb: C \times K \times M \rightarrow S,\ f(c, k_A, m) = s \\ &Ext: S \times K \rightarrow M,\ g(s, k_B) = m \end{aligned} \tag{1}$$

Alice needs to covertly transmit secret messages m in secret message space M to *Bob*. First, *Alice* selects a cover c from the cover space C, and then maps c to the steganographic carriers s in the hidden space S by mapping function f under the guidance of key k_A in the key space K. The receiver *Bob* extracts secret messages m from s by using extraction function

g under the guidance of key k_B in the key space K.

Based on the above the model and theory, we summarize general framework of a steganography system as shown in **Fig. 1**, which is divided into the following three stages:

(1) **Information hiding**: *Alice* secretly embeds true secret messages into the carriers by hiding algorithm to form the steganographic carriers.

(2) **Information transmission**: *Alice* transmits the steganographic carriers to the receiver *Bob* via the public channel, which needs to have certain ability of anti-steganalysis to ensure the security of the secret information.

(3) **Information extraction**: *Bob* extracts secret messages from the steganographic carriers received from *Alice* through extraction algorithm.

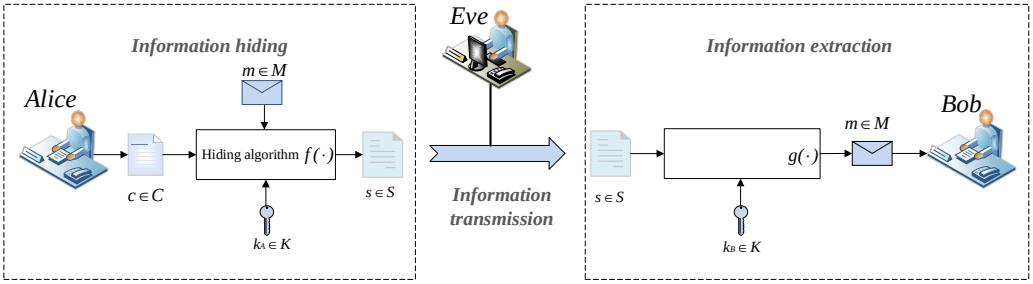


Fig. 1. The general framework of steganography

Referring to the above framework, existing research mainly designs steganography from three aspects: steganography by cover modification, cover selection and generation [14]. At the same time, based on the type of C or S, steganography has also divided into different types, such as image, text, video and audio steganography. Text steganography is a mechanism to embed secret information covertly by using the redundancy of text in format, structure, language or generation processing. At present, text steganography is mainly divided into the following three categories:

(1) Text-modification-based steganography (TMS): This type of steganography realizes the embedding of secret messages by slightly modifying the characteristics of the cover, and it mainly maintains the high similarity between the cover and stego text as much as possible to ensure imperceptibility.

(2) Coverless linguistic steganography (CLS): This type of steganography achieves information hiding by located and retrieved texts from original text-datasets as stego texts driven by secret messages, rather than modifying the original texts, and its imperceptibility and security are largely correlated with the size of datasets.

(3) Generative linguistic steganography (GLS): This type of steganography does not give cover in advance, but generates stego text according to secret messages with different text generation methods and encoding algorithms. GLS mainly ensure perceptual, statistical and cognitive-imperceptibility of stego text by considering the word with the best local and global structure and semantic information as the next candidate as far as possible.

3. Evaluation Metrics for GLS

GLS is a special type of text steganography, which generates corresponding steganographic texts on the basis of secret messages, rather than giving cover in advance, namely $C = \varnothing$. The framework of GLS is shown in **Fig. 2**, with the idea of general steganography framework, we

also divide GLS into three parts. The information hiding and extraction part mainly include two processes: text generation, steganographic encoding or decoding, the text generation process refers to the generated model well trained by text corpus and text generation technology, the steganographic encoding process is to choose the appropriate encoding algorithm according to the secret information and the steganographic decoding process is the reverse operation, these two processes can be performed either independently [13] or simultaneously [15] to generate stego text for the purpose of hiding secret messages. Although GLS does not give a text carrier in advance, some methods need to generate stego text in the term of the constraint information [16]. GLS is designed to improve the hidden capacity of text steganography, and it can resist the previous various steganalysis methods to ensure the imperceptibility of the secret messages.

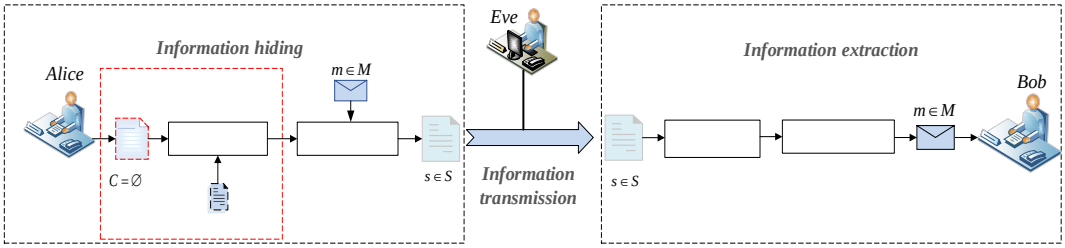


Fig. 2. The framework of GLS

GLS is an important branch of information hiding, which aims to hide the existence of secret information into the automatically generated stego text. GLS's goal is different from that of watermarking. Generally speaking, most researchers always employ payload size (hidden capacity) and security to evaluate the performance of the steganographic methods, while the robustness is used for watermarking to a greater extent [17]. Thus, considering the particularity of GLS, this paper only adopts the security and hidden capacity as evaluation metrics to measure the performance of GLS, we focus on the perceptual, statistical and cognitive-imperceptibility [18] to quantify the security, which means eavesdropper cannot detect the existence of steganography by any means [17].

3.1 Hidden Capacity

Hidden capacity refers to the number of secret bits that can be carried in stego text, researchers mainly use the embedding rate (ER) to measure it [19]. ER mainly has the following two calculation methods (the calculation method of Chinese special genre is not considered in this paper):

(1) ER of steganography embedded in words is calculated as Equation (2) [20]:

$$ER = \frac{1}{N}\sum_{i=1}^{N}\frac{(L_i - 1)\cdot k}{B(s_i)} \tag{2}$$

where N is the number of generated steganographic sentences, L_i is the length of the i-th sentence, k represents the number of bits embedded in each word, $B(s_i)$ is the number of bits occupied by the i-th sentence.

(2) ER of steganography embedded in characters is calculated as Equation (3) [21]:

$$ER = \frac{S}{L} \tag{3}$$

S is the number of secret bits embedded, and L is the bit length of the whole stego text.

3.2 Perceptual-imperceptibility

Perceptual-imperceptibility [18] requires that the generated stego text is semantically complete and natural enough so that it is not easy to be recognized by human monitors. For GLS, the higher the quality of stego text (complete semantics, correct grammar and sufficient naturalness), the more it can meet the perceptual-imperceptibility. We mainly measure the quality of stego text from the perspectives of subjective and objective [22], subjective evaluation refers to manual evaluation of text quality [13] or judging whether it is normal text [23], while objective evaluation uses indicators such as perplexity and mutual information (MI) to evaluate the quality of stego text.

(1) Perplexity [21]: Perplexity is the degree of certainty of whether the trained and tested texts contain secret messages, which is a standard measure of quality of stego text in NLP.
(2) MI [24]: MI is the measure of mutual dependence between two words, it has been widely used in NLP, for example, using MI to help realize song generation [25].

The quality of English texts in general genres generated by GLS is measured by *perplexity* (perplexity is inversely proportional to the quality of the text), while for Chinese poetry, MI (MI is proportional to the quality of the poem) is used to evaluate the quality of Chinese characters generated by GLS.

3.3 Statistical-imperceptibility

Statistical-imperceptibility requires that the statistical distribution difference between the generated stego text and cover is small enough to resist the existing statistical steganalysis methods. For GLS, the closer the statistical distribution between cover and stego text, the more it can satisfy the statistical-imperceptibility [18]. Kullback–Leibler divergence (KLD) and Jensen–Shannon divergence (JSD) [21] can evaluate statistical-imperceptibility by calculating the difference of probability distribution between cover and stego text, it also uses the ability of anti-steganalysis to evaluate statistical-imperceptibility.

(1) KLD: KLD, also called relative entropy, is one of indicators that measures difference of the overall distribution between steganographic sentences and normal sentences.
(2) JSD: JSD, a variant of KLD, is one of indicators that measures the similarity of the probability distribution between cover and stego text.
(3) Anti-steganalysis ability: Anti-steganalysis ability is a measure used to measure whether certain steganography can resist a variety of steganalysis methods. This problem is regarded as a binary classification problem to distinguish cover and stego text, so we use accuracy (Acc), recall (R), precision (P) and F1-score [19] to evaluate that.

A large number of researchers use KLD to measure the difference between cover and stego text (KLD is proportional to the difference of statistical distribution between the two categories texts). But KLD is not a strict distance measurement, if there is great difference in probability distribution between the two, that will be more suitable to use JSD (JSD is proportional to the difference). Acc is the simplest and most intuitive evaluation metric for classification, but different indicators should be selected according to the corresponding application scenarios, for example, when P and R need to be considered at the same time, we can use F1-score (the higher the value, the more robust the classification model, and the stronger the ability of anti-steganalysis) to evaluate steganography.

3.4 Cognitive-imperceptibility

Cognitive-imperceptibility, which was first defined by Yang and Xiang [18], requires the semantic expression of the generated stego text meets the contextual semantic constraints, so that there will be no cognitive inconsistencies in a specific context of use. GLS ensures semantic expression by controlling the contextual semantic constraints of stego text, so as to meet the cognitive-imperceptibility to a higher extent. Cognitive-imperceptibility uses bilingual evaluation understudy (BLEU) [26], recall-oriented understanding for gisting evaluation-longest common subsequence (ROUGH-L) [27], metric for evaluation of translation with explicit ordering (METEOR) [28] and consensus-based image description evaluation (CIDEr) [29] and other indicators to measure cognitive-imperceptibility, mainly by evaluating the semantic relevance of the stego text and the reference text.

(1) BLEU: BLEU is a measurement used to analyze the co-occurrence of n-grams between stego text and the reference text.
(2) ROUGE-L: ROUGE-L is the sub-series method of ROUGE, and it is a statistical method based on the co-occurrence precision and recall rate of the longest common sequences (LCS).
(3) METEOR: METEOR is an evaluation method based on single-precision weighted harmonic mean and single word recall, it uses three modules to count the co-occurrence times: extract, porter stem and WordNet synonym module.
(4) CIDEr: CIDEr regards each sentence as a document, calculates the cosine angle of its TF-IDF vector to obtain the similarity between the stego text and the reference. This measurement proposed in computer vision field for generation of image summarization.

BLEU is based on the n-gram method, but the larger the n, the worse the matching of the method at the sentence-level. ROUGE-L can reflect the sentence-level word order, wheras it only calculates a longest subsequence while ignores the influence of other candidate subsequences. Compare with BLEU, METEOR also is sensitive to length, however, it not only considers both the Acc and R of the entire corpus, but also considers the impact of synonyms on semantics. CIDEr combines BLEU and vector space, and its advantage is that different n-grams have different weights with different TF-IDF.

4. Classification and Analysis of the Existing GLS

GLS is mainly based on text generation techniques, which is an important research direction in the field of NLP, mainly including three steps: content planning, sentence planning and text realization to solve the problem of “what to say”, “how to say” and “how to achieve text generation”. Text generation tasks are implemented in different stages using different generation technologies or models: the template generation, pattern generation, phrase or planning extension, attribute feature generation technologies, the Markov model and neural network model [30]. Referring to the text generation methods above, this paper divides GLS into four categories according to the different methods used in text generation process, which will elaborate in detail as followed.

4.1 Template-generation-based Unconstrained GLS

Template-generation-based unconstrained GLS is to achieve the embedding of secret information under the guidance of specific sentence pattern, template structure and dictionary encoding tables. This type of steganography follows template generation technology in the task of text generation, firstly summarizes the fixed template-set from a large number of text

corpus, guides text generation with template, and then controls the choice of filling content during the template filling process on the basis of the confidential information, so as to generate stego text to realize information hiding. Based on whether the text genre is general or not, we subdivided this type of steganography into two categories: context-free-grammar (CFG)-based and specific-genre-template-based GLS.

4.1.1 CFG-based GLS

CFG-based GLS uses the principle of CFG to design and construct the corresponding CFG rules in advance to guide the generation of sentences. When selecting the next position word in each time step, there are usually multiple options that can be encoded, and then select different words to form the final stego text according to the secret information to be embedded.

The following is an example of generating sentences from CFG, assuming that the summarized ruleset is as follows, where S is the most initial state, A and B both are transfer state, a and b are terminal state. For $S \to AB|BA$, "AB" can be encoded as "0" and "BA" as "1", if the binary bit of "0" is needed to embed, we can select "AB".

$$S \to AB|BA$$

$$A \to Aaa|AA|Aab|ab$$

$$B \to Bbb|BA|Bba|ba$$

$$S \to AB \to AaaB \to AabaaB \to ababaaB \to ababaaBbb \to ababaababb$$

$$S \to BA \to Bab \to baab$$

Based on the idea above, Wayner [31] firstly proposed the mimic function based on CFG, which learns the statistical distribution of each character in the training sample and the rules for CFG in advance, and then constructs a huffman tree to encode the waiting options selected to form sentences in each time step. However, their method only considers the statistical distribution characteristics of the characters, Chapman et al. [15] constructs syntactic templates with the help of the CFG to simulate writing style to generate texts closer to natural language, and then encoded replaceable candidate words, but in fact the stego text generated by the method does not consider semantics and syntax.

4.1.2 Specific-genre-template-based GLS

Specific-genre-template-based GLS summarizes and designs specific metrical intonation template from the original poetry corpus in advance (the specific genre template is concentrated in the poetry template), and then concludes the corresponding dictionary as candidate words according to the characteristics of each location, selects words can be encoded by coding algorithm to fill the corresponding position of the template to generate poetry at each time step.

Based on the above ideas, Yu et al. [32] uses firstly ci-poetry of the song dynasty as stego text, this steganography resets dictionary with hash function, encodes candidates of each position one by one, finally selects words to generate stego text according to secret messages.

4.1.3 Analysis of Template-generation-based Unconstrained GLS

Template-generation-based unconstrained GLS is relatively simple and widely used in the early stage. The grammar rule is constructed by CFG-based GLS are general, so the stego text is difficult to meet certain grammar and semantic rules, lacking of logic between paragraphs and sentences, which resulting in poor text quality. The poetry generated based on specific-genre-template GLS is directly pieced together, which cannot meet certain emotions and

artistic conception to a great extent, and does not have corresponding literary value. This paper focuses on the ER and statistical-imperceptibility of unconstrained GLS based on template generation.

ER: This part compares the average ER of 10 samples generated by two steganographic methods Nicetext [15] (less than 2%) and Ci-stega [32] (about 16.5%). On the premise of the same amount of secret information, the length of the text generated by Nicetext is longer than that of Ci-stega, so ER based on Nicetext is relatively low.

Statistical-imperceptibility: **Table 1** measures anti-steganalysis ability of stego text mixed three steganographic methods (Nicetext, texto and Markov based) [33] by the value of the detection accuracy. As the size of stego text becomes larger, its anti-steganalysis ability becomes weaker and easier to be detected.

Table 1. Comparison of *Detection accuracy* on template-generation-based unconstrained GLS

Steganalysis	Size of text	Detection accuracy of stego text (%)	Detection accuracy of cover (%)
Method in [33]	1KB	**88.5**	86.5
	2KB	**93.5**	90.5
	3KB	**96.0**	95.5
	4KB	98.0	**100.0**
	5KB	99.0	**100.0**

4.2 Markov-based Unconstrained GLS

Markov-based unconstrained GLS follows the nonparametric Markov model used to text generation task in the NLP field, which firstly obtains the transition probability distribution between words by counting the words in the corpus, then predicts the next word in the sentence through the current n words (n is the order of the Markov model, which is the first-order by default), repeat the process to generate sentences. This paper summarizes the steps of Markov-based unconstrained GLS as follows:

(1) Dictionary construction: determining text corpus firstly, then building dictionary library for selection of later word;

(2) Status determination: choosing n consecutive words from text as the starting state, the last n words of the already generated sequence constitute the current state;

(3) Word encoding: for generating the next word, viewing firstly the corpus to find the candidate words followed by this n words and encoding them according on the number of candidates;

(4) Stego text generation: selecting the matching next word according to the desired embedded bit stream information, then repeating step (2) (3) until the secret information is completely hidden and the stego text is generated.

By analyzing the generation process of above Markov-based unconstrained GLS, this paper subdivides it into three types on the basis of steganographic encoding algorithm: the GLS based on status selection, status transformation and interval division.

4.2.1 Status Selection Based Steganographic Encoding Algorithm

Encoding algorithm based on status selection means directly encoding multiple candidate word items at a given moment. Candidates with the initial state of “it is” are “a”, “not”, “an”

and “me”, and encoded as “00”, “01”, “10”, “11” respectively, if we should embed “10”, the sentence becomes “it is an” at this time, and so on to generate stego text.

Dai et al. [34] proposed the Markov-based unconstrained GLS for the first time using the encoding algorithm mentioned above, and then the team encrypts the secret information with DES algorithm before embedding it covertly [35]. After that, Luo et al. [24] presented a steganographic method named Ci-stega based on poetry, which needs to select the words that meet the requirements according to the tone pattern and put them into the stacklist as the candidate thesaurus to be encoded. Wu et al. [36] introduces half frequency crossover rule, namely, if the next candidate word appears in the next state of the 3-gram and 4-gram Markov state transition diagrams, they take the average of the two probabilities as the probability value, and then embed secret information in the encoding algorithm based on state selection. Later, Wu et al. [37] changed the rule that uses Markov transition diagrams based on 3-gram and 4-gram alternately. Yang et al. [19] dynamically builds the candidate thesaurus of each time step according to the state transition diagram and encodes them by Huffman coding.

4.2.2 Status Transformation Based Steganographic Encoding Algorithm

Encoding algorithm based on status transformation refers to the related encoding algorithm on the entire Markov state transition diagram to realize the coding.

In order to ensure that the phrases generated at each step are the best choices, that is, to generate the best quality text finally, Wu et al. [20] proposed a steganographic method based on state transition-binary sequence (STBS-stega), which creates multiple STBS diagrams and numbered in turn. Among these STBS diagrams, different diagrams adopt different coding values for different word probability distribution paths, and encode the sequence of diagram using DES encryption which can select the best word selecting and generate the best quality stego text.

4.2.3 Interval Division Based Steganographic Encoding Algorithm

Encoding algorithm based on interval division firstly determines the mapping interval according to the fixed packet length of secret information, and then divides the mapping interval by calculating the probability ratio until the fixed-bit secret bitstream can be uniquely represented by the character sequence. For example, the binary bit stream that needs to be embedded is “100001”, the fixed packet length is 3 and the mapping interval is [0,7], which is converted into the binary string containing 000 to 111, there are two possibilities with the same transition probability from the initial state to the next state, s1 and s2, so the division range [0,3] corresponds to s1, [4,7] corresponds to s2, and so on.

Moraldo [38] designed a method combines the Markov model and the encoding algorithm. On this basis, Shniperov et al. [39] proposed a steganographic method based on second-order and third-order Markov models, and it can also process Russian text.

4.2.4 Analysis of Markov-based Unconstrained GLS

Compared with template-generation-based unconstrained GLS, Markov-based unconstrained GLS generates better text quality and can guarantee the security of secret information better. However, two approximate estimates are required when calculating the conditional probability for the k-order Markov model: one is only the first k states are considered when predicting the n-th candidate word; the other is Markov model directly regards word frequency as probability distribution. Therefore, Markov model cannot obtain the optimal conditional probability estimation and ideal language model. With the increase of secret messages need to be

embedded, the probability distribution of selected words becomes lower, so the quality of stego text becomes worse. This paper mainly evaluates from ER and statistical-imperceptibility of Markov-based unconstrained GLS.

ER: This part hlights the comparision of hidden capacity of several steganographic methods based on this type [19] [20] [36] [37], and their ER are 2.78%, 2.85%, 2.71%, 7.34% (bpw=3) respectively. The Markov-based GLS with huffman coding [19] can adjust dynamically ER to 7.34% (when bpw=3bits/word), which is the highest in these methods.

Statistical-imperceptibility: **Table 2** shows the comparison of perplexity of stego text generated by five different methods. Their difference lies in the difference in steganographic encoding, obviously, text steganography based on the encoding of state transition diagrams [20] aims to ensure that the phrases generated at each step are the best collocation, so the quality of the text is also better. **Table 3** show the comparison of detection accuracy to evaluate the ability of anti-steganalysis for different methods, which shows that STBS-stega [20] can better resist the attacks of steganalysis method in [40].

Table 2. Comparison of *perplexity* on Markov-based unconstrained GLS

Reference	IMDB[41]	News[43]
Method in [35]	418.70±105.32	470.54±122.73
Method in [38]	161.92±143.31	175.42±126.28
Method in [37]	52.05±35.80	20.52±13.98
Method in [36]	15.97±7.57	17.41±8.91
Method in [20]	13.34±9.90	**12.89±8.75**

Table 3. Comparison of *detection accuracy* on Markov-based unconstrained GLS

Reference	bpw (bits/word)	Text steganalysis method	IMDB[41]	Twitter[42]	News[43]
Method in [35]	-	Method in [40]	0.632	0.693	0.690
Method in [38]	-		0.665	0.678	0.723
Method in [19]	-		0.530	0.560	0.560
Method in [20]	3		**0.502**	**0.515**	**0.525**
	4		**0.515**	**0.510**	**0.522**
	5		**0.532**	**0.505**	**0.500**

4.3 Neural network-based Unconstrained GLS

Neural network-based unconstrained GLS follows the parameterized neural network language model used on text generation in NLP field, which firstly performs feature learning and obtains the probability distribution of all words at each time step. Based on the above idea, secret information can be embeded by fusing different encoding algorithms in the process of predicting the next word. According to the different neural network language models and frameworks, this paper subdivides this type of steganography into several categories: recurrent neural network (RNN) language model variants-based, transformer language model variants-based, generative adversarial network (GAN) language model-based and sequence to sequence model-based GLS.

4.3.1 RNN Language Model Variants-based GLS

RNN language model variants-based GLS is a kind of text steganography that learns the probability distribution of words based on neural network models such as RNN and its variant long short-term memory (LSTM). RNN has strong ability to capture sequential data features and generate word sequences. Later, LSTM further improves the memory ability of RNN and reduce the problems of gradient explosion and disappearance by modifying the cyclic function

of RNN from simple full connection to the memory unit of three control gates to recombine the current input and memory history information at each time step.

Based on the above theory, Fang et al. [11] introduces the LSTM into GLS for the first time. They use word-level LSTM for language modeling and group coding of dictionaries. However, if ER of steganography mentioned above needs to be adjusted, the grouping of all words and their corresponding coding values need to be adjusted. In order to improve this problem, Yang et al. [13] proposes FLC and VLC two dynamic coding algorithms to encode the conditional probability of words. And for improving the hidden capacity and text quality, Xiang et al. [44] proposes an LSTM text generation steganography based on character-level, which designs the selection strategy for the state selection coding results of characters, in other words, it will select the lowest perplexity and highest quality stego text. Since lyrics are more casual and flexible than poetry, Tong et al. [45] proposed two models based on char-RNN and word-RNN to generate lyrics combined with Huffman coding.

4.3.2 Transformer Language Model Variants-based GLS

Transformer language model variants-based GLS is a kind of text steganography that learns the probability distribution of words based on transformer [46] and its variants. Transformer is the actually the Encoder-Decoder network, which abandons the basic paradigm of using cyclic recursive structure to encode word sequences, but calculates the hidden states of sequences completely based on the global attention mechanism, obviously, it can better model the dependencies in long sequences than RNN.

The model mentioned above has achieved good results in the text generation task, therefore, Zachary et al. [47] introduced GPT-2 neural language model into text steganography to realize information hiding combined with arithmetic coding (AC), at the same time, they add common equivalent variant based on fixed precision binary fractions to alleviate the limitation of AC accuracy. Similarly, based on GPT-2 model, Dai et al. [48] proposed the patient-huffman algorithm to dynamically build candidate pool to realize information hiding, and dynamically adjust the ER of words by using the KLD of word probability distribution. So as to further improve the semantic quality of stego text, Yang et al. [49] firstly implements the controllable text generation model CTRL to generate stego text and the semantic classifier BERT to extract information. However, for ensuring the correct extraction of stego text, they propose the rejection sampling strategy, that is, adding a semantic classifier in the process of information embedding and judging whether the extracted content is consistent with the initial secret information.

4.3.3 GAN Language Model-based GLS

GAN language model-based GLS is a kind of text steganography based on generative adversarial network (GAN), which is subdivided into two parts: discriminant network and generative network. The discriminant network tries to distinguish cover and stego text, while the generative network generates stego text similar to the normal text to deceive the discriminant network. Through two parts of antagonistic training, the discrimination network can continuously improve the resolution, and the generation network can generate more natural stego text.

Based on the above ideas, Yang et al. [50] introduces the strategy update algorithm aiming to solve the problem that the traditional GAN is difficult to generate discrete data, it encodes the probability distribution of words at each iteration and then chooses the corresponding word

in terms of secret message each time. The innovative of the method proposed above lies in the antagonistic training between generator and discriminator.

4.3.4 Sequence to Sequence (seq2seq) Model-based GLS

Seq2seq model-based GLS refers to a kind of text steganography based on seq2seq [51] to achieve information hiding, which maps one sequence to another using neural network. Encoder of the framework vectorizes the input sentences and obtains the representation information of the sentences, and Decoder obtains the probability distribution of words finally. Because of its flexibility, this framework is now the preferred in NLP for the task of text generation, and different neural network models assume the roles of Encoder and Decoder respectively.

Based on the above ideas, Luo et al. [23] proposes a BiLSTM Encoder-Decoder model with attention mechanism as the text generation model, and combines with huffman encoding algorithm to generate quatrains, at the same time, they also use template-constrained generation method and a word-choosing approach using inner-word mutual information to focus on inter-word correlation. Yang et al. [52] first attempts to generate stego text with context sematic constraints, it uses BiRNN and self-attention to encode texts and then decode the encoded information by using another RNN model. In order to generate more semantic stego text, Yang et al. [18] puts forward the concept of cognitive-imperceptibility for the first time, presents two typical ways to realize that, namely IH-Enc and IH-Dec, combines GRU, transformer and topic-aware neural network and LSTM respectively to realize information hiding. Yang et al. [21] proposes a steganographic method-variational automatic encoder steganography (VAE-stega), to properly balance perceptual and statistical-imperceptibility, so as to better improve the ability of anti-steganalysis. In fact, VAE can also be regarded as framework of Encoder-Decoder, VAE-stega experiments that BERT-LSTM and LSTM-LSTM combined with huffman and AC coding respectively.

4.3.5 Analysis of Neural network-based Unconstrained GLS

Neural network-based unconstrained GLS first uses RNN to model sequence, but transformer can model the long-distance dependency in the sequence than RNN more effectively, which is soon widely used in text generation task to promote the development of GLS. Because text steganography and steganalysis are just antagonistic training, GAN provides a new perspective for GLS, but in fact, it is still a great challenge to optimize the generation network under the condition of how to select and generate discrete data. After that, the domain references the seq2seq to implement steganography tasks. This paper mainly evaluates ER and statistical-imperceptibility of neural network-based unconstrained GLS to evaluate the perceptual, statistical-imperceptibility and hidden capacity.

ER: This part evaluates mainly the ER of steganographic methods, RNN-stega (7.34%) [13], CLLS-128 (12.56%) and CLLS-256 (12.59%) [44], obviously, the steganography based on character-level owns the higher capacity.

Statistical-imperceptibility: **Table 4** and **Table 5** compare the perplexity, KLD, JSD and anti-steganalysis of this type of steganography. As shown in **Table 4**, with the increasing of ER, the text quality becomes worse and worse, and we can see clearly the quality based on GPT-2 (AC) [47] is the best, but in fact, under the same conditions, ER is not as good as other methods. KLD and JSD reflect the difference of the probability distribution between stego text and cover, the divergence between the two texts is greater with the increase of ER seen from **Table 4**. **Table 5** shows that the anti-steganalysis ability based on [13] (VLC) is stronger than [11].

Table 4. The comparison of *perplexity（ppl）*, *KLD and JSD* of neural network-based unconstrained GLS

Reference		IMDB[41]			Twitter[42]		
LSTM[11]	bpw	1.000	2.000	3.000	1.000	2.000	3.000
	ppl	30.665	40.027	74.543	62.205	54.181	96.30
	KLD	11.845	18.560	29.692	8.641	16.550	31.02
	JSD	12.115	15.101	19.362	12.039	14.847	20.24
RNN-stega (HC)[13]	bpw	1.000	1.845	2.565	1.000	1.873	2.656
	ppl	20.915	24.839	29.187	23.307	27.993	32.05
	KLD	19.507	18.628	17.551	8.295	8.161	8.195
	JSD	14.478	14.388	14.080	11.450	11.630	11.72
VAE-stega (LSTM-LSTM) (HC)[21]	bpw	1.000	1.863	2.577	1.000	1.890	2.687
	ppl	45.115	49.511	59.532	36.817	41.570	49.38
	KLD	**10.834**	10.902	10.878	**5.708**	5.621	5.530
	JSD	12.063	12.125	12.308	11.163	11.112	11.08
VAE-stega (BERT-LSTM) (HC) [21]	bpw	1.000	1.866	2.596	1.000	1.954	2.748
	ppl	30.266	36.349	40.832	35.076	37.887	46.75
	KLD	11.277	**10.193**	9.853	6.107	5.820	5.344
	JSD	11.979	11.803	11.809	11.024	10.963	10.87
GPT-2(AC)[47]	bpw	0.441	1.353	2.174	0.176	1.091	2.005
	ppl	**18.061**	**21.563**	**26.528**	**21.918**	**24.974**	**30.70**
	KLD	17.964	16.200	15.565	9.622	7.768	6.608
	JSD	13.288	13.150	13.325	11.563	11.164	11.02
VAE-stega (LSTM-LSTM) (AC) [21]	bpw	0.423	1.351	2.194	0.294	1.252	2.228
	ppl	39.166	37.817	51.090	29.083	34.336	42.88
	KLD	11.296	10.312	8.489	7.251	**5.166**	**4.462**
	JSD	**11.618**	11.468	11.373	11.067	**10.715**	**10.61**
VAE-stega (BERT-LSTM) (AC) [21]	bpw	0.409	1.358	2.205	0.337	1.182	2.169
	ppl	28.879	31.572	47.611	25.927	30.493	36.44
	KLD	11.912	9.337	**8.201**	6.914	6.573	5.621
	JSD	11.641	**11.319**	**11.185**	**10.995**	10.872	10.72

Table 5. The comparison of *anti-steganalysis ability* of neural network-based unconstrained GLS

Reference	Text steganalysis method	IMDB	Twitter	News
[11]	Method in [53]	**0.551**	**0.497**	**0.535**
	Method in [54]	0.615	0.642	0.665
	Method in [40]	0.830	0.759	0.718
	Method in [55]	0.823	0.778	0.743
[13]-FLC	Method in [53]	0.465	**0.480**	0.473
	Method in [54]	**0.553**	0.520	0.572
	Method in [40]	0.552	0.659	0.548
	Method in [55]	0.642	0.625	0.652
[17]-VLC	Method in [53]	0.475	**0.480**	0.485
	Method in [54]	0.562	**0.513**	**0.531**
	Method in [40]	**0.527**	**0.559**	**0.493**
	Method in [55]	**0.527**	**0.560**	**0.507**

4.4 Constrained GLS

Constrained GLS refers to a kind of steganography that realizes information hiding under some form of carrier constraint, and the constraint carrier includes image, text and so on. Different from unconstrained GLS, the constrained follows the similar idea of multimodal language generation and generates stego text under some form of carrier constraints. In this paper, we

subdivide this type of method into text-based constrained GLS and non-text-based constrained GLS based on the constraint carrier.

4.4.1 Text-based Constrained GLS

Text-based constrained GLS is a type of text steganography that realizes information hiding under the constraint of a certain format of text carrier, such as lists, jokes and so on. In the earliest, Desoky [56] promotes a constrained information-based paradigm of GLS, Nostega, and then presents many steganographic methods based on that. Later, such methods based on deep learning are also proposed successively.

Alice and *Bob* communicate with each other in different professions through Internet or Email. Under the certain constrained text in the specific field, *Alice* generates stego text according to the two modules of text generation and steganographic encoding and then sends them to many customers, but only *Bob* knows the rules to extract secret information. Desoky presents the steganography named Listega [57], which takes the subject of list items as constraint information and selects the encoded list candidate according to secret information to generate stego text. Subsequently, Desoky [58] proposes Chestega, which encodes the players or the position of the chessboard, and then generates training documents, game analysis, news articles and other texts under the constraints of these specific lists or chessboard steps. [59] designs Jokestega to generate steganographic textual joke under the constraint of keywords of joke. [60] proposes Educatega takes the questions, competitions and exams related education as constraint information to generate stego text to hide the existence of secret messages. This method generates questions containing secret information under the constraints of the question bank generated by the system. Desoky [61] also proposes a constrained GLS called Sumstega, which extracts secret information by comparing the differences between the cover and stego text under the constraints of the original text abstract.

Based on the idea above, Yang et al. [62] refers to the seq2seq of story generation, they use convolution language model to generate the premise firstly, and then generates the stego text through a CNN-CNN sequence under the constraint of the premise.

4.4.2 Non-text-based Constrained GLS

Non-text-based constrained GLS is a type of text steganography under the constraint of non-text carrier, which mainly refers to knowledge graphs (KGs) based constrained GLS and images based constrained GLS. The former generates stego text under the constraint of KGs, and the later constrains the generation of image title by extracting image features and hides the secret information.

Based on the constraint of KGs, Yang et al. [63] proposes Graph-stega using KG to guide sentence generation, which first encodes the path of KG and extracts the corresponding knowledge subgraph according to the secret information, then extracts semantics with graph state LSTM [64] under the constraint of knowledge subgraph and generates stego text with LSTM decoding end. Based on the same principle, Li et al. [65] selects the knowledge subgraph of a specific topic, and then accompanies with secret information as the input of the encoder based on the transformer. Under the constraint of the KG, the decoder LSTM was used to generate the stego text with a specific topic, this method combined with huffman encoding to realize information hiding.

Based on constrained of images, Wen et al. [16] constructs an end-to-end CNN-LSTM model, called neural image caption (NIC) to realize the task of generating steganographic titles under image information constraints, they also combining three encoding algorithms: word by word hiding (WWH), sentence by sentence hiding (SSH) and Hash hidding (HH). Different

from the steganography mentioned above, Li et al. [22] embeds secret information in image description by dynamic synonym substitution to overcome the drawback that the statistical characteristics of synonym frequency change after embedding.

4.4.3 Analysis of Constrained GLS

The linguistic steganography from unconstrained to constrained represents researchers pay more and more attention to this field. The existing restricted linguistic steganography is based on some widely used format carriers to realize information hiding, due to the massive existence of these carriers in the network and the novelty of steganography, the existing steganalysis methods are not very targeted, so it has good security and imperceptibility.

Cognitive-imperceptibility: This paper mainly compares the cognitive-imperceptibility of the steganography based on NIC [16] and SIC [22]. As can be seen from **Table 6**, the performance of the two frameworks is almost the same.

Table 6. The comparison of cognitive-imperceptibility of constrained GLS based on NIC and SIC.

Reference	BLEU-1	BLEU-2	BLEU-3	BLEU-4	CIDEr	ROUGE-L	METEOR
[16]	0.723	0.555	0.420	0.321	0.991	0.540	0.259
[22]	0.720	0.550	0.416	0.317	0.967	0.537	0.256

5 Conclusion and Future Work

Text steganography is a challenging and promising problem in the field of information security. Thanks to the emergence of GLS which aims to generate stego text close to cover, it greatly improves the hidden capacity and can resist previous steganalysis methods. As a comprehensive review on GLS, this paper focus on the framework of the GLS, classifies and evaluates properly the existing methods according to generation model, technique and encoding algorithm. GLS aims to ensure its security and improve the hidden capacity at a higher level, despite the tremendous success achieved of these issues in past years, there are remains a huge margin for improvement and development. This paper looks forward to the following domains after combing the existing contributions:

1. **Control text semantics to improve the quality of steganographic text:** The existing has made more efforts to improve the quality of generated steganographic text, but they just carry out predictive control during generation process, which is difficult to generate high-quality long text. Therefore, the next work will not only focus on the quality, but also consider the sematic consistency of the generated content to control the semantics necessarily, we should pay attention to the development of the controlled text generation (CTG) over the next few years.
2. **Minimal steganographic distortion to improve the ability of anti-steganalysis:** Most GLS only consider the difference between local candidate choices during the controlling generation process according to secret information, namely steganography distortion, resulting in poor ability of anti-steganalysis. Therefore, in order to reduce steganographic distortion, the next work will consider to adaptively control the steganographic text during the generation process from the global point of view, for example, optimizing the encoding algorithm to minimize distortion by using dynamic programming algorithm.
3. **Increase adaptation scene of constrained GLS to improve imperceptibility and security:** Compared with the unconstrained GLS, the constrained has higher imperceptibility, the next work will increase the categories of the constrained GLS to make

it has a broader application scenario, meanwhile, we should ensure the quality of the steganographic text and its imperceptibility.

References

[1] Y. J. Luo, J. H. Qin, X. Y. Xiang and Y. Tan, "Coverless image steganography based on multi-object recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 31, no. 7, pp. 2779-2791, July 2021. Article (CrossRef Link)

[2] Y. F. Huang, S. Tang and J. Yuan, "Steganography in inactive frames of VoIP streams encoded by source Codec," *IEEE Transactions on information forensics and security*, vol. 6, no. 2, pp. 296–306, 2011. Article (CrossRef Link)

[3] Z. H. Li, L. J. Meng, S. T. Xu, Z. H. Li and Y. Q. Shi, "A hevc video steganalysis algorithm based on pu partition modes," *Computers, Materials & Continua*, vol. 59, no. 2, pp. 563–574, 2019. Article (CrossRef Link)

[4] L. Y. Xiang, W. S. Wu, X. Li and C. F. Yang, "A linguistic steganography based on word indexing compression and candidate selection," *Multimedia Tools and Application*, vol. 77, no. 21, pp. 28969-28989, 2018. Article (CrossRef Link)

[5] R. Bergmair, "A comprehensive bibliography of linguistic steganography," in *Proc. of the SPIE International Conference on Security, Steganography, and Watermarking of Multimedia Contents IX*, San Jose, USA, vol. 6505, 2007. Article (CrossRef Link)

[6] T. Y. Liu and W. H. Tsai, "A new steganographic method for data hiding in Microsoft word documents by a change tracking technique," *IEEE Transactions on Information Forensics and Security*, vol. 2, no. 1, pp. 24-30, 2007. Article (CrossRef Link)

[7] L. Y. Xiang, X. H. Wang, C. F. Yang and P. Liu, "A novel linguistic steganography based on synonym run-length encoding," *IEICE Transactions on Information and Systems*, vol. E100-D, no. 2, pp. 313-322, 2017. Article (CrossRef Link)

[8] H. M. Meral, B. Sankur, A. S. Ozsoy, "Natural language watermarking via morphosyntactic alterations," *Computer Speech and Language*, vol. 23, no. 1, pp. 107-125, 2009. Article (CrossRef Link)

[9] X. Y. Chen, H. Y. Sun, Y. Tobe, Z. L. Zhou and X. M. Sun, "Coverless information hiding method based on the chinese mathematical expression," in *Proc. of International Conference on Cloud Computing and Security*, Cham, Switzerland, pp. 133–143, 2015. Article (CrossRef Link)

[10] L. Y. Xiang, G. Q. Guo, J. M. Yu, V. S. Sheng and P. Yang, "A convolutional neural network-based linguistic steganalysis for synonym substitution steganography," *Mathematical Biosciences and Engineering*, vol. 17, no. 2, pp. 1041-1058, Jan. 2020. Article (CrossRef Link)

[11] T. Fang, M. Jaggi and K. Argyraki, "Generating steganographic text with LSTMs," in *Proc. of 55th Proc. of ACL*, Vancouver, Canada, pp. 100-106, Jul. 2017. Article (CrossRef Link)

[12] G. J. Simmons, "The prisoners' problem and the subliminal channel," in *Advances in Cryptology, Proc. of Crypto 83*, Springer, Boston, MA, Jan. 1984, pp. 51–67. Article (CrossRef Link)

[13] Z. L. Yang, X. Q. Guo, Z. M. Chen, Y. F. Huang and Y. J. Zhang, "RNN-stega: Linguistic steganography based on recurrent neural networks," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 5, pp. 1280-1295. May, 2019. Article (CrossRef Link)

[14] J. Fridrich, *Steganography in digital media: principles, algorithms, and applications*, Cambridge University Press, 2009. Article (CrossRef Link)

[15] M. Chapman and G. Davida, "Hiding the hidden: a software system for concealing ciphertext as innocuous text," in *Proc. of International Conference on Information and Communications Security*, Beijing, China, pp. 335-345, 1997. Article (CrossRef Link)

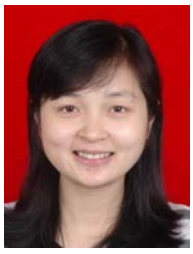
[16] J. Wen, X. J. Zhou, M. D. Li, P. Zhong and Y. M. Xue, "A novel natural language steganographic framework based on image description neural network," *Journal of Visual Communication and Image Representation*, vol. 61, pp. 157-169, 2019. Article (CrossRef Link)

[17] H. Q. Wang and S. Z. Wang, "Cyber warfare: steganography vs. steganalysis," *Communications of the ACM*, vol. 47, no. 10, pp.76-82, Oct. 2004. Article (CrossRef Link)

[18] Z. L. Yang, L. Y. Xiang, S. Y. Zhang, X. M. Sun and Y. F. Huang, "Linguistic generative steganography with enhanced cognitive-imperceptibility," *IEEE Signal Processing Letters*, vol. 28, pp. 409-419, 2021. Article (CrossRef Link)
[19] Z. L. Yang, S. Y. Jin, Y. F. Huang, Y. J. Zhang and H. Li, "Automatically generate steganographic text based on markov model and huffman coding," *arXiv preprint arXiv:1811.04720*, 2018. Article (CrossRef Link)
[20] N. Wu, Z. L. Yang, Y. Yang, L. Li, P. L. Shang and Z. R. Liu, "STBS-stega: coverless text steganography based on state transition-binary sequence," *International Journal of Distributed Sensor Networks*, vol. 16, no. 3, Mar. 2020. Article (CrossRef Link)
[21] Z. L. Yang, S. Y. Zhang, Y. T. Hu, Z. W. Hu and Y. F. Huang, "VAE-Stega: linguistic steganography based on variational auto-encoder," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 880-895, Sep. 2020. Article (CrossRef Link)
[22] M. Li, K. Wu, P. Zhong, J. Wen and Y. M. Xue, "Generating steganographic image description by dynamic synonym substitution," *Signal Processing*, vol. 164, pp. 193-201, Nov. 2019. Article (CrossRef Link)
[23] Y. B. Luo and Y. F. Huang, "Text steganography with high embedding rate: using recurrent neural networks to generate chinese classic poetry," in *Proc. of 5th ACM*, New York, USA, pp. 99–104, 2017. Article (CrossRef Link)
[24] Y. B. Luo, Y. F. Huang, F. F. Li and C. Chang, "Text steganography based on ci-poetry generation using Markov chain model," *Ksii Transactions on Internet and Information Systems*, vol. 10, no. 9, pp. 4568-4584, Sep. 2016. Article (CrossRef Link)
[25] C. L. Zhou, W. You and X. J. Ding, "Genetic algorithm and its implementation of automatic generation of chinese SONGCI," *Journal of Software*, vol. 21, no. 3, pp. 427-437, Mar. 2010. Article (CrossRef Link)
[26] K. Papineni, S. Roukos, T. Ward and W. J. Zhu, "BLEU: a method for automatic evaluation of machine translation," in *Proc. of the 40th ACL*, USA, pp. 311–318, 2002. Article (CrossRef Link)
[27] C. Y. Lin, G. H. Cao, J. F. Gao and J. Y. Nie, "An information-theoretic approach to automatic evaluation of summaries," in *Proc. of HLT-NAACL '06*, New York, USA, pp. 463–470, Jun. 2006. Article (CrossRef Link)
[28] M. Denkowski and A. Lavie, "Meteor universal: language specific translation evaluation for any target language," in *Proc. of the Ninth Workshop on Statistical Machine Translation*, Baltimore, Maryland, USA, pp. 376–380, Jun. 2014. Article (CrossRef Link)
[29] R. Vedantam, C. L. Zitnick and D. Parikh, "Cider: consensus-based image description evaluation," in *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, Boston, Massachusetts, pp. 4566–4575, Oct. 2015. Article (CrossRef Link)
[30] A. Gatt and E. Krahmer, "Survey of the state of the art in natural language generation: core tasks, applications and evaluation," *Journal of Artificial Intelligence Research*,vol. 61, no. 45, pp. 1-16, 2017. Article (CrossRef Link)
[31] P. Wayner, "Mimic functions," *cryptologia*, vol. 16, no. 3, pp. 193-214, 2010. Article (CrossRef Link)
[32] Z. S. Yu, L. S. Huang, Z. L. Chen, L. J. Li, W. Yang and X. X. Zhao, "High embedding ratio text steganography by ci-poetry of the Song dynasty," *Journal of Chinese Information Processing*, vol. 23, no. 4, pp. 55-63, 2009. Article (CrossRef Link)
[33] H. Yang and X. Cao, "Linguistic steganalysis based on meta features and immune mechanism," *Chinese Journal of Electronics*, vol. 19, no. 4, pp. 661-666, 2010.
[34] W. H. Dai, Y. Yu and B. Deng, "Bintext steganography based on Markov state transferring probability," in *Proc. of ICIS '09*, New York, USA, pp. 1306-1311, Nov. 2009. Article (CrossRef Link)
[35] W. H. Dai, Y. Yu and B. Deng, "Text steganography system using Markov chain source model and DES algorithm," *Journal of Software*, vol. 5, no. 7, pp. 785-792, Jul. 2010. Article (CrossRef Link)
[36] N. Wu, W. B. Ma, Z. R. Liu, P. L. Shang, Z. L. Yang and J. Fan, "Coverless text steganography based on half frequency crossover rule," in *Proc. of 4th ICMCCE*, Hohhot, China, pp. 726-7263, Oct. 2019. Article (CrossRef Link)

[37] N. Wu, Z. Liu, W. Ma, P. Shang, Z. Yang and J. Fan, "Research on coverless text steganography based on multi-rule language models alternation," in *Proc. of 4th ICMCCE*, Hohhot, China, pp. 803-8033, Oct. 2019. Article (CrossRef Link)
[38] H. H. Moraldo, "An approach for text steganography based on Markov chains," *arXiv preprint arXiv:1409.0915*, 2014. Article (CrossRef Link)
[39] A. N. Shniperov and K. A. Nikitina, "A text steganography method based on Markov chains," *Automatic Control and Computer Sciences*, vol. 50, no.8, pp. 802-808, 2016. Article (CrossRef Link)
[40] S. Samanta, S. Dutta and G. Sanyal, "A real time text steganalysis by using statistical method," in *Proc. of 2016 ICETECH*, Coimbatore, India, pp. 264-268, Mar. 2016. Article (CrossRef Link)
[41] A. L. Maas, R. E. Daly, P. T. Pham, D. Huang and A. Y. Ng, "Learning word vectors for sentiment analysis," in *Proc. of HLT '11*, Portland, Oregon, vol. 1, pp. 142–150, June. 2011.
[42] A. Go, R. Bhayani and L. Huang, "Twitter sentiment classification using distant supervision," *CS224N Project Report*, vol. 1, no. 12, Jan. 2009.
[43] A. Thomson, "News dataset". [Online]. Available: https://www.kaggle.com/snapcrack/all-the-news/data.
[44] L. Y. Xiang, S. H. Yang, Y. H. Liu, Q. Li and C. Z. Zhu, "Novel linguistic steganography based on character-level text generation," *Mathematics*, vol. 8, no. 5, pp. 1558, 2020. Article (CrossRef Link)
[45] Y. J. Tong, Y. L. Liu, J. Wang and G. J. Xin, "Text steganography on RNN generated lyrics," *Mathematical Biosciences and Engineering*, vol. 16, no. 5, pp. 5451-5463, Jun. 2019. Article (CrossRef Link)
[46] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit and L. Jones, "Attention is all you need," in *Proc. of the 31st International Conference on NIPS'17*, Red Hook, NY, USA, pp. 6000–6010, 2017. Article (CrossRef Link)
[47] Z. M. Ziegler, Y. T. Deng and A. M. Rush, "Neural linguistic steganography," in *Proc. of EMNLP-IJCNLP*, pp. 1210-1215, Sep. 2019. Article (CrossRef Link)
[48] F. Dai and Z. Cai, "Towards near-imperceptible steganographic text," in *Proc. of the 57th ACL*, Florence, Italy, pp. 4303-4308, Jul. 2019. Article (CrossRef Link)
[49] S. Y. Zhang, Z. L. Yang, J. S. Yang and Y. F. Huang, "Linguistic steganography: from symbolic space to semantic space," *IEEE Signal Processing Letters*, vol. 28, pp. 11-15, Dec. 2020. Article (CrossRef Link)
[50] Z. L. Yang, N. Wei, Q. H. Liu, Y. F. Huang and Y. J. Zhang, "GAN-TStega: text steganography based on generative adversarial networks," in *Proc. of Digital Forensics and Watermarking, 18th International Workshop*, Chengdu, China, vol. 12022, pp. 18-31, 2020. Article (CrossRef Link)
[51] I. Sutskever, O. Vinyals and Q. V. Le, "Sequence to sequence learning with neural networks," *Advances in neural information processing systems*, vol. 2, pp. 3104-3112, Dec. 2014. Article (CrossRef Link)
[52] Z. L. Yang, P. Y. Zhang, M. Y. Jiang, Y. F. Huang and Y. J. Zhang, "RITS: real-time interactive text steganography based on automatic dialogue model," in *Proc. of ICCCS*, Haikou, China, vol. 11065, pp. 253-264, 2018. Article (CrossRef Link)
[53] P. Meng, L. S. Hang, W. Yang, Z. L. Chen and H. Zheng, "Linguistic steganography detection algorithm using statistical language model," in *Proc. of ITCS '09*, Kiev, Ukraine, vol. 2, pp.540–543, Jul. 2009. Article (CrossRef Link)
[54] Z. Chen, L. Huang, Z. Yu, W. Yang and L. Li, "Linguistic steganography detection using statistical characteristics of correlations between words," in *Proc. of IH 2008: Information Hiding*, Santa Barbara, CA, USA, pp. 224–235, Oct. 2008. Article (CrossRef Link)
[55] A. Joulin, E. Grave, P. Bojanowski and T. Mikolov, "Bag of tricks for efficient text classification," in *Proc. of the 15th Conference of the European Chapter of the Association for Computational Linguistics*, Valencia, Spain, vol. 2, pp. 427-431, 2017. Article (CrossRef Link)
[56] A. Desoky, "Nostega: a novel noiseless steganography paradigm," *Journal of Digital Forensic Practice*, vol. 2, no. 3, pp. 132-139, Dec. 2008. Article (CrossRef Link)
[57] A. Desoky, "Listega: list-based steganography methodology," *International Journal of Information Security*, vol. 8, pp. 247-261, Apr. 2009. Article (CrossRef Link)

[58] A. Desoky and M. Younis, "Chestega: chess steganography methodology," *Security and Communication Networks*, vol. 2, no. 6, pp. 555-566, Mar. 2009. Article (CrossRef Link)
[59] A. Desoky, "Jokestega: automatic joke generation-based steganography methodology," *International Journal of Security & Networks*, vol. 7, no. 3, pp.148-160, 2013. Article (CrossRef Link)
[60] A Desoky, "Edustega: an education-centric steganography methodology," *International Journal of Security and Networks*, vol. 6, no. 2-3, pp. 153-173, 2011. Article (CrossRef Link)
[61] A. Desoky, "Sumstega: summarisation-based steganography methodology," *International Journal of Information and Computer Security*, vol. 4, no. 3, pp. 234-263, May, 2011. Article (CrossRef Link)
[62] R. Yang and Z. H. Ling, "Linguistic steganography by sampling-based language generation," in *Proc. of APSIPA ASC*, Lanzhou, China, 2019. Article (CrossRef Link)
[63] Z. L. Yang, B. T. Gong, Y. M. Li, J. S. Yang and Y. F. Huang, "Graph-Stega: semantic controllable steganographic text generation guided by knowledge graph," Jun. 2020. Article (CrossRef Link)
[64] L. F. Song, X. C. Peng, Y. Zhang, Z. G. Wang and D. Gildea, "AMR-to-text generation with synchronous node replacement grammar," in *Proc. of the 55th ACL*, Vancouver, Canada, vol. 2, pp. 7-13, Jul. 2017. Article (CrossRef Link)
[65] Y. M. Li, J. Zhang, Z. L. Yang and R. Zhang, "Topic-aware neural linguistic steganography based on knowledge graphs," *ACM/IMS Transactions on Data Science*, vol. 2, no. 2, pp. 1-13, Apr. 2021.



Lingyun Xiang received her Ph.D. degree in computer science and technology from Hunan University in 2011. She is an associate professor in the School of Computer and Communication Engineering, Changsha University of Science and Technology. Her research interests include information security, information hiding and digital watermarking, steganalysis, natural language processing, text content security.

Rong Wang is currently pursuing the M.S. degree in electronic information with the School of Computer and Communication Engineering, Changsha University of Science and Technology, China. Her research interests include information hiding and natural language processing.

Zhongliang Yang received his B.S. degree in School of Electronic Science and Technology from Sichuan University in 2015, and Ph.D. degree in the Department of Electronic Engineering from Tsinghua University in 2020. He is now continuing his postdoctoral research at Tsinghua University. His research interests include steganography, steganalysis and natural language processing.

Yuling Liu is currently an Associated professor in the College of Computer Science and Electronic Engineering at Hunan University, China. She was Visiting scholar at UMASS Lowell in 2016. She received the Ph. D degree in Computer Science from Hunan University, China, in 2008. Her research interests include network and information security, information hiding based on big data, text analysis.