DOI: 10.3969/j.issn.0255-8297.2021.06.004

文本隐写及隐写分析综述

康慧娴， 易 标， 吴汉舟
上海大学 通信与信息工程学院，上海 200444

摘 要： 梳理了文本隐写与隐写分析的发展脉络，将文本隐写算法分为两类：修改式文本隐写和生成式文本隐写。归纳了两类算法的实现过程，并从率失真性能和安全性等方面分析主流算法的优势与不足。针对两类文本隐写算法，总结了对应的隐写分析算法实现过程，并对文本隐写与隐写分析的发展趋势进行了展望。

关键词： 网络安全；信息隐藏；文本隐写；文本隐写分析；深度学习
中图分类号： P751.1 **文章编号：** 0255-8297(2021)06-0923-16

Recent Advances in Text Steganography and Steganalysis

KANG Huixian, YI Biao, WU Hanzhou
School of Communication and Information Engineering, Shanghai University, Shanghai 200444, China

Abstract: This paper sorts out the development context of text steganography and steganalysis, and divides text steganography algorithms into two categories: modified text steganography and generative text steganography. The implementation process of the two types of algorithms is summarized, and the advantages and disadvantages of mainstream algorithms are analyzed from the aspects of rate-distortion performance and safety. Aiming at the two types of text steganography algorithms, the realization process of the corresponding steganalysis algorithms is summarized, and the development trend of text steganography and steganalysis is prospected.

Keywords: network security, information hiding, text steganography, text steganalysis, deep learning

香农将网络空间中的信息安全系统分为以下 3 种：加密系统、隐私系统和隐蔽通信系统 [1]。在保护信息安全时，加密系统和隐私系统暴露了秘密信息的存在，因此容易受到针对性攻击，而隐蔽通信系统则通过将秘密信息嵌入到特定的载体中来确保信息的安全。

隐写作为实现隐蔽通信的重要方法，是信息安全领域中一个热门的研究方向。隐写是将秘密信息隐藏在公开载体中并通过公开渠道传送，该过程不仅隐藏了秘密信息，还隐藏了秘密通信的行为。随着互联网的兴起与发展，文字、图片、视频和音频等多媒体为隐写提供了丰富的载体。由于文本高度编码的特性，能够供以修改的冗余信息较少，因此难以将秘密信息嵌

收稿日期： 2021-06-01
基金项目： 国家自然科学基金（No.61902235）；上海市人才项目“晨光计划”（No. 19CG46）资助
通信作者： 吴汉舟，博士，副教授，研究方向为信息隐藏。E-mail: hanzhou@shu.edu.cn

入其中，其研究成果远远少于以视频和图像为载体的信息隐藏研究成果。但文字作为人们日常生活中使用最频繁的信息传输载体，研究文本隐写仍具有重要意义。近年来，飞速发展的自然语言处理技术使文本隐写逐渐成为隐写领域中新兴的研究热点。

隐写分析是指对隐写的攻击，目的是为了检测秘密信息的存在甚至破坏秘密通信。隐写分析分为主动分析和被动分析。主动分析是以提取秘密信息为目标，估计嵌入的秘密信息长度、嵌入位置、嵌入算法使用的密钥、删除或破坏隐写对象中的秘密信息等；而被动分析只需要检测到秘密信息和秘密通信的存在。目前，隐写分析大多以被动分析为主。其中，文本隐写分析本质是一个二分类任务，主要目标是将文本区分为含密文本和不含密文本。为了对文本进行分类，需要提取出能够表征文本的统计特征，然后分析隐写前后这些特征的变化，最后设计相应的鉴别器。传统的文本隐写分析方法通常需要手动构建和选择合适的文本统计特征，随着深度学习的发展，越来越多的神经网络模型可以用于自动提取合适的文本特征。

隐写与隐写分析是矛与盾的关系，两者螺旋上升发展。当今社会中每个人都非常注重个人信息安全，文本隐写的研究非常重要。但是为了防止一些不法分子如黑客和恐怖分子等使用文本隐写技术传输危险信息从而危害公共安全，研究性能强大的文本隐写分析技术同样十分重要。本文梳理了文本隐写与隐写分析的发展脉络，归纳了各类算法的实现过程，分析总结了各类算法的优势和不足，并给出了展望，以方便该领域的研究者了解其研究进展。

1 文本隐写算法分类与分析

本文将现有文本隐写算法划分成两类，即修改式隐写算法和生成式隐写算法，如图 1 所示。修改式隐写算法通过分析文本格式特征或者字符属性特征，寻找冗余信息位来隐藏秘密信息，也可以通过分析文本语法或语义特征，利用同义词替换或句法变换来隐藏秘密信息。生成式隐写算法通过构建语言模型自动生成文本载体，并在生成过程中利用熵编码方法嵌入秘密信息。

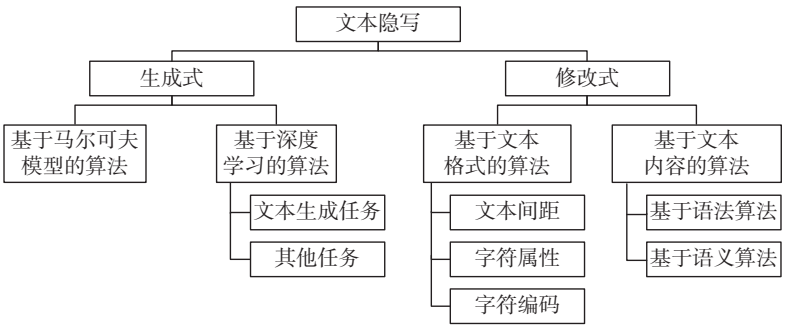

图 1 文本隐写算法分类

Figure 1 Classification of text steganography algorithms

1.1 修改式文本隐写

修改式文本隐写指的是在已有文本载体的基础上，对文本格式或者文本内容加以修改并嵌入秘密信息。其中，基于文本格式的信息隐藏利用文本间距、字符属性以及字符在计算机中编码特征来隐藏信息；而基于文本内容的算法以自然语言处理技术为支撑，通过对文本语法和句法的分析，发掘出文本内容特征，构造合适的算法隐藏信息，该类算法可以划分成基于语法的修改和基于语义的修改。

1.1.1　基于文本格式的算法

文本都是由字、行和段等有规律的结构组成，因此文献 [2] 提出了字移编码和行移编码算法，在不可查的前提下，字移编码通过字符左移或右移微小的距离来嵌入信息，行移编码则通过行上移或下移微小的距离来嵌入信息。在该算法中行移编码嵌入容量小于字移编码，但字移编码隐蔽性较低。

除了文本间距，字符的属性如字体大小、样式和颜色等也可以用来隐藏秘密信息。针对英文文本富含很多空格这一特性，文献 [3] 利用空格字符的字体大小来隐藏秘密信息，略微更改空格字符字体大小隐藏“1”，不修改隐藏“0”。对此文献 [4] 进行了改进，不仅利用空格字符的字体大小，还利用了空格字符的类型，其中字体类型可以编码 6 比特的秘密信息，字体大小编码 1 比特的秘密信息，从而提升了信息隐藏容量，但该算法受到字体类型种类的限制。

文本除了空格等不可见字符，还包含许多可见字符，为了充分利用文本中每个字符，文献 [5] 提出了基于混合大小写字体的信息隐藏算法，利用英文单词的每个字母包含大小写两种样式的特性，用大写字母隐藏“1”，小写字母隐藏“0”，理论上可以让每个字母都隐藏秘密信息，可以实现大容量，但是隐写文本视觉上容易被发觉，隐蔽性较差。

考虑在保证嵌入容量的基础上提升安全性，研究人员提出了修改字符颜色来隐藏秘密信息，文献 [6] 将二进制比特流按 8 位划分转换成 10 进制 RGB 值，以 RGB 值的形式将秘密信息隐藏在不可见的字符中，每个字符可以隐藏 24 比特。上述算法基本利用空格字符隐藏信息，更加适合英文文本。文献 [7] 通过修改每个字符和下划线 RGB 的 3 个通道最低位的值来嵌入信息，可以应用于中文文本，相比上述仅利用空格字符的算法，该算法大大扩展了可使用的字符，提升了信息隐藏容量，但不适用于隐蔽性要求较高的场合，因为过度的修改字符颜色易被攻击者发现。文献 [8] 提出针对中文文本字体的信息隐藏方法，通过对中文字符字体的修改嵌入秘密信息。

上述算法基于字符的不同属性隐藏秘密信息，本质上是对人眼可感知的文本属性进行修改。而文本在计算机中是以二进制形式存储的，那么就可以从计算机的角度出发，提出一些基于字符编码的信息隐藏算法。Unicode 编码为各种语言的字符设定了统一且唯一的二进制码，具有很大的通用性，因此它被认为是基于字符编码隐写的常用编码方式。文献 [9] 提出通过 Unicode 码的奇偶性隐藏信息，该算法将文本载体用十进制 Unicode 编码表示，秘密信息视为二进制比特串，对比秘密比特与单个字符 Unicode 编码的奇偶：若同为奇数则嵌入 1，若同为偶数则嵌入 0，若奇偶性不一致则通过视觉上不可察的方法对文字格式进行修改，使算法对嵌入位置进行标记。该算法的嵌入位置从文本第 1 个字符开始，若文本传输丢失了部分文字，则无法提取完整秘密信息。为了解决上述非顺序提取导致的秘密信息不能复原的问题，文献 [10] 提出插入分隔符作为区分标志，同样利用了 Unicode 的奇偶性，遇到奇偶性不一致时，通过设置标记位的方法进行隐藏，即使部分载体遭到破坏，仍然可以提取完整秘密信息，因此提升了稳健性。文献 [11] 将文本信息和隐藏信息都转换成 Unicode 编码，并全都转化成二进制形式，按位异或得到新的二进制串后再转成十进制 Unicode 编码，得到隐写文本。相比于利用 Unicode 奇偶性的方法，载体字符和隐藏字符因为都采用了 Unicode 编码方式，所以明显提升了信息隐藏容量。

研究人员通过研究文本字符编码表发现，有些编码后的字符插入文本后不会被人眼感知，此类字符被称为“不可见字符”。利用这个特性，文献 [12] 提出了一种基于 Hash 函数与不可见 ASCII 字符替换的信息隐藏算法，该算法用“SOH”这一不可见字符来替换文本分段中的空格，对替换后的分段文本进行 Hash 运算，然后将 Hash 值与隐藏信息进行比较，根据设定规则嵌入信息。文献 [13] 根据约束函数找到嵌入位置，再根据秘密信息为 0 或 1 分别嵌入空

格或“SOH”字符，该算法的嵌入能力受到了约束函数的限制。文献 [14] 扩展了编码方式，提出了基于 Unicode 编码的不可见字符嵌入算法，该算法将 Unicode 不可见字符编码两两组合表示成二进制序列，从而可以将秘密比特串翻译成不可见字符插入到文本每个句子的句号前，完成秘密信息隐藏。理论上利用不可见字符的嵌入秘密信息方法具有很大的容量，独立于文本文件的格式和排版，但一般的文本文件不允许存在过多空字符，因此限制了此类算法的应用。

1.1.2 基于文本内容的算法

随着自然语言处理技术的发展，研究者们应用文本分词、句法分析等技术可以更好地分析文本的统计特征，在对文本特征建模后，能够构造出合适的算法来嵌入秘密信息。根据目前的研究成果，可以将基于文本内容的算法划分成基于语法的修改和基于语义的修改两类。

基于语法的信息隐藏技术以自然语言处理技术为基础，利用句中词语的依赖关系，或者句式变换等语法规则，研究者们构造了特定的规则用来嵌入秘密信息。文献 [15] 提出了一种基于虚词变换的隐写方法，该算法以助词“的”为例，在保证文本原意的前提下，通过增删“的”的方式嵌入信息，该算法定义了模板给增删提供依据，具有一定的灵活性，但其信息隐藏容量不高。文献 [16] 将变换提升至句子层，提出了一种基于移位变换的句子层隐写方法，该算法以词性标注为基础，定义了句子中词序变换的规则，根据该规则改变句子中词语的顺序来隐藏秘密信息，增大了隐藏容量，但可能导致句子语义不通。在英文文本方面，文献 [17] 提出了一种基于上下文替换的隐写方法，通过英文载体中定冠词和指示性形容词可以互相替代的特点嵌入信息，该算法保证了隐写文本的语义通顺，但其信息隐藏复杂度较高。

基于语法的方法为了保证修改后的文本安全，通常在句子级别嵌入秘密信息，牺牲了信息隐藏容量；基于语义的隐写方法更加细化了对文本内容的研究，从字词层面嵌入以提高信息隐藏容量。其目的是为了保证修改后的载体符合真实文本的语义，其中一个热门的研究点是基于同义词替换隐藏信息，包括同义词库构建，同义词编码方法设计以及同义词替换规则设计在内的 3 个优化方向。

文献 [18] 研究了人们在社交网络中沟通的习惯，发现人们会用缩写代替完全形式词语，因此提出了通过缩写和完全形式词语替换的方法隐藏信息，保证了语义不变性。文献 [19] 在上述算法的基础上增强了安全性，对单词缩写列表进行了加密，使其从静态列表变为动态列表，这使攻击者很难在仅知悉算法的情况下从文本中提取秘密信息。虽然保证了修改后文本的语义，但在这种利用缩写替换完全形式词语的隐写算法中，信息隐藏容量受到了缩写词语数目的限制，并且只适用于英语等有缩写词语的语言。

为了在语义不变的前提下提升信息隐藏容量，研究者们想到了利用同义词替换的方法，并产生了不同的优化方向。文献 [20] 提出了基于同义词替换的隐写方法，该算法先将文本载体分词，再判断每个词语是否在同义词库中，如果在，就根据该词所在的同义词组中词的个数以及当前所需嵌入的二进制位串，使用霍夫曼编码进行同义词替换。鉴于上述方法的同义词组中存在不可替代的词，经过替换可能破坏了句子的语义，文献 [21] 提出了一种改进的同义词替换信息隐藏算法，将同义词分为完全可替换词组、不完全可替换词组和歧义词组 3 类，根据需要选择不同方法替换。该算法降低了替换后语义失真程度。

除了构建同义词库，文献 [22] 从信息论角度出发，通过改进编码方式来提高同义词的嵌入率，设计了一种矩阵编码与同义词替换结合的信息隐藏方法，该算法在嵌入相同秘密信息的前提下，减少了对文本的修改。而文献 [23] 重点考虑替换的同义词是否合适，根据 WordNet 字典和 Internet 上收集的统计信息计算出给定单词和上下文单词的组合频率，选取出最合适

的同义词用以替换。文献 [20] 应用了依存句法分析获取同义词与搭配词的相容情况，相比于文献 [23] 仅考虑相邻单词的相关程度，还考虑到文本中远距离的词也可能对单词产生影响。图 2 给出了个基于同义词替换的文本隐写算法示例，如果载体文本的句子中出现同义词“情况”，那么可以用“情形”代表编码 0，“情况”代表编码 10，“状况”代表编码 11，根据需要嵌入的信息选择合适的同义词进行替换。

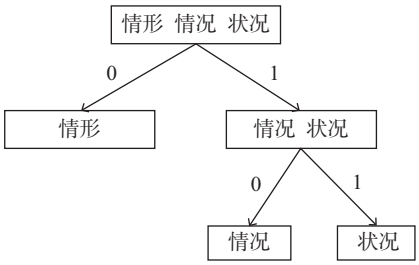

图 2　基于同义词替换的文本隐写算法示例

Figure 2　Example of text steganography algorithm based on synonym substitution

1.2　生成式文本隐写

随着自然语言处理技术的发展，文本生成技术愈加成熟。修改式隐写算法需要已有的文本载体，与此不同的是，生成式的文本隐写以文本生成技术为基础，自动生成含密载体，提高了秘密信息在传输中的安全性，保护了发送方和接收方。

生成式文本隐写通常分为两个步骤，第 1 步是自动文本生成，第 2 步是秘密信息嵌入。目前的研究主要是对这两个步骤进行优化。

1.2.1　基于马尔可夫模型的算法

马尔可夫模型可用于建模自然文本的生成过程，利用该模型进行文本生成时可以使用定长编码、霍夫曼编码等编码方式来嵌入秘密信息。

文献 [24] 根据一些公共文本建立状态转移图，对每个分支使用二比特定长编码，使用美国数据加密标准（data encryption standard, DES）算法将秘密信息从字节流转变成比特流，再根据状态转移图选择单词，生成隐写文本。该算法没有考虑到状态转移图中词与词之间概率大小的不同，因此文献 [25] 改进了编码方式，根据需要编码的位数 n，将转移图中的状态按照概率分配码字，直至每个码都对应唯一的一个短句，该短句即为隐写文本。除此之外，还使用标识符代替了发送方自己设定的开头词语，增加了随机性。该算法考虑到了状态转移概率，提升了文本质量，但是牺牲了信息隐藏容量。为了提升隐写文本的质量，研究者们考虑一些特殊体裁的文本，在中文领域如诗和词。文献 [26] 采用宋词这个体裁，用平仄过滤候选词语，将候选词语使用霍夫曼编码，根据编码值选择词语生成隐写文本。该算法提升了文本质量，但受体裁的限制，其实用性较差。文献 [27] 打破了上述算法的限制，提出了一种根据马尔可夫链模型和霍夫曼编码自动生成隐写文本的算法，将状态转移图中的条件概率进行霍夫曼编码，从而选择出合适的词语用以生成隐写文本。该算法提升了文本质量和信息隐藏容量，生成的文本也不受体裁的限制，具有普适性。

1.2.2　基于深度学习的算法

按照生成式隐写方法的两个步骤，研究者们分别对语言模型和建模方式进行了优化，基于深度学习的算法主要以文本预训练模型为基础。在语言模型方面，采用了长短

期记忆网络（long short-term memory, LSTM）、生成式对抗网络（generative adversarial network, GAN）、生成式预训练（generative pre-trained transformer, GPT）和变分自编码器（variational auto-encoder, VAE）等模型；编码方面采用了定长编码、霍夫曼编码和算术编码等熵编码方法。基于深度学习的隐写算法根据任务不同又可以划分成通用文本隐写和基于特定任务的文本隐写。图 3 是一个生成式文本隐写算法示例，该模型根据当前词计算出候选词并依据其概率分布构建霍夫曼树，将候选词作为叶子节点，然后从根节点开始，将其左孩子节点和右孩子节点分别编码成 0 和 1，根据秘密信息选择适合的词语（图中词 3）作为输出。

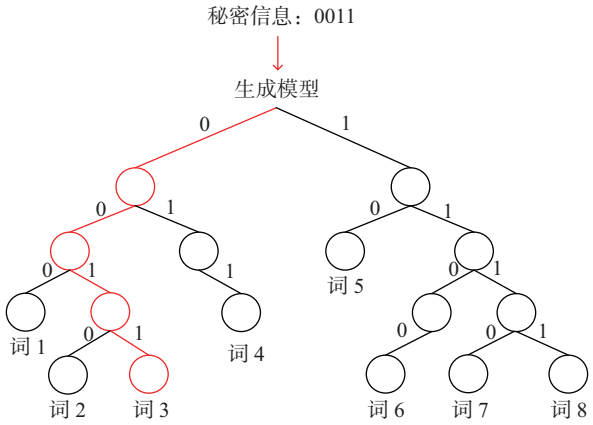

图 3 生成式文本隐写算法示例

Figure 3 Example of generative text steganography algorithm

通用文本隐写算法适用于所有文本，其基础是神经网络预训练模型的高速发展，这些隐写算法能够最大程度地模拟真实文本的统计特征，并自动生成文本，而各种熵编码方法能巧妙地将秘密信息嵌入到文本载体中。文献 [28] 最先提出了基于 LSTM 网络的文本隐写方法，该算法构建了一个发送方和接收方共享的词典，给该词典中的每个词一个固定长度的编码，根据秘密比特流和 LSTM 的概率转移关系挑选词语来构成隐写文本。相比于马尔可夫模型，该算法明显地提升了文本质量。但是如果需要调整嵌入率，或者需要重新构建词库并编码，其计算复杂度则很高。为此，文献 [29] 进行了改进，提出了针对转移概率的动态编码方法，其中包括定长编码方法和变长编码方法，两者都基于 LSTM 网络。定长编码是将每一步得到的候选词用相同位数的码进行编码，再根据秘密信息选择合适的词；变长编码是根据候选词的概率采用霍夫曼编码，再结合秘密信息选出合适的词。该类算法采用的动态编码方式提高了生成文本的质量，增强了文本的不可感知性和信息隐藏容量。但是当嵌入量增大时，定长编码降低了文本的质量，变长编码又降低了嵌入效率。除此之外，由于 LSTM 网络对长文本的注意力不够，无法保证句子与句子之间的连贯性。为了解决这一问题，文献 [30] 采用 GAN 模型来模拟整段文字的统计特性，生成器是文本生成模型 GRU，判别器是一个隐写分析模型。将生成器得到的文本序列输入到判别器中，即可输出一个二分类的结果，然后判断是否为隐写文本，当判别器区分不出时，生成器即可用于生成隐写文本。编码方法采用了文献 [22] 提出的定长编码方式，该算法加强了句子之间的联系，提升了文本质量，训练中的副产物判别器可以直接作为隐写分析模型。文献 [31] 基于 LSTM 网络，通过引入关键词和注意力机制来增强句子间连贯性，在每次迭代的时候始终受到关键词的控制，保证生成的文本不过分地偏离主题，同样采用文献 [29] 提出的两种编码方式，提升了生成文本尤其长文本的质量。上述算法评价

文本质量的指标都是 perplexity，该指标更符合人类判断文本质量的标准，文献 [32] 提出了一种基于算术编码的文本隐写算法，该算法更加注重客观判别指标 KL 散度 (Kullback-Leibler divergence)，采用表现更加优秀的语言模型 GPT-2 作为文本生成模型，将秘密信息用算术编码成二进制的小数，再根据语言模型的概率分布构建同心圆，根据秘密信息找到唯一路径，该路径上的词语构成了隐写文本。该算法生成的文本从概率分布上与真实文本更加相似。文献 [33] 进一步提出了基于自适应算术编码的文本隐写算法，相比于文献 [32] 每次解码都选择相同 K 个单词作为候选词构建同心圆，该算法每次解码时自适应选择不同数目的词，并按概率划分成不同的子区间，再根据自适应算术编码后的秘密信息选择合适区间对应的词作为输出，直至隐藏完所有的秘密信息。该算法提升了信息隐藏容量，但也增加了计算复杂度。上述文章有的侧重感官不可察觉性，有的侧重统计不可察觉性，为了平衡两者，文献 [34] 提出了基于变分自编码器（variational auto-encoder，VAE）的文本隐写算法，编码器采用双向编码器（bidirectional encoder representations from transformer, BERT）模型，解码器采用 LSTM 模型，用编码器学习到不含密文本的统计分布特征。该算法的生成阶段不同于直接指定输入，它是从编码器学习到的样本空间中随机采样出隐向量作为输入，并在生成阶段采用霍夫曼编码和算术编码完成信息隐藏。

除了纯文本之外，许多特定任务如图像理解、视觉故事和对话生成等都包含文本部分，这些文本部分也可以进行隐写。文献 [35] 提出了一种实时交互式文本隐写模型，应用双向 RNN 模型和注意力机制将输入编码成隐向量，单向循环神经网络（recurrent neural network, RNN）对隐向量解码获得响应文本，并在解码过程中应用定长编码嵌入秘密信息。文献 [36] 提出了一种基于图像描述的文本信息隐藏算法，首先利用卷积神经网络 (convolutional neural network, CNN) 对输入图像进行编码得到隐向量，再用 LSTM 对隐向量解码获得描述，在解码过程中，分别设计了基于句子和基于单词的秘密信息嵌入算法。文献 [37] 提出了一种基于视觉故事生成的隐写算法，该算法应用 ResNet 对多幅图像进行编码，然后利用双向 LSTM 对图像序列之间的关系进行建模，最后使用单向 LSTM 对隐向量解码得到视觉故事，并在解码阶段根据概率自适应嵌入不同比特。

2　现有文本隐写算法对比

文本隐写的评价指标与隐写相同，即需要判断隐写的不可感知性、安全性以及嵌入容量。文本隐写不可感知性评价包括主观评价和客观评价，主观评价即人类对隐写文本的流畅度、语义相关性等方面进行综合打分，若打分过低，视为隐写文本，不可感知性较差；而客观评价使用文本质量评价的标准指标困惑度（用 P_{er} 表示）来衡量文本隐写的不可感知性，对于隐写文本 $\boldsymbol{S}=\{\boldsymbol{\omega}_1,\boldsymbol{\omega}_2,\cdots,\boldsymbol{\omega}_n\}$ 而言，若单词 $\boldsymbol{\omega}_i$ 出现的概率为 $P(\omega_i)$，则 P_{er} 的计算公式为

$$P_{\mathrm{er}}=2^{-\frac{1}{n}\lg P(\omega_1,\omega_2,\cdots,\omega_n)} \tag{1}$$

文本隐写的安全性通过 KL 散度来评判，设 X 与 Y 是属于同一样本空间 $\boldsymbol{\chi}$ 的随机变量，分别表示隐写文本和不含密文本，且分别服从 $P(X)$ 与 $P(Y)$ 的概率分布，$P(X)$ 与 $P(Y)$ 的 KL 散度 $D_{\mathrm{KL}}(P(X)\|P(Y))$ 的公式为

$$D_{\mathrm{KL}}(P(X)\|P(Y))=\sum P(X)\lg\frac{P(X)}{P(Y)} \tag{2}$$

当 $D_{\mathrm{KL}}(P(X)\|P(Y))=0$ 时，认为隐写系统是绝对安全的，即不含密文本与隐写文本有完全相同的分布。不含密文本与隐写文本差异越大则 $P(X)$ 与 $P(Y)$ 差异越大，散度值也越

大。文本隐写的嵌入容量指隐写传输的信息量，即在满足不可感知性及安全性的前提下，文本载体中能够嵌入秘密信息比特的最大值。

本文对各种算法进行了分类整理，结果如表 1 所示。

表 1 文本隐写分析算法对比

Table 1 Comparison of text steganalysis algorithms

实验配置			FCN[44]			CNN[35]			GNN[36]		
数据集	方法	嵌入率	AC	PR	RE	AC	PR	RE	AC	PR	RE
IMDB	Bins[7]	1.000	0.789	0.806	0.748	0.835	0.861	0.799	**0.859**	**0.864**	**0.851**
		2.000	0.895	0.907	0.880	0.922	0.945	0.895	**0.939**	**0.947**	**0.929**
		3.000	0.946	0.951	0.941	0.959	0.962	0.955	**0.967**	**0.967**	**0.967**
	FLC[43]	1.000	0.705	0.731	0.649	0.773	**0.791**	0.743	**0.788**	0.789	**0.786**
		2.000	0.844	0.866	0.813	0.882	0.895	0.867	**0.901**	**0.921**	**0.878**
		3.000	0.926	0.935	0.915	0.946	0.949	0.942	**0.959**	**0.958**	**0.961**
	VLC[43]	1.000	0.708	0.732	0.656	0.770	0.792	0.733	**0.784**	**0.793**	**0.769**
		2.183	0.848	0.857	0.837	0.896	0.911	0.877	**0.913**	**0.914**	**0.911**
		3.285	0.919	0.937	0.900	0.943	0.952	0.933	**0.960**	**0.956**	**0.964**
Twitter	Bins[7]	1.000	0.765	0.802	0.903	0.822	0.859	0.772	**0.835**	**0.874**	**0.783**
		2.000	0.876	0.910	0.834	0.919	**0.947**	0.889	**0.929**	0.945	**0.911**
		3.000	0.930	0.954	0.903	0.957	**0.977**	0.937	**0.962**	0.970	**0.953**
	FLC[43]	1.000	0.680	0.694	0.645	0.760	0.791	0.708	**0.744**	**0.813**	**0.713**
		2.000	0.824	0.839	0.803	0.888	**0.914**	0.857	**0.897**	0.912	**0.878**
		3.000	0.915	0.935	0.893	0.947	**0.967**	0.926	**0.949**	0.962	**0.936**
	VLC[43]	1.000	0.690	0.700	0.664	0.760	**0.833**	0.650	**0.771**	0.820	**0.694**
		2.179	0.838	0.860	0.807	0.892	0.916	0.863	**0.895**	**0.922**	**0.864**
		3.261	0.913	0.935	0.887	0.944	**0.960**	0.927	**0.954**	0.957	**0.952**

从表 1 可以看出，文本隐写算法的嵌入能力具有显著差异，具体分析如下：

1）基于文本格式的算法

该类算法研究成果较多，不同算法嵌入能力差异较大，多数算法都只针对某一特定属性提出。例如，基于字符编码的算法以每个字符嵌入 1 比特居多 [9-10,12-13]，而基于字符属性的算法主要通过改动不可见字符的大小 [3]、颜色 [6-7] 或类型 [4] 来隐藏信息，使得算法不再受到文件编码格式的限制，但也牺牲了信息隐藏容量；若对可见字符 [5,8] 进行改动虽然保证了隐藏容量，但是无法保证隐蔽性。此类算法操作简单，容量较大，但鲁棒性不强，多数不能抵抗重新排版的攻击，抗统计分析攻击能力差。

2）基于文本内容的算法

该类算法以修改文本内容为主，对于基于语法的算法[15-17]和基于语义的算法[20-23]都以自然语言处理技术为基础。虽然在嵌入级别上有所不同，但都需要对文本进行分词和分析，因此会产生较为复杂的计算过程。基于语法的算法大多倾向于研究者自己手动定义的模板进行变换，生成的文本容易引起语法错误或者常识错误，上下文语义也常常不连贯，统计特征与正常文本差异较大。与之相比，基于语义的算法具有更强的鲁棒性和更好的隐蔽性。而无论哪种算法，嵌入的信息都容易受到对文本内容增加、删除、替换的攻击，使得隐写文本的稳定性不高。为了提高算法的安全性，该类算法要么牺牲容量保证隐蔽，要么对文本特征进行更加细致的分析，才能使嵌入的秘密信息难以被察觉。

3）基于马尔可夫模型的算法

这类算法[24-27]将马尔可夫模型作为模拟生成文本的统计语言模型，在保证文本质量的基础上，依据不同的编码方式嵌入机密信息。马尔可夫链的两次近似使得该模型不能很好地替代统计语言模型，需要进一步优化所生成的隐写文本质量。该类算法在特殊的文本载体如诗歌[26]上展现出较好的隐写效果，使得研究者们将研究重点转移到生成式隐写算法。

4）基于深度学习的算法

该类算法遵循两个步骤，即自动文本生成和秘密信息嵌入。对于通用文本隐写算法和基于特定任务的文本隐写算法，嵌入容量受到了不同熵编码方式的影响，起初该类算法的优化目标是保证一定嵌入量的同时达到感官不可察觉的目的[28-31]，后来为了满足统计不可察觉性，又牺牲了一定的嵌入量[32-33]，最后需要平衡感官不可察觉及统计不可察觉[34]。除了通用的文本隐写算法，其他涉及文本的任务[35-37]也可以对文本进行隐写，从而扩展文本隐写的应用场景，但是与通用的方法相比，嵌入量相对较少。从总体上看，该类算法生成的文本质量较高，信息隐藏容量和安全性较高，抗统计分析攻击能力较强，但由于构建语言模型需要大量参数，此类算法的计算过程较复杂。

3 文本隐写的主要问题及解决办法

归纳起来，现阶段文本隐写存在以下几个主要问题：

1）信息隐藏容量较小

无论是修改式的方法还是生成式的方法，由于文本载体的冗余空间较少，为保证隐蔽性，嵌入算法均不能达到嵌入的上限，从而导致信息隐藏容量较小。

2）稳健性差

目前，大多数算法的稳健性都不够高，具体表现在文件格式整体被替换，文本内容被增加、删除或修改，都会导致接收者无法完整提取甚至错误提取秘密信息。

3）计算复杂度高

为了让隐写文本的统计特征与自然文本尽可能相近，并具有抗统计分析的能力，所构建的语言模型需要花费大量的运算资源。而且，模型参数作为密钥虽然保证了秘密信息在解码过程中的安全性，但需要占据大量的存储空间。

相应地，未来领域的研究重点可围绕以下几个方向展开：

1）提高信息隐藏容量

生成式文本隐写算法相比于修改式文本隐写算法的安全性更高，能够产生更多的冗余空间，为了提高信息隐藏容量，可以更多地考虑生成式的文本隐写方法。

2）增强稳健性

对于格式攻击造成的信息缺失，需要对文本进行加密处理，防止未经授权的用户对隐写

文本进行修改。而对于增删、替换词语攻击，可通过对句子中词语按重要度进行编码嵌入信息，实现有效的抵御。

3）降低计算复杂度

统计语言模型的优化不是文本隐写研究的主要优化目标，因此可将语言模型共享到公共网络云平台，利用可靠的第三方来代替保管网络参数，即公钥，同时发送方、接收方应准备一副私钥用以保证秘密通信。

4 文本隐写分析算法分类与分析

隐写分析与隐写是一个对抗的过程，针对不同的隐写算法，研究人员也提出了多种隐写分析算法。本文在广泛调研现有算法的基础上，将文本隐写分析算法划分成传统文本隐写分析算法和基于深度学习的文本隐写分析算法。

传统的文本隐写分析算法通常手动构建一系列的文本特征，然后分析隐写前后这些特征的变化，最后设计相应的二分类器来区分隐写文本和正常文本。对于不同的隐写算法，分别设计对应的特征，如文献 [38] 对相邻单词之间长度差异建模，通过对字移编码隐写算法进行统计分析，可以检测到秘密信息的存在并估计秘密信息的大小。对字符类型改动的隐写算法，文献 [39] 将每个字符属性特征映射成向量，利用改动字符和未改动字符的特征向量训练一个基于支持向量机（support vector machine, SVM）的二分类器，进而检测秘密信息的存在并且估计秘密信息的长度。文献 [40] 针对虚词变换的隐写算法，将语言特征映射成特征向量，再用最近邻分类器区分隐写文本和正常文本。针对同义词替换的隐写算法，文献 [41-42] 设计了不同的特征，前者从 3-gram 语言模型中提取特征向量，而后者基于高频词的数量在嵌入后总是减少的事实，利用由同义词位置及其同义词的数量组成的属性对的相对频率来构建特征向量。

除了上述针对特定隐写算法的隐写分析算法，研究者们也提出了许多通用的隐写分析算法，同样满足手动构建特征向量并进行分类的基本流程。文献 [43] 针对多种隐写算法，提出了基于单词分布的统计检测算法，该算法将单词的分布作为特征向量用以分类，并用 SVM 作为分类器。文献 [44] 改进了上述算法，用基于不同自然频率区域中的单词分布作为特征向量，提升了检测的精度。对于不同的隐写算法，应用不同特征检测得到结果有所差异，因此，文献 [45] 提出了一种基于元特征征和免疫克隆机制的隐写分析方法。该算法定义了 57 种特征，包括平均单词长度、空格率、字母百分比等，然后利用免疫克隆机制选择合适的特征用以隐写分析，该算法的普适性更强。

除了手动提取特征训练分类器之外，研究者们还提出了根据统计特性差异设置阈值区分正常文本和隐写文本，而不必训练分类器，降低了计算复杂度。如文献 [46] 基于 N-gram 语言模型计算正常文本和隐写文本的困惑度，并通过设定阈值进行隐写分析。文献 [47] 提出了基于贝叶斯估计和相关系数的隐写分析算法。

由于生成式隐写算法的不断发展，基于深度学习的隐写算法能够生成统计上与真实文本相似的隐写文本，传统的手动提取特征的方法很难在该类算法上达到很高的精度。因此，研究者们充分利用神经网络可以自动提取文本特征这一特性，提出了许多基于深度学习的隐写分析算法。文献 [48] 首先分析了文本相邻两个单词之间的关联性，然后将所有的单词映射到一个语义空间，并使用一个隐藏层来提取词与词之间的关联特征，最终将提取到的特征送入一个 softmax 分类器进行分类。文献 [49] 考虑到文本相邻多个单词之间的相关性，利用词嵌入层提取单词的语义和语法特征，然后利用不同尺寸的矩形卷积核学习句子特征，将提取出的所有特征拼接起来送入 softmax 分类器进行分类。文献 [50] 简化了隐写分析过程，采用文献 [49] 提出的算法进行文本特征的提取，通过比对该文本特征和真实文本特征在二维空间中

的分布判断其是否含有秘密信息。文献 [51] 分析了自动生成的隐写文本，发现每个单词的条件概率分布在嵌入秘密信息后会失真，为了提取到文本的长期依赖关系，提出了基于双向递归神经网络来提取文本中每个单词的条件分布特征，并将该特征送入到 softmax 分类器进行分类。文献 [52] 结合了上述算法的优点，使用含注意力机制的双向长短期记忆递归神经网络从文本中捕获长期上下文信息，并使用具有不同内核大小的卷积神经网络提取局部特征，将两个特征合并再送入到 softmax 分类器进行分类。文献 [53] 分析了文本单词之间的潜在关系，同时考虑到了单词周围词对其影响，以及不同距离单词对语义的潜在影响，提出了使用具有多种大小的卷积滑动窗口提取文本特征，将提取出的特征和真实文本特征在二维空间的分布进行对比得到检测结果。文献 [54] 考虑到文本多层特征，相比于上述只用最后一层隐藏层特征作为特征向量的算法，该算法利用特征金字塔提取文本不同维度的特征，再基于该特征进行二分类。考虑到自然语言中单词与单词之间有着复杂的关联结构，用序列的形式不足以进行有效建模，且无法有效地利用全局信息。文献 [55] 提出了一种基于图神经网络的隐写分析方法，首先为每一个文本创建一个与之对应的文本图，然后将文本图输入图卷积神经网络中进行特征提取。

相比于手动选择特征向量，基于深度学习自动提取特征向量的方法普适性更强，检测精度也更高。文本隐写分析的评价指标与普通二分类任务的评价指标基本一致，即准确率、精准率、召回率。唯一需要注意的地方是，由于隐写分析的主要目的是找出隐写文本，因此将隐写文本视为正例，也就是说标签设置为 1。表 2 给出了一些基于深度学习的隐写分析方法用在 IMDB 和 Twitter 两个数据集上，检测基于递归神经网络的生成式隐写算法的结果。

未来，隐写分析算法的发展方向是不断优化文本特征提取算法。另外相较于载体文本而言，显示场景中收集到的隐写文本数据量往往较少，用以训练分类器的数据不充分，使得隐写分析算法会向半监督甚至无监督的方向优化。

表 2　文本信息隐藏算法对比

Table 2　Comparison of text information hiding algorithms

分类	文献	原理	嵌入能力	局限
修改式文本隐写算法	[2]	基于字移编码、行移编码	一个字符或一行嵌入 1 比特	隐藏容量低
	[9]	基于 Unicode 编码奇偶性与字符属性修改	每个字符嵌入 1 比特	严格按照初始位置提取，稳健性差
	[10]	基于 Unicode 编码奇偶性	每个字符嵌入 1 比特	只适用于 Unicode 编码字符
	[11]	基于 Unicode 编码的二进制异或	每个字符嵌入 15 比特	隐蔽性低
	[12]	基于 Hash 函数与不可见 ASCII 字符替换	与文本空格数量与分组大小有关	受字符编码方式限制
	[13]	基于不可见 ASCII 字符替换	与文本空格数量及约束函数选择相关	受字符编码方式限制

续表 2

分类	文献	原理	嵌入能力	局限
修改式文本隐写算法	[14]	基于 Unicode 编码的不可见字符嵌入	每个句子嵌入 2 比特	稳健性差
	[3]	基于空格字符字体的大小	每个段落嵌入 1 比特	隐藏容量低
	[4]	基于空格字符字体类型和大小	每个空格嵌入 7 比特	受文档版本限制
	[5]	基于混合英文字母大小写	每个字母嵌入 1 比特	隐蔽性低
	[6]	基于不可见字符的前景色	每个空格嵌入 24 比特	普适性低
	[7]	基于字符和下划线 RGB 的 3 个通道最低位值	每个字符嵌入 3 比特	隐蔽性低
	[8]	基于中文字体修改	每个字符嵌入 1 比特	受文档版本限制
	[15]	增删句中助词“的”字	每次修改嵌入 1 比特	分析句子过程复杂
	[16]	改变句子的词序	每次修改嵌入 1 比特	隐蔽性低
	[17]	句子中定冠词和指示性形容词的替换	每次修改嵌入 2lb3 比特	信息隐藏效率低
	[18]	缩写与完全形式词语替换	取决于单词编码位数	隐藏容量低
	[19]	动态的缩写与完全形式词语替换	取决于单词编码位数	隐藏容量低
	[20]	筛选出适合替换的中文同义词	取决于同义词编码位数	无法抵抗同义词替换攻击
	[21]	从不同类型同义词库中筛选出适合替换的中文同义词	取决于同义词编码位数	无法抵抗同义词替换攻击
	[22]	基于矩阵编码的同义词替换	取决于矩阵编码效率	无法抵抗同义词替换攻击
	[23]	基于单词组合频率的同义词替换	取决于同义词编码位数	无法抵抗同义词替换攻击
	[20]	基于单词与搭配词相容的同义词替换	取决于同义词编码位数	无法抵抗同义词替换攻击
生成式文本隐写算法	[24]	定长编码马尔可夫状态转移概率	每个词语嵌入 2 比特	文本质量低
	[25]	变长编码马尔可夫状态转移概率	每个句子嵌入一定比特	隐藏容量低
	[26]	霍夫曼编码候选词语	每个词语嵌入霍夫曼编码比特数	体裁受到限制
	[27]	霍夫曼编码马尔可夫状态转移概率	每次词语嵌入一定比特	文本质量低
	[28]	基于 LSTM 的固定词表编码	取决于每个词语编码的位数	计算复杂度高
	[29]	基于 LSTM 的定长和变长编码	每个词语嵌入 1∼5 比特	句子间关联性差
	[30]	基于 GAN 的动态编码	每个词语嵌入 1∼5 比特	计算复杂度高
	[31]	主题引导式动态编码	每个词语嵌入 1∼5 比特	统计上易察觉

续表 2

分类	文献	原理	嵌入能力	局限
生成式文本隐写算法	[32]	基于 GPT-2 的算术编码	每个句子嵌入少量比特	隐藏容量低
	[33]	基于 GPT-2 的自适应算术编码	每个句子嵌入少量比特	计算复杂度高
	[34]	基于 VAE 的动态编码	每个词语嵌入 1∼5 比特	语义不可感知性差
	[35]	实时交互文本隐写算法	每个句子嵌入 2 比特	通用性差
	[36]	基于图像描述的隐写算法	每个句子或单词嵌入少量比特	隐藏容量低
	[37]	基于视觉故事生成的隐写算法	每个短句嵌入 1∼4 比特	鲁棒性差

5　结　语

随着互联网的飞速发展，个人信息安全意识的不断增加，隐写作为保证秘密通信的一个重要手段愈发成了研究热点，而文本的高流动性使得文本隐写受到越来越多的关注，相关理论研究也会愈加完善。未来，其研究成果也会逐步产业化，以更好地应对当今的信息安全问题。为了防止隐写算法被不法分子用于传递危险信息从而危害公共安全，而作为隐写对立面的隐写分析，必然要随着隐写的发展而不断发展。

参考文献:

[1] Shannon C E. Communication theory of secrecy systems [J]. The Bell System Technical Journal, 1949, 28(4): 656-715.

[2] Low S H, Maxemchuk N F, Brassil J T, et al. Document marking and identification using both line and word shifting [C]//Proceedings of INFOCOM. IEEE, 1995, 2: 853-860.

[3] Mahato S, Yadav D K, Khan D A. A novel approach to text steganography using font size of invisible space characters in Microsoft word document [M]. New Delhi: Springer, 2014.

[4] Kumar R, Malik A, Singh S, et al. A space based reversible high capacity text steganography scheme using font type and style [C]//International Conference on Computing, Communication and Automation (ICCCA). IEEE, 2016: 1090-1094.

[5] Ali A A, Saad A. New text steganography technique by using mixed-case font [J]. International Journal of Computer Applications, 2013, 62(3): 6-9.

[6] Khairullah M D. A novel text steganography system using font color of the invisible characters in Microsoft word documents [C]//2009 second international conference on computer and electrical engineering. IEEE, 2009: 482-484.

[7] Tang X, Chen M S. Design and implementation of information hiding system based on RGB [C]//2013 3rd International Conference on Consumer Electronics, Communications and Networks. IEEE, 2013: 217-220.

[8] 陈芳, 王冰. 基于文本字体的信息隐藏算法 [J]. 计算机技术与发展, 2006, 16(1): 20-22.
Chen F, Wang B. An algorithm of text information hiding based on font [J]. Computer Technology and Development, 2006, 16(1): 20-22. (in Chinese)

[9] 付兵. 基于字符 Unicode 编码奇偶性的文本信息隐藏算法研究 [J]. 福建电脑, 2008, 24(12): 66.
Fu B. Research on text information hiding algorithms based on Unicode coding parity [J]. Fujian Computer, 2008, 24(12): 66. (in Chinese)

[10] 陆绿, 方勇. 基于字符 Unicode 奇偶性的数字水印设计与实现 [J]. 计算机技术与发展, 2010, 20(8): 176-179.

Lu L, Fang Y. Design and implementation of digital watermark based on character Unicode parity [J]. Computer Technology and Development, 2010, 20(8): 176-179. (in Chinese)

[11] 黄国超, 王衍波, 张凯泽. 基于 Unicode 编码的信息隐藏算法研究与设计 [J]. 计算机技术与发展, 2011, 21(10): 233-236.
Huang G C, Wang Y B, Zhang K Z. Research and design of information hiding algorithm based on encoding of Unicode [J]. Computer Technology and Development, 2011, 21(10): 233-236. (in Chinese)

[12] Liu F, Luo P P, Ma Z J, et al. Security secret information hiding based on hash function and invisible ASCII characters replacement [C]//2016 IEEE Trustcom/BigDataSE/ISPA. IEEE, 2016: 1963-1969.

[13] 崔光明, 洪星, 袁翔, 等. 基于不可见字符替换的信息隐藏方法研究 [J]. 计算机应用与软件, 2016, 33(4): 277-280.
Cui G M, Hong X, Yuan X, et al. Research on information hiding based on invisible characters replacement [J]. Computer Applications & Software, 2016, 33(4): 277-280. (in Chinese)

[14] 张震宇, 李千目, 戚湧. 基于不可见字符的文本水印设计 [J]. 南京理工大学学报（自然科学版）, 2017, 41(4): 405-411.
Zhang Z Y, Li Q M, Qi Y. Text watermarking design based on invisible characters [J]. Journal of Nanjing University of Science and Technology, 2017, 41(4): 405-411. (in Chinese)

[15] 赵敏之, 孙星明, 向华政. 基于虚词变换的自然语言信息隐藏算法研究 [J]. 计算机工程与应用, 2006, 42(3): 158-160.
Zhao M Z, Sun X M, Xiang H Z. Research on the Chinese text steganography based on the modification of the empty word [J]. Computer Engineering and Application, 2006, 42(3): 158-160. (in Chinese)

[16] 刘玉玲, 孙星明, 辛国江. 基于移位变换的句子层自然语言信息隐藏算法 [J]. 控制与决策, 2009, 24(12): 1861-1864.
Liu Y L, Sun X M, Xin G J. Algorithm of natural language information hiding based on shift conversion in sentence level [J]. Control and Decision, 2009: 24(12): 1861-1864. (in Chinese)

[17] Wang F, Huang L S, Chen Z L, et al. A novel text steganography by context-based equivalent substitution [C]//2013 IEEE International Conference on Signal Processing, Communication and Computing (ICSPCC 2013). IEEE, 2013: 1-6.

[18] Shirali-Shahreza M, Shirali-Shahreza M H. Text steganography in SMS [C]//International Conference on Convergence Information Technology (ICCIT 2007). IEEE, 2007: 2260-2265.

[19] Rafat K F. Enhanced text steganography in SMS [C]//2009 2nd International Conference on Computer, Control and Communication. IEEE, 2009: 1-6.

[20] 甘灿. 基于同义词替换的自然语言文本信息隐藏技术研究 [D]. 长沙: 湖南大学, 2008.

[21] 甘灿, 孙星明, 刘玉玲, 等. 一种改进的基于同义词替换的中文文本信息隐藏方法 [J]. 东南大学学报(自然科学版), 2007, 37(增刊 1): 137-140.
Gan C, Sun X M, Liu Y L, et al. Improved steganographic algorithm based on synonymy substitution for Chinese text [J]. Journal of Southeast University (Natural Science Edition), 2007, 37(Suppl.1): 137-140. (in Chinese)

[22] 杨潇, 李峰, 向凌云. 基于矩阵编码的同义词替换隐写算法 [J]. 小型微型计算机系统, 2015, 36(6): 1296-1300.
Yang X, Li F, Xiang L Y. Synonym substitution-based steganographic algorithm with matrix coding [J]. Journal of Chinese Mini-Micro Computer Systems, 2015, 36(6): 1296-1300. (in Chinese)

[23] Bolshakov I A, Gelbukh A. Synonymous paraphrasing using word net and Internet [C]//International Conference on Application of Natural Language to Information Systems. Heidelberg: Springer, 2004: 312-323.

[24] Dai W H, Yu Y, Dai Y H, et al. Text steganography system using Markov chain source model and DES algorithm [J]. Journal of Software, 2010, 5(7): 785-792.

[25] Moraldo H H. An Approach for text steganography based on Markov chains [EB/OL]. [2014-09-02]. https://arxiv.org/abs/1409.0915.

[26] Luo Y B, Huang Y F, Li F F, et al. Text steganography based on ci-poetry generation using Markov chain model [J]. KSII Transactions on Internet and Information Systems (TIIS), 2016, 10(9): 4568-4584.

[27] Yang Z L, Jin S Y, Huang Y F, et al. Automatically generate steganographic text based on Markov model and Huffman coding [EB/OL]. [2018-11-12]. https://arxiv.org/abs/1811.04720.

[28] Fang T, Jaggi M, Argyraki K. Generating steganographic text with LSTMs [EB/OL]. [2017-05-30]. https://arxiv.org/abs/1705.10742.

[29] Yang Z L, Guo X Q, Chen Z M, et al. RNN-Stega: linguistic steganography based on recurrent neural networks [J]. IEEE Transactions on Information Forensics and Security, 2018, 14(5): 1280-1295.

[30] Yang Z L, Wei N, Liu Q S, et al. GAN-TStega: text steganography based on generative adversarial networks [C]//International Workshop on Digital Watermarking. Cham: Springer, 2019, 12022 LNCS: 18-31.

[31] Kang H X, Wu H Z, Zhang X P. Generative text steganography based on LSTM network and attention mechanism with keywords [J]. Electronic Imaging, 2020(4): 2911-2918.

[32] Ziegler Z M, Deng Y, Rush A M. Neural linguistic steganography [EB/OL]. [2019-09-03]. https://arxiv.org/abs/1909.01496.

[33] Shen J M, Ji H, Han J W. Near-imperceptible neural linguistic steganography via self-adjusting arithmetic coding [EB/OL]. [2020-10-01].https://arxiv.org/abs/2010.00677.

[34] Yang Z L, Zhang S Y, Hu Y T, et al. VAE-Stega: linguistic steganography based on variational auto-encoder [J]. IEEE Transactions on Information Forensics and Security, 2020, 16: 880-895.

[35] Yang Z L, Zhang P Y, Jiang M Y, et al. Rits: real-time interactive text steganography based on automatic dialogue model [C]//International Conference on Cloud Computing and Security. Cham: Springer, 2018, 11065: 253-264.

[36] 薛一鸣, 周雪婧, 周小诗, 等. 基于图像描述的文本信息隐藏 [J]. 北京邮电大学学报, 2018, 41(6): 7-13.

Xue Y M, Zhou X Q, Zhou X S, et al. Text information hiding based on image caption [J]. Journal of Beijing University of Posts and Telecommunications, 2018, 41(6): 7-13. (in Chinese)

[37] Guo Y Y, Wu H Z, Zhang X P. Steganographic visual story with mutual-perceived joint attention [J]. EURASIP Journal on Image and Video Processing, 2021(1): 1-14.

[38] Li L J, Huang L S, Zhao X X, et al. A statistical attack on a kind of word-shift text-steganography [C]//International Conference on Intelligent Information Hiding and Multimedia Signal Processing, IEEE, 2008: 1503-1507.

[39] Xiang L Y, Sun X M, Luo G, et al. Research on steganalysis for text steganography based on font format [C]//Third International Symposium on Information Assurance and Security. IEEE, 2007: 287-294.

[40] 曾莱蓓, 孙星明, 向凌云, 等. 基于虚词变换的文本隐藏信息检测方法研究 [J]. 计算机研究与发展, 2009, 46(增刊 1): 177-181.

Zeng L B, Sun X M, Xiang L Y, et al. Research on text hidden information detection method based on virtual word transformation [J]// Journal of Computer Research and Development, 2009, 46(Suppl.1): 177-181. (in Chinese)

[41] Taskiran C M, Topkara U, Topkara M, et al. Attacks on lexical natural language steganography systems [C]//Security, Steganography, and Watermarking of Multimedia Contents VIII. International Society for Optics and Photonics, 2006, 6072: 607209.

[42] Ling Y X, Sun X M, Gang L, et al. Linguistic steganalysis using the features derived from synonym frequency [J]. Multimedia Tools and Applications, 2014, 71(3): 1893-1911.

[43] Chen Z L, Huang L S, Yu Z S, et al. A statistical algorithm for linguistic steganography detection based on distribution of words [C]//2008 Third International Conference on Availability, Reliability and Security. IEEE, 2008. 558-563.

[44] Chen Z L, Huang L S, Yu Z S, et al. Effective linguistic steganography detection [C]//IEEE International Conference on Computer & Information Technology Workshops. IEEE, 2008: 224-229.

[45] Yang H, Cao X B. Linguistic steganalysis based on meta-features and immune mechanism [J]. Chinese Journal of Electronics, 2010, 19(4): 661-666.

[46] Meng P, Hang L S, Yang W, et al. Linguistic steganography detection algorithm using statistical language model [C]//2009 International Conference on Information Technology and Computer Science. IEEE, 2009: 540-543.

[47] Samanta S, Dutta S, Sanyal G. A real time text steganalysis by using statistical method [C]//2016 IEEE International Conference on Engineering and Technology (ICETECH). IEEE, 2016: 264-268.

[48] Yang Z L, Huang Y F, Zhang Y J. A fast and efficient text steganalysis method [J]. IEEE Signal Processing Letters, 2019, 26(4): 627-631.

[49] Wen J, Zhou X J, Zhong P, et al. Convolutional neural network based text steganalysis [J]. IEEE Signal Processing Letters, 2019, 26(3): 460-464.

[50] Yang Z L, Wei N, Sheng J Y, et al. TS-CNN: text steganalysis from semantic space based on convolutional neural network [EB/OL]. [2018-10-18]. https://arxiv.org/abs/1810.08136.

[51] Yang Z L, Wang K, Li J, et al. TS-RNN: text steganalysis based on recurrent neural networks [J]. IEEE Signal Processing Letters, 2019, 26(12): 1743-1747.

[52] Bao Y J, Yang H, Yang Z L, et al. Text steganalysis with attentional LSTM-CNN [C]//International Conference on Computer and Communication Systems (ICCCS 2020). IEEE, 2020: 138-142.

[53] Yang Z L, Huang Y F, Zhang Y J. TS-CSW: text steganalysis and hidden capacity estimation based on convolutional sliding windows [J]. Multimedia Tools and Applications, 2020, 79(25): 18293-18316.

[54] Yang H, Bao Y J, Yang Z L, et al. Linguistic steganalysis via densely connected LSTM with feature pyramid [C]//Proceedings of the 2020 ACM Workshop on Information Hiding and Multimedia Security, 2020: 5-10.

[55] Wu H Z, Yi B, Ding F, et al. Linguistic steganalysis with graph neural networks [J]. IEEE Signal Processing Letters, 2021, 28: 558-562.

(编辑：管玉娟)