可规避文本信息拦截的“火星文”生成隐写系统

朱嘉豪[1], 张玉书[1], 刘 哲[1], 张新鹏[2]

[1]南京航空航天大学计算机科学与技术学院 南京 中国 211106

[2]复旦大学计算机科学与技术学院 上海 中国 200433

摘要 近些年来, 由于互联网企业竞争激烈, 各平台文本信息存在着相互恶意拦截的问题, 这往往给用户带来不便甚至造成损失。目前, 在中文文本信息过滤领域中,“火星文”在规避关键词屏蔽方面效果显著。然而, 随着人工智能的快速发展, 检测技术不断提升, 仅仅依靠规避关键词屏蔽已然不足以确保文本信息传递的安全性, 文本关键信息仍然存在着被拦截的风险, 这是由于这类关键信息的呈现模式通常具有规律性。为了解决这类问题, 本文采用了文本信息隐藏技术。鉴于传统文本隐写算法的局限性, 本文提出了一种基于“火星文”生成的文本隐写系统。该文本隐写系统利用“火星文”较于传统平面媒介的语言形式而言, 信息冗余度高的特点, 将重要内容隐藏至文本中。该文本隐写系统主要由预处理、控制以及隐写三大基本模块组成。通过对汉字结构特征的研究以及“火星文”构字方式的分析, 本文设计出了 6 种隐写子模块以供信息嵌入与提取。实验结果分析, 所提出的隐写方案的嵌入容量高于同类型隐写方案, 且具有较强的鲁棒性。此外, 我们给出该文本隐写系统在互联网中的一个具体应用, 从而体现其实用性。

关键词 文本信息过滤; 关键词屏蔽; 文本信息隐藏; 火星文

中图法分类号 TP37 DOI 号 10.19363/J.cnki.cn10-1380/tn.2022.09.01

A Steganographic System based on “Martian” Generation for Avoidance of Text Information Interception

ZHU Jiahao[1], ZHANG Yushu[1], LIU Zhe[1], ZHANG Xinpeng[2]

[1] College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China

[2] School of Computer Science, Fudan University, Shanghai 200433, China

Abstract In recent years, due to fierce competition among Internet companies, there exists a phenomenon that text information on various platforms has been maliciously intercepted by each other, which may bring inconvenience and even loss to users. Nowadays, “Martian” is effective in avoiding keyword blocking in the field of Chinese text information filtering because of the complexity of the formation of “Martian” characters and the variety of “Martian” characters. However, with the rapid development of artificial intelligence, information filtering technology has improved dramatically. Hence it is far from enough to maintain the security of text information transmission only by keyword blocking avoidance. The vital text content still has the risk of being intercepted by information filtering systems owing to the regularity of the presentation pattern of such content. To address this problem, we adopt text information hiding technology. Given that “Martian” has more redundant information than the language form of traditional plane medium and the traditional text information hiding schemas are not suitable for the avoidance of text information blocking, we select “Martian” as the steganographic carrier and propose a “Martian” generation based steganographic system. Through the combination of “Martian” and text information hiding technology, some vital and sensitive information can be free from being exposed in the text, improving the security of information transmission. The text steganographic system proposed in this paper consists of three basic modules: preprocessing module, control module, and steganographic module, where the preprocessing module is mainly responsible for the normalization of data to be embedded, the control module is in charge of task assignment during data embedding and extraction process, and the steganographic module is designed for some specific data embedding and extraction tasks. Through the research on the structural characteristics of Chinese characters and the analysis of the construction pattern of “Martian”, we designed six kinds of steganographic sub modules for data embedding and extraction. Extensive experiments and theoretical analysis demonstrates that the proposed text steganographic system achieves higher embedding capacity than other similar text steganographic schemes and possesses good robustness. In

通讯作者: 张玉书, 教授, 博士生导师, Email: yushu@nuaa.edu.cn。

本课题得到国家自然科学基金项目(No. 62072237)的资助和江苏省自然科学基金面上项目-基于混沌压缩感知的物联网数据安全低能耗获取技术研究(No. BK20201290)的资助。

收稿日期: 2021-09-17; 修改日期: 2021-11-11; 定稿日期: 2022-07-15

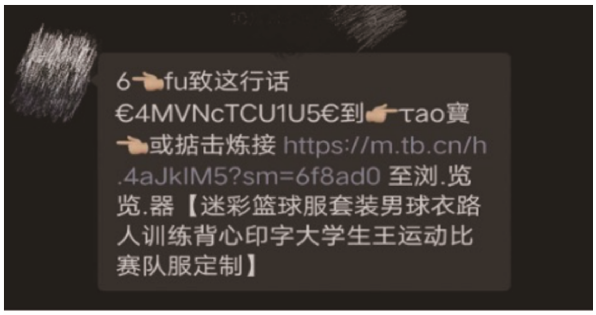
addition, we present a specific Internet application example of the text steganography system based on “Martian” generation, which reflects its practicality.

Key words text information filtering; keyword blocking; text information hiding; Martian

1 引言

随着互联网的普及与发展, 信息传递与交流方式日益多样, 信息量也呈指数式增长。不法分子凭借便捷的传播途径并以文本、图片、视频、音频等媒介为载体进行不良信息传播, 对国家网络空间安全造成严重的危害[1]。因此, 针对这些媒介的内容过滤系统应运而生。文本作为互联网信息传递与信息交流的重要载体, 在互联网中的使用最为广泛, 针对文本信息过滤的研究已然成为信息过滤领域中的研究重点。信息过滤这一概念最早由 Denning 在 1982 年提出[2], 作者以电子邮件为特例, 描述了如何利用过滤机制对重要邮件和普通邮件进行识别。起初, 针对文本的过滤大多依靠简单的关键词匹配算法[3-5], 这类算法具有较快的过滤速度, 然而并没有考虑到字词与上下文的关联性。此外, 一些文本内容散布者会有意避开使用敏感词, 这便使得基于关键词匹配的信息过滤算法失效。随着自然语言处理技术的发展, 基于内容理解的过滤算法逐渐成为文本过滤的研究重点。研究者开始对文本中的字、词和短语等文本基本单元(即特征项)进行划分, 并进行一系列特征提取操作以获取具有判别文本内容性质的特征集[6-7], 主要步骤包括: 文本特征项分离、文本特征项量化, 以及文本特征项筛选与提取。面对海量文本特征集合, 学者们纷纷提出了不同的文本内容匹配模型。早些年, 贝叶斯决策[8-10]、向量空间[11-12]、支持向量机[13-15]等模型被广泛利用于文本过滤领域。现如今, 随着人工智能兴起, 深度学习在文本分析领域发挥着不可或缺的作用, 文本信息过滤领域也因此取得了新的突破[16-17]。以上文本过滤方案虽然在一定程度上净化了网络环境, 然而, 互联网中仍存在着信息过滤系统误判或者恶意拦截的问题, 例如, 社交网络平台之间的链接屏蔽。这类拦截问题严重影响了用户体验, 损害了用户权益, 扰乱了市场秩序。目前, 在中文文本信息过滤领域中, “火星文”在躲避关键词屏蔽方面效果显著, 例如, 淘宝、拼多多等电商平台为避免商品分享链接被 QQ、微信等社交软件拦截而使用“火星文”, 如图 1(a)所示。然而, 随着检测技术的发展, 这类利用“火星文”的分享链接仍然存在着被拦截的可能[18]。针对图 1(a)所示的分享链接, 技术人员可通过对口令“￥PKMAcWP9cyX￥”进行模式匹配从而拦截此类关键信息, 这是由于其呈现模式通常具有一定的规律性, 易被机器检测。为了有效应对这类文本信息拦截问题, 本文设计出了一个针对性的文本信息隐藏系统。

信息隐藏作为保障信息传递安全的一种重要技术手段, 对国家网络空间安全具有重要意义。目前, 传统中文文本的信息隐藏可分为三类, 即基于文本图像的算法[19-21]、基于文本格式的算法[22-29], 以及基于文本内容的算法, 其中基于文本内容的算法又可分为基于语义的算法[30-31]、基于语法的算法[32-33], 以及基于汉字结构特征的算法[34-38]。近些年来, 凭借文本生成进行信息隐藏的算法逐渐成为主流[39-45]。然而, 以上文本信息隐藏算法所生成的含密文本大多以传统平面媒介语言形式呈现, 因此, 并不适用于规避文本内容拦截, 并且这些算法大多基于对原始文本进行微小的修改, 存在着嵌入容量小鲁棒性弱的缺点。因此, 如何设计出一个能够规避文本信息拦截且具有良好隐写性能的文本隐写系统便成为亟待解决的问题。



(a) 社交软件中的“火星文”

(b) 视频内容平台中的“火星文”

图 1 网络中的“火星文”

Figure 1 “Martian” in the Internet

目前, 尚未存在成熟的技术手段可有效过滤含有“火星文”的文本。在研究过程中我们发现, 相较

于传统平面媒介语言形式，“火星文”具有更多的冗余空间。因此，本文将“火星文”作为隐写载体，提出了一种可规避文本信息拦截的“火星文”生成隐写系统。该文本隐写系统由预处理、控制以及隐写三大基本模块组成。通过对汉字结构特征的研究以及“火星文”构字方式的分析，本文设计出 6 种隐写子模块以供信息嵌入与提取。本文主要贡献有:

(1) 针对互联网中文本信息误拦截以及恶意拦截问题提供了一个实际的解决方案。

(2) 将“火星文”作为隐写载体，提出了基于“火星文”生成的文本隐写系统。该隐写系统较与同类型文本信息隐藏方案在嵌入容量上提高了 52%，中文单字符嵌入容量可达 1.87 比特，且与基于文本生成的算法相比，该隐写系统又具有较强的鲁棒性。

本文的剩余内容如下: 第二节将介绍中文文本信息隐藏，“火星文”以及汉字的编码的相关知识; 第三节描述了“火星文”生成隐写系统的大致框架，并简单介绍信息嵌入流程与信息提取流程; 第四节将详细介绍该隐写系统各个模块及子模块的功能; 第五节将对实验结果进行分析; 最后，在第六节进行总结并展望未来的研究。

2 预备知识

2.1 中文文本信息隐藏

信息隐藏作为保障信息传递安全的一种重要技术手段，其利用人类感官冗余与载体数据冗余，将信息以特定方式嵌入至所选载体中，从而实现隐蔽通信[46]。如今，以图片、视频、音频为载体的信息隐藏研究已取得不少学术成果。然而，文本信息隐藏研究显得相对滞后，其中针对中文文本信息隐藏的研究更是少之又少。虽然与图片、视频、音频等载体相比，文本存在着信息冗余度低、数据量少的缺点，但作为互联网信息传递与交流的重要载体，以文本为载体的信息隐藏仍具有一定的研究价值。在中文文本信息隐藏中，部分研究人员将文本存储为图像格式，从而对文本图像进行隐藏操作。Zhao 等[19]通过改变每行文本图像中上下两半部分黑色像素之和的比值进行隐藏操作。该嵌入算法具有较强的稳健性，然而嵌入容量受限于文本行数。Qi 等[20]针对扫描打印文本，通过翻转字符图像黑色点来进行信息嵌入，此算法具有较高的隐蔽性，然而嵌入容量低。Tan 等[21]利用傅里叶描述子，通过修改扫描文本图像中字符边界点来进行信息嵌入，进一步提升了隐写容量和鲁棒性。由于对文本图像进行信息隐藏操作并没有充分利用到文本的属性，因此，部分研究学者开始从文本内容的组织方式中探寻冗余空间。其中,一部分算法利用 Word[22]、PDF[23]、XML[24]等文档中未使用空间进行信息嵌入，一部分算法通过修改字、行和段在文本中的排版间距进行信息隐藏，如文献[25]。然而，这类算法对文件格式的依赖性极高，且隐藏位置暴露的概率较高。除此之外，一些研究者利用字符的属性进行信息操作，如: 字符编码[26] 、字符色彩[27]、字符不可见性[28-29]等，这类算法的嵌入容量普遍高于基于调整文本间距的隐藏算法，然而伴随着隐蔽性差的缺点。随着自然语言处理技术的发展，中文文本信息隐藏逐渐关注文本的内容。其中基于语义的信息隐藏技术主要以同义词替换为主[30-31]。这类算法的嵌入能力往往取决可供替换的词汇数量。针对语法的文本信息隐藏主要从句式以及词性着手进行信息隐藏[32-33]。这些算法具有较强的隐蔽性，然而嵌入能力较差。由于基于语义与语法的信息隐藏算法普遍依赖于自然语言处理技术，在嵌入操作前往往会对文本中的字、词以及句进行分析，从而产生复杂的计算过程。因此，部分学者转而利用中文汉字的特征进行信息嵌入。Sun 等[34]利用文本中左右结构形式作为嵌入位置，Wang 等[35]在 Sun 的基础上增加了上下结构汉字的形式作为信息嵌入点。文献[36]通过简繁体的交错使用进行信息嵌入，并提出了简单替换嵌入算法、高效替换嵌入算法以及基于模板的嵌入算法。文献[37]提出一种基于多音字的文本水印方法，文献[38]给出了一种基于汉字笔画的文本水印算法。

以上算法的隐藏操作普遍依赖于原始文本，在文本信息隐藏领域中，存在着不依赖于原始文本的隐藏算法，即文本生成隐写。在中文文本生成隐写算法中，以诗词为隐写载体的构造式文本信息隐藏成为了研究热点。Yu 等[39]最先将宋词作为隐写载体，其提取不同的宋词格律模板并根据待嵌信息生成含密宋词，然而，所生成的宋词质量不高且信息嵌入率较低。Liu 等[40]对文献[39]所提出的算法做出改进，提高了信息嵌入率，然而宋词生成质量仍然不佳。为了提高所生成诗词的质量, Luo 等[41]提出了基于马尔科夫链的宋词生成模型。随着深度学习的发展，基于深度神经网络的诗词生成隐写模型也被设计出[42-43]，这进一步提高了生成质量以及信息嵌入率。此外，Yang 等[44]设计出一种基于循环神经网络的语言隐写术，其可根据待嵌秘密比特流自动生成英文文本，同时也适用于中文。然而，所生成的文本质量并不能完全保证隐写的安全性，因此，Yang 等[45]又提出基于变分自编码器文本隐写方案，进一步提升了隐写

文本的隐蔽性和安全性。

2.2 隐写载体——“火星文”

“火星文”作为网络语言的一种存在形式, 是网络语言发展到一定阶段的产物, 其普遍存在于互联网中[47]。早期的“火星文”是社会青年群体为追求个性、新颖而设计出的语言符号, 如图 1(b)所示。现如今, “火星文”又增添了新的用途, 其被广泛用于躲避关键词屏蔽, 例如淘宝、拼多多等网购零售平台为避免商品分享信、QQ 等社交软件拦截而使用“火星文”, 如图 1(a)所示。“火星文”字符种类繁多, 构字复杂多样。目前, “火星文”常见的构成方式有如下 5 种: 网络符号构成、数字组合构成、拼音字母构成、简繁体汉字构成, 以及生造字构成。由于人具有认知推理能力, 在“火星文”中, 任意常规汉字可以由多种方式进行表示且不影响人的理解, 例如汉字“皓”, 可被转化为拼音“hao”, 也可被拆分成汉字“白”与“告”, 亦或是被替换偏旁生成“浩”。因此, 较于传统平面媒介的语言形式, “火星文”具有更高的信息冗余度, 也存在更大的空间可供秘密信息嵌入。

图 2 汉字部件编码

Figure 2 Chinese character component codes

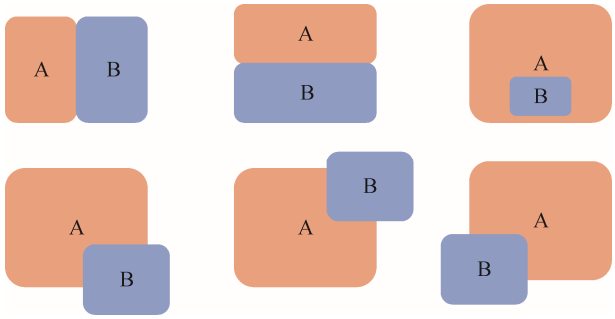


图 3 汉字部件的 6 种空间关系

Figure 3 Six spatial structures between Chinese character components

啊	38lr422lr148	阿	422lr148	艾	313ud300
挨	430lr(308ud170)	哎	38lr(313ud300)	哎	38lr(308ud170)
哀	304ud38ud354	皑	173lr(53ud51)	蔼	313ud(419lr261)
鞍	244lr(330ud33)	氨	110ru(330ud33)	安	330ud33
岸	52ud(7lu36)	案	330ud33ud86	暗	102lr(134ud87)
凹	161	敖	(335ud302)lr45	盎	152ud153
按	430lr(330ud33)	袄	470lr108	昂	97ud(426lr421)

图 4 汉字编码

Figure 4 Chinese character codes

2.3 汉字编码

对汉字进行有效的编码可使汉字的操作与处理更为简便。Sun 等[48]提出了一种汉字编码方式, 使汉字能够以简便的数学形式呈现。不失一般性, 设 Ω 为汉字集合, Θ 为汉字部件集合, Ξ 为 Θ 中任意两个汉字部件位置关系集合, 则存在如下两种情况:

- Θ = {“点”, “横”, “竖”, “撇”, “捺”, “折”, “钩”, “提”}, 则 $\Omega=(\Theta,\Xi)$。
- $\Theta=\Omega$, $\Xi=\varnothing$, 则 $\Omega=(\Theta,\Xi)$。

第一种情况下, Θ 为 8 种汉字基本笔画, 构造最为简单, 任意汉字部件可由此 8 种笔画组合而成。然而, 汉字部件繁杂, 组合方式繁多。因此, 位置关系集合 Ξ 便极其复杂。第二种情况下, Ξ 为空集, 则 Θ 为整个汉字集合 Ω, 这便使得 Θ 中的元素过多。为了平衡 Θ 与 Ξ 构造时的复杂度问题, 作者对汉字的组成部件进行了统计分析, 从中选取了 505 个汉字部件作为集合 Θ 的元素, 图 2 展示了部分汉字部件的编码。针对此 505 个汉字基本部件, 作者进而定义了 6 种汉字部件空间位置关系作为集合 Ξ 的元素, 如图 3 所示。将 Θ 中的元素作为操作对象, Ξ 中的元素作为操作符, 根据表 1 所提供的符号优先级以及运算方向, 则每个汉字都有其唯一的编码形式。本文选取了 2500 个常用简体汉字, 记为 Ω_{sc}, 其对应的繁体字集合为 Ω_{sc}, 且有 $\Omega_{sc}\cup\Omega_{tc}=\Omega$。图 4 展现了部分所选简体字表达式。

表 1 算符优先级表

Table 1 Operator priority table

操作符	优先级	运算方向
( )	1	
we, lu, ld, rd	2	从左至右
lr, ud	3	从左至右

3 文本隐写系统框架

本文提出的隐写系统主要由三大模块组成: 预处理模块、控制模块, 以及隐写模块。根据“火星文”的构字方式, 本文设计出 6 种隐写子模块, 分别为: 简繁体转换模块(s-t transformation, S-TT)、字音转换模块(c-p transformation, C-PT)、字体重构模块(font reconstruction, FR)、同音字替换模块(homophone substitution, HS)、字体拆分模块(font separation, FS)以及非汉字字符替换模块(non-Chinese character substitution, N-CCS), 如图 5 所示。根据以上模块的设计该文本隐写系统的信息嵌入过程如图 6(a)所示。首先选取原始文本, 与此同时, 将待嵌信息输入至预处理模块进行数据规范化。将原始文本与预处理后的待嵌信息作为控制模块的输入, 控制模块会根据待嵌信息的内容并结合输入字符的特征将隐写任务分配至相应的隐写子模块中。最后, 隐写子模块执行信息嵌操作, 并将反馈信息传递至控制模块以便执行下一步信息嵌入操作。图 6(b)描述了该文本隐写系统信息提取流程。首先获取含密文本“火星文”, 判断当前含密字符的类型。控制模块会根据含密载体字符的类型将信息提取任务分配至相应的隐写子模块中。最后,隐写子模块执行相应的信息提取操作,并将反馈信息发送至控制模块, 以便控制模块执行下一步信息提取操作。其中, 隐写模块中的标号表示为其所对应的隐写子模块, 控制模块在信息嵌入与提取过程中的任务分配细节将在第四节讨论。

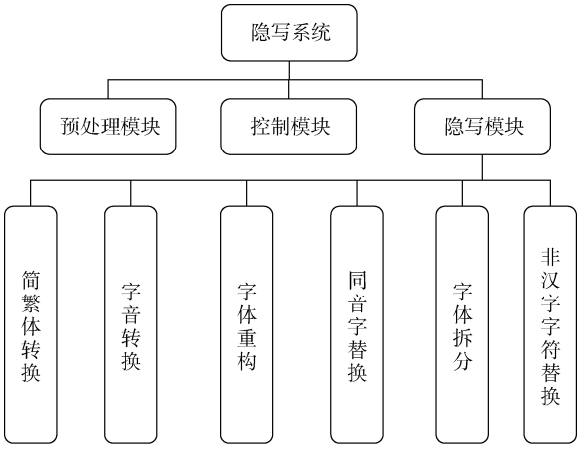


图 5 “火星文”隐写系统框架

Figure 5 Steganographic framework of “Martian” system

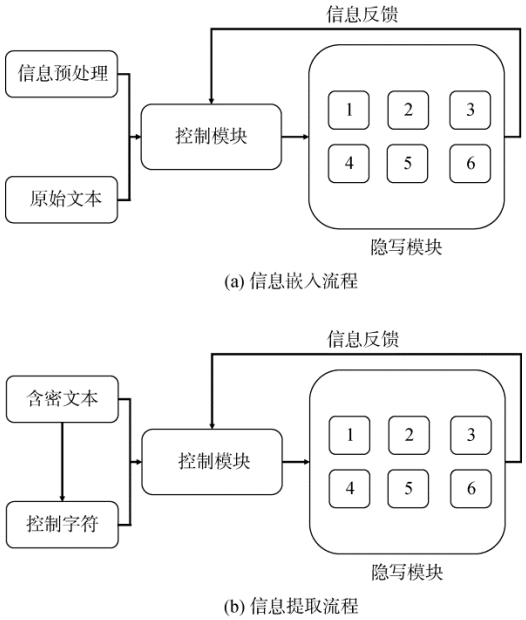


图 6 信息嵌入与提取流程图

Figure 6 Flow chart of information embedding and extraction

4 隐写系统模块设计

此节将详细描述该文本隐写系统的各个模块的功能以及算法实现。不失一般性, 设英文字符集 E, 数字集合为 N, 标点符号集合为 Ψ, 原始输入 $C=\{c_1,\cdots,c_n\}$, 其中 $c_i\in E\cup N\cup\Psi\cup\Omega_{sc}$, 待嵌信息为 M, $M=\{0,1\}^q$, q 为待嵌消息的长度, 含密文本为 S, $S=\{s_1,\cdots,s_d\}$, 其中 s_i 为“火星文”字符。值得注意的是, 标点符号不进行信息嵌入与提取处理。

4.1 预处理模块

该模块主要负责以下两个部分工作:

4.1.1 辅助信息增添

信息嵌入过程中, C 中字符可能会遗留部分未被使用, 这会给信息提取带来操作上的不便。为了解决此问题, 本文采取如下步骤:

Step 1: 计算 M 的长度 l。

Step 2: 将 l 转换为二进制 l_b, 若 l_b 不足 λ 位, 则采取高位补 0。

Step 3: 生成预处理信息 M', $M'=l_b\cup M$。

4.1.2 待嵌信息加密

为了保证秘密信息内容的隐私性, 首先应对待嵌信息 M'进行加密, 得到密文信息比特流 M_e,

$$M_e=\mathrm{Enc}_K(M',P) \tag{1}$$

其中, K 为密钥, P 为辅助参数。所选择的加密方案只需满足如下等式即可

$$M'=\mathrm{Dec}_K(\mathrm{Enc}_K(M',P),P) \tag{2}$$

其中, λ 是预先设置的参数, 具体细节将在第五节讨论。

4.2 控制模块

此模块主要负责解决信息嵌入以及信息提取过程中任务分配的问题, 任务分配细节如下。

4.2.1 信息嵌入任务分配

为解决简繁体字信息提取过程中的混淆问题, 例如简繁体转化模块与字体重构模块都有可能生成简体字, 则必定会给信息提取带来困扰, 因此, 须在每一次隐写操作过程中引入一个隐写控制符。本文采用了 Unicode 不可见控制字符(零宽度字符)[29]。由于原始字符 $c_i \in E \cup N \cup \Psi \cup \Omega_{sc}$, 这些不可见控制字符并不会对文本的显示造成太大的影响。为了增加隐写容量, 使引入的不可见控制字符也能携带信息, 我们对其进行了编码。表 2 显示了不可见控制字符 $C_{invisible}$ 的编码和所属类别的标号, 并给出所适用的隐写子模块。然而, 使用以上特殊 Unicode 字符存在着隐写安全性问题, 在第五节中, 我们将对其进行分析, 并给出一种可行的解决方案。在每轮信息嵌入中, 信息嵌入方式取决于随机数 α 的值, $\alpha \in (0,1)$。为了表达简便, 设当前待嵌字符为 c_i, 待嵌信息比特为 m_j, $Module_k$, $k=1,\cdots,6$, 对应 6 个隐写子模块且接受参数 c_i 与 $C_{invisible}$, $C_{invisible}$ 的取值由表 2 与待嵌比特串 $\{m_j, m_{j+1}\}$ 得出, I_1、I_2 和 I_3 分别表示表 2 中第一类、第二类与第三类 Unicode 不可见控制字符集合, 设 *fs* 为反馈信号。例如, 当 α 的值落入 $Module_3$ 也就是字体重构模块的判定域中, 则控制模块会选取 $Module_3$ 进行信息嵌入操作。由表 2 可知, $Module_3$ 适用于第 2 类不可见控制字符, 因此, 控制模块根据 $\{m_j, m_{j+1}\}$ 的值选取对应的不可见控制字符 $C_{invisible}$ 并将其作为参数与 c_i 一并传入至 $Module_3$。图 7 左部分展示了信息嵌入任务分配流程。参数 ε_i 的设置将在第五节讨论。

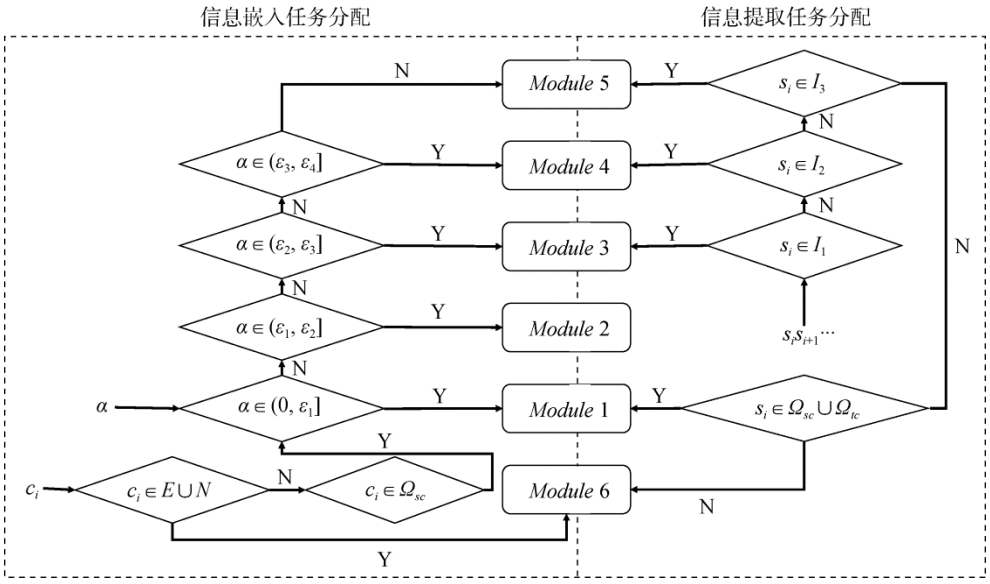


图 7 信息嵌入与提取任务分配流程

Figure 7 Allocation processes of information embedding and extraction task

4.2.2 信息提取任务分配

在信息提取过程中, 控制模块首先判别含密字符 s_i 的字符类型, 若为不可见控制字符, 则可根据表 2 查找其所属类别和对应的隐写子模块 $Module_k$, 继而将 s_i 作为参数分配至该隐写子模块。否则, 将根据 s_i 的是否为汉字进行信息提取时的任务分配。图 7 右部分展示了信息提取时的任务分配流程, 由图可知, 无论 s_i 的字符类型如何, 其都会作为参数传入至 $Module_k$。由于控制模块已经对 s_i 做出判断, 对于各个隐写子模块 $Module_k$ 而言, 所接受的字符 s_i 的类型是已知的。

4.3 隐写模块

“火星文”字符种类繁多且构字复杂多样, 本文通过对汉字结构特征的研究以及“火星文”构字方式的分析, 设计出如下 6 种隐写子模块:

4.3.1 简繁体转换模块

Sun 等[36]提出了三种针对简繁体的信息嵌入方式, 即简单替换嵌入算法(SSE)、高效替换嵌入法

(ESE)以及基于模板的嵌入算法(TBE)。为了保证较高的嵌入率, 同时又兼容其他隐写子模块, 本文参考 SSE 算法, 并设计字典 $D:\Omega_{sc}\to\Omega_{tc}$, 信息嵌入过程如下:

Step 1: 若待嵌信息 $m_j=0$, 则含密字符 $s_k=C_{invisible}\cup c_i$, $fs=\text{True}$, 否则执行步骤 2。

Step 2: 从字典 D 中找到 c_i 对应的繁体字 tc_i, 若 $c_i\neq tc_i$, 则 $s_k=C_{invisible}\cup c_i$ $fs=\text{True}$, 否则嵌入失败, $fs=\text{False}$。

信息提取过程中, 若当前含密字符 $s_i=\Omega_{tc}$, 则提取信息 $m_j=1$, 否则, $m_j=0$。

表 2 隐写控制字符表

Table 2 Steganographic control character table

标号	Unicode 字符	编码	适用子模块
1	U+200B	00	FR
	U+200C	01	
	U+200D	10	
	U+200E	11	
2	U+202A	00	HS
	U+202B	01	
	U+202C	10	
	U+202D	11	
3	U+206A	00	FS
	U+206B	01	
	U+206C	10	
	U+206D	11	
4	ϕ	ϕ	S-TT,C-PT,N-CCS

表 3 相似字符表

Table 3 Similar character table

00	01	10	11
a	A	α	_
b	B	β	
⋮	⋮	⋮	⋮
1	①	Y	|
2	②	γ	γ
⋮	⋮	⋮	⋮
0	◦	O	⊙

4.3.2 字音转换模块

简体字与拼音的转化本质为简体字与英文字符的替换, Rizzo 等[49]利用 Unicode 协会所提出的“混淆”字符表进行信息嵌入, 该方案对字符的外形相似度具有较高要求。然而, “火星文”对字符的相似度约束更加宽泛, 只需满足外形相似或者语义相似即可。本文并没有对相似字符表中的具体内容做出约束, 只需满足通信双方的约定即可, 表 3 显示了部分相似字符以及对应编码,算法 1 展示了信息嵌入过程。

算法 1. 字音转换信息嵌入算法.

输入: $c_i, m_j, C_{invisible}$,待嵌信息长度 $l_{message}$

输出: fs.

1: $CP_i=\text{Phonetic\_alphabet}(c_i)$; //获取$c_i$的拼音

2: $l=\text{Length}(CP_i), M_{substr}=\{m_j,\cdots,m_{j+2l}\}$;

3: IF THEN

4: ω=1;

5: WHILE $\omega<2l$ DO

6: 根据 $\{M_{substr}[\omega], M_{substr}[\omega+1]\}$ 从 Table 3 中查找与 $CP_i[\lceil\omega/2\rceil]$ 相似的字符 $simc$;

7: $CP_i[\lceil\omega/2\rceil]=simc;\omega=\omega+2;$

8: ENDWHILE

9: $s_k=C_{invisible}\cup CP_i;\ fs=\text{True}$;

10: RETURN fs.

对于信息提取, 只需在表 3 中查找当前含密字符 s_i 所对应的编码即为提取信息 m_j。

4.3.3 字体重构模块

在“火星文”中, 字体重构的表现形式分为两种, 第一种为偏旁增添, 如: 打→哳, 第二种为偏旁替换, 如: 嗷→遨, 然而无论是偏旁增添还是偏旁替换, 为了不造成过大的感官差异, 增添或替换的偏旁对于整个汉字的其他部件而言应在结构上显得更加简单, 笔画数更少。根据第二节所介绍的汉字编码知识, 本文采用一种汉字的二叉树表示形式, 图 8 展示了部分汉字的树形结构。为简便表达, 设 $\alpha,\beta\in\Theta^*$, $f(\cdot)$ 为求笔画数函数, $h(\cdot)$ 为求树高函数, 且有如下定义:

定义 1. 若 $h(\beta)-h(\alpha)\leqslant\theta_1$, 定义: $\beta\preccurlyeq_{\mathrm{H}}\alpha$。

定义 2. 若 $f(\beta)-f(\alpha)\leqslant\theta_2$, 定义: $\beta\preccurlyeq_{\mathrm{S}}\alpha$。

其中参数 θ_1、θ_2 的设置将在第五节进行讨论。

本文在 Θ 上建立一个以汉字部件笔画数为索引的字符表 Γ, 并遵从如下两个设计准则:

• 汉字部件笔画数的分类应使得表中每行元素分配尽可能均匀。

• 汉字部件笔画数的分类应使得表中每行元素尽可能多。

考虑到单字符嵌入容量对系统鲁棒性的影响, 本文对 505 个汉字部件笔画数进行了统计, 对 Γ 进行如表 4 所示的设计。算法 2 展示了偏旁增添情况下信息嵌入的过程。算法 3 则适用于偏旁替换情况下的信息嵌入。在具体的信息嵌入操作中, 该隐写子模块会优先执行偏旁增添信息嵌入算法, 若此算法失

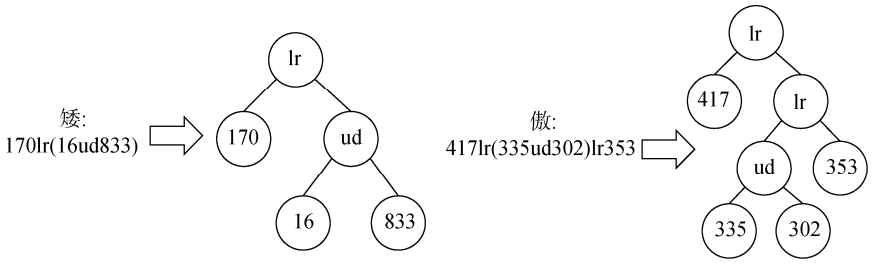


图 8　汉字的树形结构

Figure 8　Tree structure of Chinese characters

败则执行偏旁替换信息嵌入算法。

信息提取过程中, 由于已经知道当前含密字符 s_i 为不可见控制字符, 因此 s_{i+1} 必定为经过偏旁增添或者偏旁替换修改过的汉字, 具体的信息提取过程如算法 4 所示。

表 4　以笔画数为索引的汉字部件表

Table 4　Chinese character component table indexed by stroke number

笔画数	汉子部件	编码
⩽2	一、乙、二、十、…	00
3	广、干、土、士、…	01
4	王、卅、丰、车、…	10
⩾5	瓜、乐、母、央、…	11

算法 2. 偏旁增添信息嵌入算法.

输入: $c_i, m_j, C_{invisible}$,待嵌消息长度 $l_{message}$,汉字部件笔画数索引表 Γ .

输出: fs .

1: $BT = TreeStructure(c_i)$; //获取c_i的树形结构

2: $M_{substr} = \{m_j, \cdots, m_{j+1}\}$;

3: IF $j+1 < l_{message}$ THEN

4: $cmp = \Gamma[M_{substr}]; \omega = 1;$

5: WHILE $\omega < \text{Length}(cmp)$ DO

6: $character = Combine(cmp[\omega], BT, \Xi)$

7: IF $cmp[\omega] \preccurlyeq_{\text{H}} c_i$ THEN

8: IF $character \in \Omega$ THEN

9: $s_k = C_{invisible} \cup character;$

10: RETURN $fs = \text{True}$;

11: IF $cmp[\omega] \preccurlyeq_{\text{S}} c_i$

12: IF $character \in \Omega$ THEN

13: $s_k = C_{invisible} \cup character;$

14: RETURN $fs = \text{True}$

15: $\omega = \omega + 1;$

16: ENDWHILE

17: RETURN $fs = \text{False}$

算法 3. 偏旁替换信息嵌入算法.

输入: $c_i, m_j, C_{invisible}$,待嵌消息长度 $l_{message}$,汉字笔画数索引表 Γ .

输出: fs .

1: $BT_l, BT_r = \text{LRtrees}(c_i)$; //获取 c_i 左右子树结构

2: $M_{substr} = \{m_j, \cdots, m_{j+1}\}$;

3:IF $j+1 < l_{message}$ THEN

4: $cmp = \Gamma[M_{substr}]; \omega = 1;$

5: IF $BT_l \preccurlyeq_{\text{H}} BT_r$ OR $BT_r \preccurlyeq_{\text{H}} BT_l$ THEN

6: $complex = BT_r \preccurlyeq_{\text{H}} BT_l \,?\, BT_l : BT_r$;

7: WHILE $\omega < \text{Length}(cmp)$ DO

8: $character = Combine(cmp[\omega], complex, \Xi)$;

9: IF $character \in \Omega$ THEN

10: $s_k = C_{invisible} \cup character$;

11: RETURN $fs = \text{True}$;

12: $\omega = \omega + 1$;

13: ENDWHILE

14: $complex = BT_r \preccurlyeq_{\text{S}} BT_l \,?\, BT_l : BT_r$;

15: $\omega = 1;$

16: WHILE $\omega < \text{Length}(cmp)$ DO

17: $character = Combine(cmp[\omega], complex, \Xi)$

18: IF $character \in \Omega$ THEN

19: $s_k = C_{invisible} \cup character$;

20: RETURN $fs = \text{True}$;

21: $\omega = \omega + 1$;

22: ENDWHILE

23: RETURN $fs = \text{False}$.

算法 4. 字体重构信息提取算法.

输入: s_i, 隐写控制字符表 Table 2, 汉字部件笔画数索引表 Γ.

输出: fs.

1: $controlcode = \text{Getcode}(s_i, \text{Table 2})$; //从表 2 中获取$s_i$所对应的编码

2: $BT_l, BT_r = \text{LRtrees}(s_{i+1})$;// s_{i+1} 左右子树结构

3:IF $BT_l \preccurlyeq_{\text{H}} BT_r$ OR $BT_r \preccurlyeq_{\text{H}} BT_l$ THEN

4: $easy=BT_r \preccurlyeq_{\text{H}} BT_l ? BT_r : BT_l$;

5: $code = Getcode(easy, \Gamma)$; //获取简单组件所对应的编码

6: $m_j = controlcode \cup code; fs = \text{True};$

7: ELSE

8: $easy=BT_r \preccurlyeq_{\text{S}} BT_l ? BT_r : BT_l$;

9: $code = Getcode(easy, \Gamma)$;

10: $m_j = controlcode \cup code; fs = \text{True};$

11: RETURN fs ;

4.3.4 同音字替换模块

同音字替换普遍存在于“火星文”中, 且所替换的同音字可为简体字也可为繁体字。本文在 $\Omega=\Omega_{sc} \cup \Omega_{tc}$ 的基础上建立了一张同音字表 Σ 以供信息嵌入与提取。对于每个同音字集合, 本文提出一种完全二叉树字符编码方案, 图 9 展示“rèn”集合中汉字的编码方式。值得注意的是, 对于一个完全二叉编码树, 汉字的编码长度是变化的, 并且汉字只存储在叶子节点, 其他节点存储的数据都为空。为了解决由一字多音所造成的信息提取失败问题, 本文在设计同音字表时, 保证相同汉字不会在多行出现。表 5 展示同音字表的部分信息, 信息嵌入算法如算法 5 所示, 算法 6 描述了 Gethomophone 函数。

算法 5. 同音字替换信息嵌入算法.

输入: $c_i, m_j, C_{invisible}$.

输出: fs.

1: $CP_i = \text{Phoneticalphabet}(c_i)$; //获取$c_i$的拼音

2: $ht = \text{TreeStructure}(CP_i)$; //获取编码树

3: $hp = \text{Gethomophone}(ht, m_j)$;

4: $s_k = C_{invisible} \cup hp$;

5: RETURN $fs = \text{True}$;

算法 6. Gethomophone

输入: 树根节点 $root$, 待嵌比特 m_j.

输出: 同音汉字 hp.

1: Gethomophone($root, m_j$)

2: IF $root$ THEN

3: IF $m_j == 0$ THEN

4: $hp = \text{Gethomophone}(root->LeftNode, m_{j+1})$;

5: IF $m_j == 1$ THE

6: $hp = \text{Gethomophone}(root->rightNode, m_{j+1})$;

7: RETURN $hp = hp \cup root->data.$

信息提取过程中, 当接收到含密字符 s_i, 则可知 s_{i+1} 为经过同音字替换模块处理过的汉字, 具体信息提取步骤如下:

Step 1:据表 2, 获取 s_i 的编码 $controlcode$。

Step 2:获取 s_i 的拼音 SP_{i+1}。

Step 3:在表 5 中查找 SP_{i+1} 所在行, 在该行检索 s_{i+1}, 并获取 s_{i+1} 的二叉树编码 btc。

Step 4: 提取信息 $m_j = controlcode$, $m_j = controlcode \cup btc$, 并将反馈信号 fs 设置为 True 返回至控制模块。

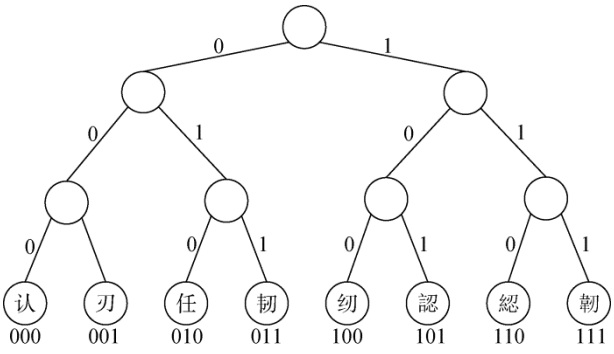


图 9 同音字完全二叉编码树

Figure 9 Complete binary coding tree for homophones

表 5 同音字表

Table 5 Homophone table

拼音	汉字组
àn	(000)岸 (001)案 (100)按 (11)黯...
⋮	⋮
rèn	(000)认 (001)刃 (010)任 (011)韧...
⋮	...
zuò	(00)做 (01)作 (10)座 (11)坐 ...
⋮	⋮

4.3.5 字体拆分模块

根据研究发现, 火星文中大部分的拆分字都是

左右形式和上下的形式, 其中以左右形式的拆分居多, 如: 行→彳亍, 好→女子。由于上下形式的拆分已经破坏汉字视觉上的结构, 一定程度上影响读者的理解, 因此本文仅考虑左右结构汉字的拆分, 文献[34-35]都提出了针对左右结构汉字的信息嵌入方案。然而, 在信息提取过程中, 需要生成原始样本作为参考, 因此操作复杂, 不适用于本文提出的文本隐写系统。具体信息嵌入过程如算法 7 所示。

算法 7. 字体拆分信息嵌入算法.

输入: $c_i, m_j, C_{invisible}$.

输出: fs.

1: $BT = TreeStructure(c_i)$; //获取 c_i 的树形结构

2: $root$ = Getrootnode(BT); //获取树的根节点

3:IF $root ==$ "lr" THEN

4: IF $m_j == 0$ THEN

5: $s_k = c_i$;

6: ELSE

7: c_{il}, c_{ir} = SplitCharacter($c_i, root$); // 拆分字体得到左右子树结构c_{il}与c_{ir}

8: $s_k = C_{invisible} \cup c_{il} \cup c_{ir}$; fs = True;

9:RETURN fs.

对于信息提取, 当该隐写子模块接收含密字符 s_i, 则可根据表 2 获取 s_i 对应的编码$controlcode$。同时我们也知 s_{i+1} 与 s_{i+2} 为一个汉字的左右两个部分, 依据字体拆分模块的嵌入算法, 最后可得 $m_j = \{controlcode, 1\}$。

4.3.6 非汉字字符替换模块

信息嵌入过程中, 该子模块主要利用相似字符表 3 对原始生成样本中英文字符或者数字字符采取相似字符替换进行信息嵌入。信息提取过程中, 该模块会查找当前含密字符在相似字符表中对应的编码, 并做出相应的信息提取操作。具体的嵌入与提取算法与字音转换模块所提供的方法相似, 此处不再赘述。

5 实验结果与分析

本节首先介绍度量指标的设计, 接着对参数设置进行讨论, 并依据设定的参数进行实验。在实验结果基础上, 进一步分析该文本隐写系统鲁棒性。

5.1 度量指标

不失一般性, 设嵌入率为 ER, 则

$$\mathrm{ER} = \frac{M}{C} \tag{3}$$

其中, M 表示嵌入比特, C 表示原始文本字符数, 设嵌入效率为 EE, 则

$$\mathrm{EE} = \frac{M}{CH} \tag{4}$$

其中 CH 为原始文本修改字符数。设文本膨胀率为 TER, 则

$$\mathrm{TER} = \frac{S}{C} \tag{5}$$

其中 S 为含密样本字符数。值得注意的是, 我们并没有利用原始文本比特数与含密文本比特数之比来计算 TER, 这是由于在不同的编码方案中, 字符所占字节数可能并不相同。

文本膨胀会增加信息传输的成本, 因此一个良好的文本隐写系统应在具有较高的嵌入率 ER 与嵌入效率 EE 的同时具有较低的文本膨胀 TER, 本文给出隐写性能衡量指标

$$\mathrm{SPI} = \frac{\mathrm{ER} \times \mathrm{EE}}{\mathrm{TER}} \tag{6}$$

为了与不同类型的文本信息隐藏算法进行比较, 本文给出单字符嵌入容量

$$\mathrm{SCEC} = \frac{M}{\mathrm{TER}} \tag{7}$$

5.2 隐写子模块嵌入能力分析

本文中, 各个隐写子模块的单字符嵌入能力并不相同, 按照前文所设计的表, 我们可以计算出各个隐写子模块的单字符嵌入容量。对于 S-TT 模块, 无论是简体字还是繁体字都只嵌入 1 比特信息, 且没有附加控制字符, 因此 SCEC(S-TT) = 1 比特。对于 C-PT 模块, 依照相似字符表的编码方案, 易得 SCEC(C-PT) = 2 比特。对于 FR 模块, 经过偏旁替换或者增添的汉字能嵌入 2 比特信息, 同时附加的一个控制字符也可嵌入 2 比特信息, 根据公式(7)可得 SCEC(FR) = 2 比特。对于 FS 模块, 由于字体是否拆分取决于当前待嵌比特m_j的取值, 则可根据公式如下公式计算出 FS 模块的单字符嵌入能力:

$$\mathrm{SCEC(FS)} = \sum_{i=0}^{1} \mathrm{SCEC(FS} \mid m_j = i) \cdot p_i \tag{8}$$

其中 p_0 与 p_1 分别表示 $m_j = 0$ 与 $m_j = 1$ 的概率, $\mathrm{SCEC}(X \mid Y)$ 表示在情况 Y 下, 模型 X 的单字符嵌入容量。不妨设 $p_0 = p_1 = 0.5$, $m_j = 0$ 时, FS 模块相当于 S-TT 模块, 因此, $\mathrm{SCEC(FS} \mid m_j = 0) = 1$ 比特。在 $m_j = 1$ 情况下, $\mathrm{SCEC(FS} \mid m_j = 1) = 1.5$ 比特, 根据公式(8)可得 SCEC(FS) = 1.25 比特。对于 N-CCS 模块, 根据前文的嵌入算法, 易得 SCEC(N-CCS) = 2 比特。

HS 模块与以上模块不同, 该模块的嵌入能力与一个汉字的同音字个数有关, 因此, 我们从 Ω_{sc} 随机抽取了 100 个汉字, 并为每个汉字随机生成了 100 个长度为 10 的比特串用以模拟待嵌比特流, 最后对这 10000 组实验结果取均值, 表 6 给出了各个隐写子模块的具体嵌入能力数据。

表 6 各隐写子模块单字符嵌入容量(单位:比特)

Table 6 Single character embedding capacity of each steganographic sub module(bit)

S-TT	C-PT	FR	HS	FS	N-CCS
1	2	2	2.86	1.25	2

5.3 参数设置

若不考虑实际应用, 可只选择单字符嵌入能力强的模块进行信息嵌入。然而, 这样生成的隐写样本不仅影响人的正常理解, 同时也可能暴露系统相关表的信息。为了让本文所提出的隐写系统能够运用于现实生活, 所生成的隐写样本需接近网络中的样本。因此, 在参数设定上, 本文参考了“火星文”在各大网络电商平台“火星文”的使用情况。

5.3.1 预处理模块参数

预处理模块参数λ的取值与待嵌信息长度有关, 通常情况下, 通信双方会约定传递消息的最大长度, 且本文所提出的隐写系统的隐写性能对此参数并不敏感。实验中, 本文设置 $\lambda=7$, 即待嵌消息长度最大为 128 比特。

5.3.2 控制模块参数

由于并不存在“火星文”数据集, 本文从互联网中搜集了 150 个使用“火星文”的电子商务平台口令链接样本 (包含 5672 个字符), 我们对这些样本进行了清洗, 去除了样本中具有传统平面媒介语言形式的部分, 最终获得包含 4896 个字符的火星文样本集。我们人工为每个样本生成对应的普通样本(即具有传统语言形式的样本)。通过比对“火星文”样本与其对应的普通样本, 经统计发现约 53.2%的“火星文”构字方式符合 S-TT 模块, 同时, 在隐写过程中, 使用其余汉字处理模块会造成文本膨胀。因此, 本文认为简繁体转换模块使用概率应大于其余 4 个汉字处理模块。实验过程中发现, 字体重构模块算法复杂度较高, 对其降低使用频率可减少含密文本生成时间, 其余模块的使用概率将参考表 6 生成。本文实验中, 为了尽可能增加隐写容量同时满足以上条件, 我们设置参数: $\varepsilon_1=0.35, \varepsilon_2=0.55, \varepsilon_3=0.65, \varepsilon_4=0.95$。

5.3.3 字体重构模块参数

字体重构模块参数 θ_i 的取值会对信息嵌入成功率造成影响。本文针对已选取的 150 个“火星文”样本, 从中提取出汉字。统计结果显示, 在提取出的汉字中有 346 个汉字进行过字体重构, 本文对这些汉字进行如下的处理:

Step 1: 计算每个汉字树形结构 BT 的左子树 BT_l 与右子树 BT_r 高度差的绝对值。

Step 2: 绘制关于汉字左右子树高度差绝对值的直方图。

实验结果如图 10 所示, 该最小值为 0, 但考虑到 $\theta_1=0$ 会使信息提取算法失效, 因此设置 $\theta_1=1$。在此实验结果上, 再次对这 346 个汉字进行如下处理:

Step 1: 选取左右子树高度差为 0 的汉字并组成新的汉字集合。

Step 2: 在新汉字集上, 计算左右子树所组成汉字部件的笔画之差的绝对值并进行直方图统计。实验结果如图 11 所示, 该最小值为 0, 但考虑到 $\theta_2=0$ 会使信息提取算法失效, 因此设置 $\theta_2=1$。

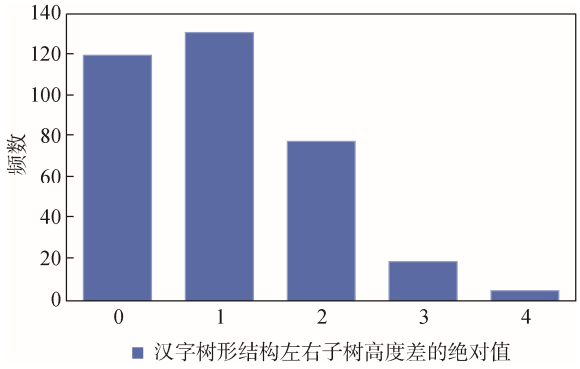


图 10 汉字左右子树高度差的绝对值直方图

Figure 10 Absolute value histogram of height difference between left and right subtrees of Chinese characters

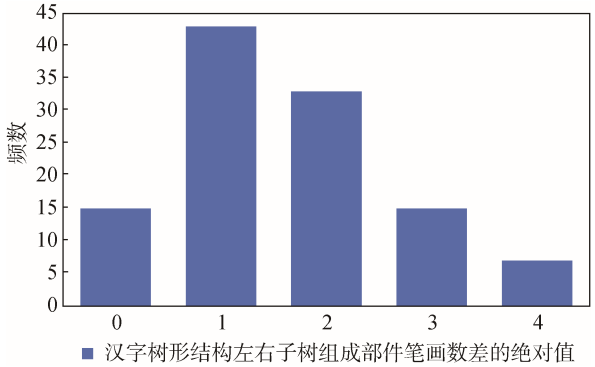


图 11 汉字左右子树组成部件笔画差的绝对值直方图

Figure 11 Absolute value histogram of stroke number difference between left and right subtrees of Chinese characters

5.4 隐写实验

5.4.1 非文本生成隐藏算法对比分析

本文所提出的隐写方案的嵌入操作依赖于对原始文本的修改, 因此仍然属于基于修改的文本隐藏算法。由于文献[34-36]与本文所提出的信息嵌入方案都是对原始文本中文字体进行修改, 在算法设计上具有相似性, 因此, 本文首先对这类算法进行实验。我们从新华网中随机选取了 10 种新闻类别, 并从各个新闻类别中的头条新闻文本中选取了 10 条文本, 其中每条文本包含 200 个字符(不含标点符号)。我们对这 100 条文本进行顺序打乱, 最终获得了一个涵盖多种新闻类别的文本数据集 NewsData。与此同时, 相应的 100 条随机生成的且长度为 400 的二进制串被用以模拟待嵌信息, 其中, 0 与 1 出现的概率都为 50%。SSE、ESE 和 TBE 的各项指标可从理论得出。值得注意的是, ESE 算法的嵌入率 ER 和嵌入效率 EE 与待嵌信息分段长度 L 有关, 根据文献[36]其隐写性能可由如下公式得出:

$$\mathrm{SPI(ESE)}=\frac{\mathrm{ER}\times\mathrm{EE}}{\mathrm{TER}}=\frac{L^2}{2^{L-1}+0.5} \tag{9}$$

对于$L\in N^*$, 易得公式(9)在 L=3 时候取最大值。SSE 算法为 ESE 的一种特殊情况, 即 L=1。TBE 算法的隐写性能与模板长度 T 和选取嵌入字符数 O 有关, 可由如下公式得出:

$$\mathrm{SPI(TBE)}=\frac{\mathrm{ER}\times\mathrm{EE}}{\mathrm{TER}}=\frac{\left\lfloor \log\binom{T}{O}\right\rfloor^2}{\left(\sum_{i=m}^{T}\frac{\binom{i-1}{O-1}}{\binom{T}{O}}\right)\times O} \tag{10}$$

根据文献[36], $T\in\{50,100,200\}$ 时, 当 $T=200$ 且 $O=52$ 时, TBE 算法可以获得最好的隐写性能。由表 7 可知, 本文所提出的文本信息隐藏模型在隐写性能上优于其他基于字体修改的隐藏方案, 且相较于 TBE 算法提升了 52%。本文模型中文单字符嵌入能力表达式如下:

$$\begin{aligned}&\mathrm{SCEC(Ours)}=\varepsilon_1\mathrm{SCEC(S\text{-}TT)}+(\varepsilon_2-\varepsilon_1)\mathrm{SCEC(C\text{-}PT)}+\\&(\varepsilon_3-\varepsilon_2)\mathrm{SCEC(FR)}+(\varepsilon_4-\varepsilon_3)\mathrm{SCEC(HS)}+\\&(1-\varepsilon_4)\mathrm{SCEC(FS)}\end{aligned} \tag{11}$$

根据表 6 以及参数ε_i的设置, 可得SCEC(Ours) =1.87 比特。值得注意的是, 文献[34-35]所提出的隐写方案依赖于 Word 文本, 因此不具有通用性。

除以上基于字体修改的算法外, 其他非文本生成隐藏算法也与本文模型进行了对比。本文参考文献[50], 给出各个算法的嵌入能力与适用文档类型, 如表 8 所示。值得注意的是, 文献[27]通过修改字符的 RGB 值完成单字符 3 比特的嵌入, 然而这类算法不具有通用性, 且正常情况下, Word 文档中大部分字体颜色都相同, 对颜色细微的修改虽然能躲避人类视觉的观测却难以躲避机器检测, 因此该算法隐蔽性较弱。表 8 中, 本文模型的嵌入能力依赖于前文所设计各项表, 在 5.4.2 节中, 我们将更加深入分析本文模型的嵌入能力, 同时给出提升隐写容量的方案。综上, 相较于非文本生成隐藏算法, 本文模型在嵌入能力以及通用性方面都具有明显的优势。

表 7 基于字体修改信息隐藏算法实验结果

Table 7 Experimental results of information hiding algorithm based on font modification

模型	ER	EE	TER	SPI
本文模型	1.96	2.91	1.48	3.85
文献[34]	0.43	2.01	1	0.86
文献[35]	0.54	1.98	1	1.06
SSE[36]	1	2	1	2
ESE[36](L=3)	0.67	3	1	2.01
TBE[36](T=200,O=52)	0.79	3.18	1	2.52

表 8 非文本生成隐写算法对比

Table 8 Comparison of steganographic algorithms for non-text generation

模型	嵌入容量	适用文本类型
本文模型	每个字符嵌入 1.87 比特	无限制
文献[19]	一行嵌入 1 比特	打印扫描文本
文献[20]	一个嵌入字符嵌入 1 比特	打印扫描文本
文献[21]	取决于图像字符边界黑点数	打印扫描文本
文献[25]	调整一个单词或一行嵌入 1 比特	Word 文档
文献[27]	每个字符嵌入 3 比特	Word 文档
文献[26]	每个词嵌入 2 比特	无限制
文献[28]	每个段落嵌入 1 比特	无限制
文献[29]	每句嵌入 2 比特	无限制
文献[22]	取决于未使用的文本属性空间大小	特定 Word 文档
文献[23]	每行嵌入 1 比特	PDF 文档
文献[24]	取决于标签排列组合数	XML 文档
文献[32]	修改一个虚词“的”嵌入 1 比特	无限制
文献[33]	句子长度变换嵌入 1 比特	无限制
文献[30]	受限于矩阵编码	无限制
文献[31]	受限于 LZW	无限制
文献[37]	相邻多音字嵌入 16 比特	无限制
文献[38]	单个 8 笔画汉字嵌入 16 比特	无限制

5.4.2 文本生成隐藏算法对比分析

中文文本生成隐写算法可分为两类: 诗词生成隐写与普通文本生成隐写。在诗词生成隐写领域, 相比于其他算法, Qin 等[43]模型所生成的绝句诗具有较好的可读性和较高的嵌入率。在文献[43]中, 一首绝句诗的嵌入容量由如下公式得出:

$$\begin{aligned}\mathrm{EC(Qin)} &= \log_2(K_{num}) + \log_2(P_{num}) + \log_2(R_{num}) \\ &+4\log(beam_{num})\end{aligned} \tag{12}$$

其中, K_{num}、P_{num}、R_{num}、$beam_{num}$ 分别表示主题词数、模板类数、韵律类数、候选诗句数。因此, 其单字符嵌入能力可由如下公式计算得出:

$$\mathrm{SCEC(Qin)} = \frac{\mathrm{EC(Qin)}}{4N_p} \tag{13}$$

其中, N_p 表示绝句诗每行中文字符数, 标点符号未纳入考虑范围。文中, 给定 $P_{num}=8$ 和 $R_{num}=36$ 且 $N_p \in \{5,7\}$, 因此, SCEC(Qin) 主要由 K_{num} 与 $beam_{num}$ 决定, $beam_{num}$ 设置过大会影响绝句诗的生成质量, 然而, K_{num} 的取值并没有限制, 这也是此算法具有较高嵌入率的主要原因。对于本文模型, 根据公式(11), 在参数 ε_i 确定的情况下, 其单字符嵌入能力受限于各个隐写子模块。通过分析发现, 各个隐写子模块的嵌入能力极度依赖于隐写控制字符表(表 2)、相似字符表(表 3)和同音字表(表 5)。在不修改本文模型算法的情况下, 如下方案可提升嵌入能力:

(1) 由于本文只选取了 2500 个简体字和与其对应的繁体字, 这直接影响了同音字表(见表 5)每行元素的数量, 从而影响了 HS 的嵌入能力, 因此可扩充 Ω_{sc}, 增加同音字表中每行中文字符个数, 从而提升 HS 模块的嵌入能力。

(2) 增加相似字符数, 提升 C-PT 模块与 N-CCS 模块的嵌入能力。

(3) 利用网络特殊字符作为隐写控制字符, 根据图 1(a)与图 1(b)所示, 网络特殊符号是“火星文”文本的组成成分之一, 且这些符号并没有使用限制。可利用特殊符号替换不可见控制字符, 增加表 2 中每个类别的候选字符数, 从而增加隐写控制字符的嵌入容量。

无论是基于诗词生成隐写算法还是基于普通文本生成隐写算法, 理论上, 本文模型都可以在不修改模型算法的情况下, 利用如上三种方式, 达到与之接近的嵌入能力。文献[43]所提出的隐写方案在信息提取过程中需要从含密诗中提取主题词、韵律和模板类型, 并将这些信息传入至已经训练好的语言模型中生成候选诗句, 并利用所生成的候选诗句进行信息提取操作。在面对有损信息传输以及篡改攻击, 这些信息, 例如主题词, 有一定的概率丢失或者被篡改, 这将导致整个信息提取过程的失败。同样, 对于文献[44-45], 通过分析其信息提取算法, 第 n 个字 w_n 的信息提取依赖于由概率 $p(w_n \mid w_1,\cdots,w_{n-1})$ 所决定的字候选集, 含密文本中任意一个字 w_i 的丢失与篡改都有可能导致 $p(w_n \mid w_1,\cdots,w_{n-1})$ 概率的改变, 从而可能导致后续提取任务失败。然而, 本文隐写子模块的信息嵌入与提取操作都是相互独立的, 任意一个模块的提取失败都不会影响到其他模块的信息提取, 这也是本文模型相较于文本生成隐写算法的优势。

5.4.3 “淘口令”生成案例

本文以互联网中流行的“淘口令”作为生成案例, 将重要信息嵌入至“火星文”文本中, 使其不再暴露于外界, 这在一定程度上降低了重要信息被检测的风险, 有效解决了目前互联网中常见的文本信息拦截问题, 保障了信息传递的安全。图 12 展示了含密“淘口令”的部分生成样例。其中, 字符编码采用 UTF-8, 图中半部分利用了不可见控制字符, 下半部分则利用 5.4.2 节所提出的网络特殊符号方案。值得注意的是, 图 12 中的信息嵌入过程并没有使用完整段原始文本, 因此部分含密文本仍旧保留原始字符, 可见本文模型具有较好的嵌入能力, 同时, 这也符合我们日常生活中所看见的淘口令样本。

5.5 隐写安全性分析

本文隐写方案所生成的“火星文”与网络中的“火星文”并无差别, 因此, 具有一定的隐蔽性。由于本文隐写载体特殊且所生成的含密“火星文”样本复杂多样, 在攻击者不知晓该隐写系统的编码方案以及相关表的情况下, 做出准确的隐写分析将具有相当大的难度。传统针对语义[51]、语法[52]以及文本图像的隐写分析技术[53]均不适用于本类文本检测。

然而, 本文使用不可见控制字符(见表 2)来辅助隐写, 这类字符的使用在互联网中非常罕见, 倘若隐写分析者仅针对该类字符进行检测, 则可以达到 100%的检测准确率。所幸 5.4.2 节中所提出的第三个嵌入能力提升方案可有效解决这类问题。本文所使用的网络特殊字符部分来自于网站①。由于“火星文”本身复杂多样的特性, 可见隐写控制字符的选择范围将非常广泛, 倘若隐写分析者针对这类特殊字符

①http://www.360doc.com/content/09/0820/17/180960_5085129.shtml

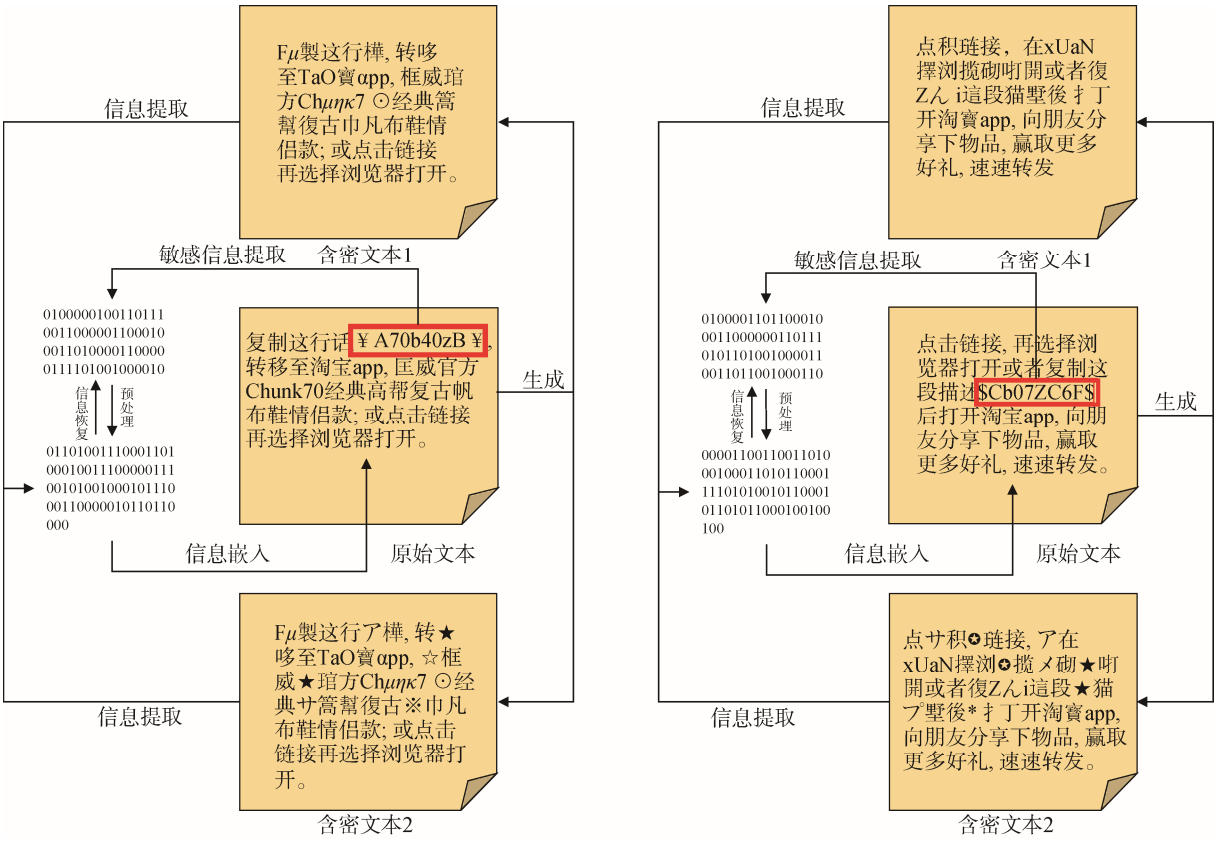


图 12　含密“火星文”文本生成案例

Figure 12　Generation case of “Martian” text with secret data

进行检测, 那么可能会将互联网中良性的“火星文”文本或者普通文本误判为隐写文本, 这是得不偿失的。倘若分析者试图通过统计的方式发掘这类字符的使用规律, 由于本文默认使用的每一类隐写控制字符的字符候选数最多为 4(见表 2), 若观测次数趋于无穷, 那么分析者发掘出隐写控制字符的使用模式的概率将会趋于 100%, 也就是说, 分析者必然能知晓这些控制字符的使用规律从而作出隐写分析判断。针对这类隐写分析攻击, 本文提出一下两种应对措施:

(1) 适当扩大每一类隐写控制字符的字符候选池。表 2 中, 每一类隐写控制字符最多只有 4 个候选字符, 若扩充其候选字符数, 可增加统计隐写分析的难度。

(2) 不定期更新隐写控制字符表。隐写控制字符表的更新主要是为了尽可能模糊隐写控制字符的使用规律。

此外, “火星文”虽然普遍存在于互联网中, 然而其并非为主流的文本语言形式, 因此, 当通信双方利用“火星文”进行消息传递时, 可能会引起第三方的怀疑, 因此, 本文所提出的文本隐写系统并不适用于对文本语言形式要求非常严格的场景。

5.6　鲁棒性分析

5.6.1　文本拦截攻击

“火星文”字符复杂多样, 且呈现方式没有规律性。目前, 市场上尚未有成熟的针对“火星文”拦截的信息过滤软件。与此同时, 本文直接将重要而又敏感的信息嵌入至文本并生成“火星文”, 如此便使得大部分针对内容屏蔽的信息过滤系统失效, 从而有效降低了文本因内容被拦截的风险。为了验证其有效性, 此小节中, 本文随机生成了 100 个不同的“口令”信息并利用图 13 中左部分的广告文本模式生成了 100 个“火星文”文本, 最后利用网易易盾所提供的文本检测接口①进行广告内容检测。图 13 可视化地给出了两个原始广告文本以及对应的两个含密文本的检测结果, 其中“口令”部分为 4 个英文或者数字字符, 编码方式为 UTF-8 且为随机生成。实验结果显示, 该 100 个“火星文”文本里有 18 个文本的检测结果显示为疑似(广告-商业推广), 其余均为通过, 其中, 这 18 个疑似样本中的“淘宝”二字均

①https://dun.163.com/trial/text

只进行了简繁体变化。基于以上结果，可以说明本文所提出的文本隐写系统在一定程度上可以规避文本信息拦截。

5.6.2 文本篡改攻击

若攻击者进行篡改字符恶意攻击，这便使得信息提取准确率大大降低。为了获得所提出文本隐写系统的抗篡改攻击能力，本文首先随机选取了 5 个普通中文文本，其中每个文本含有的字符数大于 24。接着，为每一个普通中文文本分配 30 个不同的待嵌比特串，其中每一个比特串的长度为 32 位。然后，对得到的 150 个含密“火星文”文本进行修正，即去除不含密的字符使得“火星文”文本为满嵌状态，此步骤主要是为了保证攻击者只对含密字符进行篡改。最后，我们对这 150 个满嵌文本进行 1 至 6 字符的随机篡改。图 14 直观展现了不同篡改等级下 150 个含密文本的平均信息提取成功率。可以看出，即使修改了 6 个字符，本文所提出的文本隐写系统的平均信息提取成功率仍然能保持在 70%以上，因此，可以说明，我们的隐写系统具有一定的抗篡改攻击能力。

然而，图 14 也展现出平均信息提取成功率会随着随机篡改次数的增加而近似地线性地下降的现象。为了缓解恶意攻击对该文本隐写系统的影响，本文提出如下三种应对方案:

(1) 纠错编码。若将恶意攻击看作信息传输过程中发生的错误，便可利用纠错码提升信息提取的准确率，从而在一定程度上缓解恶意攻击对隐写系统造成的影响。具体的编码方案需考虑实际应用中纠错码的纠错能力、编码效率以及解码能力，因此在纠错码的选择上，本文没有给出具体的方案。

(2) 冗余字符嵌入。在含密文本中，冗余字符不含有秘密信息，然而对于隐写攻击者来说，任何一个字符都有可能是含密载体。因此，当隐写攻击者对文本进行恶意篡改时，该方案在一定程度上能够缓解恶意攻击对信息提取准确率所带来的影响。然而，一味地增加冗余字符也会增加信息传输成本,并且会大幅度加大信息提取步骤的难度。

图 13 可视化基于不可见隐写控制字符的文本内容检测案例

Figure 13 Visualization of a text content detection case based on invisible steganographic control characters

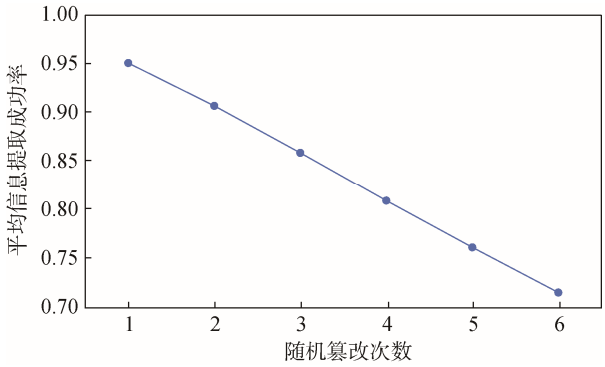


图 14 “火星文”生成隐写系统在不同篡改等级下的平均信息提取成功率

Figure 14 Average information extraction success rate of the“Martian”generation steganographic system under different tampering levels

(3) 单字符嵌入容量控制。当一个字符具有较大的嵌入容量，在面对篡改等恶意攻击时，会使提取信息的准确率大大降低，从而降低文本隐写系统的鲁棒性，因此，本文在设计相关表时，有意减少单字符嵌入容量。

6 总结与展望

本文提出了一种以“火星文”为载体的文本隐写系统，根据实验结果分析，该隐写系统具有较好的隐写性能，在实际生活中，可有效规避文本信息拦截问题，具有一定的应用价值. 据调查，这也是第一个使用网络语言为隐写载体的文本隐写系统。目前本文所提出的文本隐写系统的隐写操作对象只为汉字、英文字符以及阿拉伯数字，在未来的工作中，将会考虑“火星文”中出现的其他字符，设计出更多的隐写模块，同时在参数设置方面，本文将尝试从数值优化角度对隐写参数进行设置，从而进一步提升该文本隐写系统的隐写性能。

参考文献

[1] Sha H Z, Liu Q Y, Liu T W, et al. Survey on Malicious Webpage Detection Research[J]. *Chinese Journal of Computers*, 2016, 39(3): 529-542.
(沙泓州, 刘庆云, 柳厅文, 等. 恶意网页识别研究综述[J]. *计算机学报*, 2016, 39(3): 529-542.)

[2] Denning P J. ACM President's Letter: Electronic Junk[J]. *Commu-*

nications of the ACM, 1982, 25(3): 163-165.

[3] Song H, Dai Y Q. A New Fast String Matching Algorithm for Content Filtering and Detection[J]. *Journal of Computer Research and Development*, 2004, 41(6): 940-945.
(宋华, 戴一奇. 一种用于内容过滤和检测的快速多关键词识别算法[J]. *计算机研究与发展*, 2004, 41(6): 940-945.)

[4] Su G Y, Li J H, Ma Y H, et al. Improving the Precision of the Keyword-Matching Pornographic Text Filtering Method Using a Hybrid Model[J]. *Journal of Zhejiang University Science*, 2004, 5(9): 1106-1113.

[5] Liu Y B, Shao Y, Wang Y, et al. A Multiple String Matching Algorithm for Large-Scale URL Filtering[J]. *Chinese Journal of Computers*, 2014, 37(5): 1159-1169.
(刘燕兵, 邵妍, 王勇, 等. 一种面向大规模 URL 过滤的多模式串匹配算法[J]. *计算机学报*, 2014, 37(5): 1159-1169.)

[6] Peng Y Z, Yuan C G, Wang Y, et al. Studies on Objectionable Information Filtering Technology Based on Contents Understanding[J]. *Application Research of Computers*, 2009, 26(2): 433-438, 447.
(彭昱忠, 元昌安, 王艳, 等. 基于内容理解的不良信息过滤技术研究[J]. *计算机应用研究*, 2009, 26(2): 433-438, 447.)

[7] Liu M Y, Huang G J. Research on Harmful Text Filtering Model Based on Semantic Analysis[J]. *Journal of Chinese Information Processing*, 2017, 31(2): 126-131, 138.
(刘梅彦, 黄改娟. 面向信息内容安全的文本过滤模型研究[J]. *中文信息学报*, 2017, 31(2): 126-131, 138.)

[8] Kim Y H, Hahn S Y, Zhang B T. Text Filtering by Boosting Naive Bayes Classifiers[C]. *The 23rd annual international ACM SIGIR conference on Research and development in information retrieval*, 2000: 168-175.

[9] Ho W H, Watters P A. Statistical and structural approaches to filtering Internet pornography[C]. *2004 IEEE International Conference on Systems*, *Man and Cybernetics*, 2004: 4792-4798.

[10] Fan X H, Sun M S. A High Performance Two-Class Chinese Text Categorization Method[J]. *Chinese Journal of Computers*, 2006, 29(1): 124-131.
(樊兴华, 孙茂松. 一种高性能的两类中文文本分类方法[J]. *计算机学报*, 2006, 29(1): 124-131.)

[11] Li Q, Li J H. Method of Filting Reactionary Text Based on Vector Space Model[J]. *Computer Engineering*, 2006, 32(10): 4-5, 8.
(李强, 李建华. 基于向量空间模型的过滤不良文本方法[J]. *计算机工程*, 2006, 32(10): 4-5, 8.)

[12] Cao Y, He W H. Information Security Filtering System Based on Vector Space Model[J]. *Computer Engineering and Design*, 2006, 27(2): 224-227.
(曹毅, 贺卫红. 基于向量空间模型的信息安全过滤系统[J]. *计算机工程与设计*, 2006, 27(2): 224-227.)

[13] Sun Q, Li J H, Li S H. A Malicious Information Filtering Model Based on One-Class Classification[J]. *Journal of Shanghai Jiao Tong University*, 2005, 39(12): 1993-1996, 2001.
(孙强, 李建华, 李生红. 基于一类分类法的不良信息过滤模型[J]. *上海交通大学学报*, 2005, 39(12): 1993-1996, 2001.)

[14] Kim Y, Nam T. An efficient text filter for adult Web documents[C]. *2006 8th International Conference Advanced Communication Technology*, 2006: 20-22.

[15] Xie X L, Long Z. Implementation of Bad Information Filtering System Based on SVM Algorithm[J]. *International Journal of Security and Its Applications*, 2016, 10(9): 45-54.

[16] Hu W X, Gu Z Q, Xie Y S, et al. Chinese text classification based on neural networks and Word2Vec[C]. *2019 IEEE Fourth International Conference on Data Science in Cyberspace*, 2019: 284-291.

[17] Nismi Mol E A, Santosh Kumar M B. Study on impact of RNN, CNN and HAN in text classification[C]. *2020 Advanced Computing and Communication Technologies for High Performance Applications*, 2020: 94-102.

[18] Guo H L. *The design and implementation of Tao-secret-command serving system to business*[D]. Nanjing: Nanjing University, 2019.
(郭慧玲. 淘客密令系统的设计与实现[D]. 南京: 南京大学, 2019.)

[19] Zhao X Y, Sun J Y, Li L L. Watermarking of Text Images Using Character Step Edge Adjustment[J]. *Journal of Computer Applications*, 2008, 28(12): 3175-3178, 3182.
(赵星阳, 孙继银, 李琳琳. 基于字符阶梯边沿调整的文本水印算法[J]. *计算机应用*, 2008, 28(12): 3175-3178, 3182.)

[20] Qi W F, Li X L, Yang B, et al. Document Watermarking Scheme for Information Tracking[J]. *Journal on Communications*, 2008, 29(10): 183-190.
(亓文法, 李晓龙, 杨斌, 等. 用于信息追踪的文本水印算法[J]. *通信学报*, 2008, 29(10): 183-190.)

[21] Tan L N, Hu K, Zhou X M, et al. Print-Scan Invariant Text Image Watermarking for Hardcopy Document Authentication[J]. *Multimedia Tools and Applications*, 2019, 78(10): 13189-13211.

[22] Yang D M, Guo S. Data Hiding Method Based on Word Document[J]. *Computer Applications and Software*, 2015, 32(5): 314-318.
(杨德明, 郭盛. 基于 Word 文档的数据隐藏方法[J]. *计算机应用与软件*, 2015, 32(5): 314-318.)

[23] Zhong Z Y, Guo Y H, Xu G A. Digital Watermarking Algorithm Based on Structure of PDF Document[J]. *Journal of Computer Applications*, 2012, 32(10): 2776-2778, 2782.
(钟征燕, 郭燕慧, 徐国爱. 基于 PDF 文档结构的数字水印算法[J]. *计算机应用*, 2012, 32(10): 2776-2778, 2782.)

[24] Jie Y. Algorithm of XML document information hiding based on equal element[C]. *2010 3rd International Conference on Computer Science and Information Technology*, 2010: 250-253.

[25] 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'95*, 1995: 853-860.

[26] Zhang H L, Liu D, Wen X Q, et al. Text Information Hiding Algorithm Based on Chinese Characters Coding in Words Platform[J]. *Computer Engineering*, 2010, 36(7): 150-152.
(张洪礼, 刘丹, 温学谦, 等. 基于词平台汉字编码的文本信息隐藏算法[J]. *计算机工程*, 2010, 36(7): 150-152.)

[27] 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*, 2013: 217-220.

[28] 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*, 2016: 1963-1969.

[29] 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.
(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*, 2017, 41(4): 405-411.)

[30] Yang X, Li F, Xiang L Y. Synonym Substitution-Based Steganographic Algorithm with Matrix Coding[J]. *Journal of Chinese Computer Systems*, 2015, 36(6): 1296-1300.
(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.)

[31] Xiang L Y, Li Y, Hao W. Reversible Natural Language Watermarking Using Synonym Substitution and Arithmetic Coding[J]. *Computers Materials and Continua*, 2018, 55(3): 541-559.

[32] 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 Applications*, 2006, 42(3): 158-160.
(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.)

[33] Meng Y J, Guo X P, Zhang W, et al. Text Watermarking Algorithm Based on Sentence Length[J]. *Computer Engineering and Applications*, 2007, 43(32): 52-54, 134.
(蒙应杰, 郭喜平, 张文, 等. 一种基于句长的文本水印算法[J]. *计算机工程与应用*, 2007, 43(32): 52-54, 134.)

[34] Sun X M, Luo G, Huang H J. Component-based digital watermarking of Chinese texts[C]. *The 3rd international conference on Information security - InfoSecu '04*, 2004: 76-81.

[35] Wang Z H, Chang C C, Lin C C, et al. A Reversible Information Hiding Scheme Using Left-Right and Up-down Chinese Character Representation[J]. *Journal of Systems and Software*, 2009, 82(8): 1362-1369.

[36] Sun X M, Meng P, Huang L S. Chinese Text Steganography Based on Character Forms[J]. *Computer Engineering and Design*, 2013, 34(9): 3063-3067.
(孙新梅, 孟朋, 黄刘生. 基于字体的中文信息隐藏算法[J]. *计算机工程与设计*, 2013, 34(9): 3063-3067.)

[37] Fei W B, Tang X H. A Chinese text watermark algorithm based on pOLYPHONE[C]. *Proceedings of 2011 Cross Strait Quad-Regional Radio Science and Wireless Technology Conference*, 2011: 1215-1218.

[38] Tang X H, Wang L N. Text watermarking algorithm based on the stroke of Chinese characters[C]. *2011 International Conference on Multimedia Technology*, 2011: 794-796.

[39] Yu Z S, Huang L S, Chen Z L, et al. High Embedding Ratio Text Steganography by Ci-Poetry of the Song Dynasty[J]. *Journal of Chinese Information Processing*, 2009, 23(4): 55-62.
(余振山, 黄刘生, 陈志立, 等. 用宋词实现高嵌入率文本信息隐藏[J]. *中文信息学报*, 2009, 23(4): 55-62.)

[40] Liu Y C, Wang J, Wang Z B, et al. A Technique of High Embedding Rate Text Steganography Based on Whole Poetry of Song Dynasty[M]. Cloud Computing and Security. Cham: Springer International Publishing, 2016: 178-189.

[41] 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*, 2016, 10(9): 4568-4584.

[42] Luo Y B, Huang Y F. Text Steganography with High Embedding Rate: Using Recurrent Neural Networks to Generate Chinese Classic Poetry[C]. *The 5th ACM Workshop on Information Hiding and Multimedia Security*, 2017: 99-104.

[43] Qin C, Wang M, Si G W, et al. Constructive Information Hiding with Chinese Quatrain Generation[J]. *Chinese Journal of Computers*, 2021, 44(4): 773-785.
(秦川, 王萌, 司广文, 等. 基于绝句生成的构造式信息隐藏算法[J]. *计算机学报*, 2021, 44(4): 773-785.)

[44] 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*, 2019, 14(5): 1280-1295.

[45] 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*, 2021, 16: 880-895.

[46] Zhang H, You W K, Zhao X F. A Survey of Video Steganalysis[J]. *Journal of Cyber Security*, 2018, 3(6): 13-27.
(张弘, 尤玮珂, 赵险峰. 视频隐写分析技术研究综述[J]. *信息安全学报*, 2018, 3(6): 13-27.)

[47] Zhang F. “*martian*” *analysis*[D]. Changchun: Northeast Normal University, 2010.
(张帆. “火星文”探析[D]. 长春: 东北师范大学, 2010.)

[48] Sun X M, Chen H W, Yang L H, et al. Mathematical Representation of a Chinese Character and Its Applications[J]. *International Journal of Pattern Recognition and Artificial Intelligence*, 2002, 16(6): 735-747.

[49] Rizzo S G, Bertini F, Montesi D. Content-Preserving Text Watermarking through Unicode Homoglyph Substitution[C]. *The 20th International Database Engineering & Applications Symposium*, 2016: 97-104.

[50] Wu G H, Gong L C, Yuan L F, et al. Review of Information Hiding on Chinese Text[J]. *Journal on Communications*, 2019, 40(9): 145-156.
(吴国华, 龚礼春, 袁理锋, 等. 中文文本信息隐藏研究进展[J]. *通信学报*, 2019, 40(9): 145-156.)

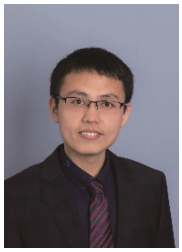
[51] Xiang L Y, Yu J M, Yang C F, et al. A Word-Embedding-Based Steganalysis Method for Linguistic Steganography via Synonym Substitution[J]. *IEEE Access*, 6: 64131-64141.

[52] Fu M, Dai Z X, Hu W T. A Syntax Checking Algorithm for Information-Hiding[J]. *Science Technology and Engineering*, 2015, 15(21): 142-145.
(付敏, 戴祖旭, 胡文涛. 一种文本信息隐藏中的语法检测算法[J]. *科学技术与工程*, 2015, 15(21): 142-145.)

[53] Jiang B, Ping X J, Zhang T. Pattern analysis applied on steganalysis for binary text images[C]. *2008 International Symposium on Electronic Commerce and Security*, 2008: 351-354.

朱嘉豪 于常州大学信息管理与信息系统专业获得学士学位。于南京航空航天大学网络空间安全专业获得硕士学位。研究领域为信息隐藏、深度学习等。研究兴趣包括：对抗样本。Email: jiahaozhu@nuaa.edu.cn

刘哲 于卢森堡大学获得博士学位。南京航空航天大学教授，博士生导师，研究领域为密码工程、AI 安全、区块链等。Email: gecp@nuaa.edu.cn



张玉书 于重庆大学获得博士学位。南京航空航天大学教授，博士生导师，研究领域为多媒体安全、区块链等。Email: yushu@nuaa.edu.cn

张新鹏 于上海大学获得博士学位。复旦大学教授，博士生导师。研究领域为多媒体信息安全、AI 安全、图像处理等。Email: zhangxinpeng@fudan.edu.cn