https://crad.ict.ac.cn
62(3): 779−789, 2025

基于对联载体的生成式信息隐藏方法

李欣然[1] 李蓉受[2] 秦 川[2] 钱振兴[3] 张新鹏[3]
[1](上海理工大学管理学院 上海 200093)
[2](上海理工大学光电信息与计算机工程学院 上海 200093)
[3](复旦大学计算机科学技术学院 上海 200433)
(xinranLi@usst.edu.cn;)

Generative Information Hiding Method Based on Couplet Carrier

Li Xinran[1], Li Rongshou[2], Qin Chuan[2], Qian Zhenxing[3], and Zhang Xinpeng[3]
[1] (*Business School*, *University of Shanghai for Science and Technology*, *Shanghai* 200093)
[2] (*School of Optical-Electrical and Computer Engineering*, *University of Shanghai for Science and Technology*, *Shanghai* 200093)
[3] (*School of Computer Science*, *Fudan University*, *Shanghai* 200433)

Abstract In order to improve the hiding capacity of information hiding algorithms while ensuring the quality of generated text, we propose a generative information hiding method based on couplet carrier. Firstly, we pre-train the couplet text data and build a couplet generation model based on a multi-flow pre-training and fine-tuning framework; secondly, we use the subject words as the input to generate the first line of a couplet, and the model can generate the first line of couplets on the same subject words; then we use the first line of a couplet as the input to generate the second line of a couplet. The method mitigates the semantic ambiguity in the current couplet generation model by utilizing the span-by-span learning approach, the padding generation mechanism and the noise perception mechanism to ensure that the generated couplets correspond to each other in terms of their metrical patterns. The secret information can be effectively hidden by different choices of subject words, candidate the first line of couplets and candidate words for generating the second line of a couplet. The experimental results show that the method can obtain high hiding capacity, and the average hiding capacity of 7-word couplets can reach 10.24B, and the generated couplets satisfy the strict form and content requirements of couplets, such as equal number of words, comparable lexicality, proportional structure and harmonious ping-ze. The overall performance of the proposed method is better than the current mainstream generative text information hiding schemes.

Key words text generation; generative information hiding; couplet carrier; text quality; hiding capacity

摘 要 为了提高信息隐藏算法的隐藏容量,同时保持生成文本的质量,提出一种基于对联载体的生成式信息隐藏方法.首先通过对对联文本数据进行预训练,在多流预训练和微调框架的基础上建立对联生成模型;其次将主题词作为输入生成上联,且模型可以在同一主题词上生成多个上联;然后将上联作为输入生成下联.通过对主题词、候选上联以及生成下联的候选词的不同选择,可有效地隐藏秘密信息.实验结果表明,该方法可获得较高的隐藏容量,7字对联的平均隐藏容量可达10.24 B,生成的对联结果满足对联严格的形式和内容要求,即字数相等、语义相对和语调相协,所提方法整体性能优于当前主流的生成式文本信息隐藏方案.

收稿日期: 2023-08-11; **修回日期**: 2024-05-23
基金项目: 国家自然科学基金项目(62172280); 国家自然科学基金重点项目(U20B2051); 上海市自然科学基金项目(21ZR1444600)
This work was supported by the National Natural Science Foundation of China (62172280), the Key Program of the National Natural Science Foundation of China (U20B2051), and the Natural Science Foundation of Shanghai (21ZR1444600).
通信作者: 秦川(qin@usst.edu.cn)

关键词 文本生成; 生成式信息隐藏; 对联载体; 文本质量; 隐藏容量

中图法分类号 TP391

DOI: 10.7544/issn1000-1239.202330663 **CSTR:** 32373.14.issn1000-1239.202330663

在当今互联网信息时代, 人们越来越关注通信传输中的信息安全. 相比传统的加密技术, 信息隐藏是一种安全通信技术[1-3], 将秘密信息除收发双方外以他人不易察觉的形式隐藏在数据载体中, 并广泛应用于军事、法律、知识产权和信息保护等领域[4-5]. 文本是互联网数据的核心形式之一, 具有格式多样、数据量小、传输效率高等优点, 该技术分为基于文本格式和基于文本内容.

1)基于文本格式. 修改字符间的距离、行距、字体大小、RGB 值和字体风格等进行信息隐藏. 文献 [6−7] 控制 HTML 文本的字体颜色和空格部位隐藏秘密信息; 文献 [8] 利用 ASCII 码中的不可见字符的独特性实现隐藏算法设计. 这些方法通过微调文本的格式特征来实现信息隐藏, 通常很难被人类视觉系统察觉[9]. 然而, 随着文本处理技术和方法手段的发展, 这些方法的安全性逐渐降低, 容易被检测出异常.

2)基于文本内容. 通过词汇、句法或语义进行操作来隐藏信息, 可分为修改式算法和生成式算法. 修改式算法充分利用了自然语言中的语义和语法冗余[10-11], 通常很难从阅读中察觉异常, 具有良好的隐蔽性. 然而, 这些算法比较复杂, 容易受到隐写分析攻击[12]. 生成式算法无需修改文本内容或格式, 以秘密信息驱动生成文本进行信息隐藏. 下面将介绍一些于文本生成的代表性算法.

文献 [13] 利用宋词词典与模板隐藏秘密信息, 却忽略了词汇之间的语义联系, 可读性较差, 同时隐藏容量也较低. 文献 [14] 在此基础上提升了隐藏容量. Luo 等人[15] 利用马尔可夫链生成了文本质量较高的宋词, 但算法隐藏容量较低. 文献 [16] 提出了一种基于混合加密模型, 增强了安全性, 但生成的宋词质量较低且隐藏容量有限. 邹孜逸等人[17] 利用宋词韵律的“平仄”奇偶性隐藏秘密信息.

Luo 等人[18] 利用编码-解码(encoder - decoder)模型生成绝句, 使用候选词隐藏信息并实现较高的隐藏容量. 然而, 该方法在将候选词组合成句子时存在较大的信息损失. 文献 [19] 通过增加主题信息来生成绝句, 从而提高了生成绝句的质量.

鉴于上述问题, 本文提出了一种基于对联载体的生成式信息隐藏算法. 首先输入主题词生成上联, 然后以上联为输入生成下联. 模型采用 span-by-span 学习方式进行预训练, “多片段-多粒度”的目标使得模型能够自然地表达主题词信息, 生成满足对联对仗特性的下联. 同时, 利用主题词、生成的上联和下联中的不同选择信息来设计信息隐藏方法, 实现了较高的隐藏容量. 除对联之外, 本文提出的生成式信息隐藏方法还可以扩展应用于其他文本, 如元曲、宋词、歌词等(需要使用相应文本的数据集进行训练).

本文贡献主要在于: 1)利用对联生成的特性, 设计了包括主题词和上联候选信息在内的上联部分, 并设计了逐字生成的下联候选信息, 有效增加了隐藏容量; 2)通过模型的微调设计, 采用 span-by-span 学习方式提升了生成对联的语义完整性, 使生成的上下联满足对仗特性.

所提基于对联生成的生成式信息隐藏方法相比于载体选择式方法更具有安全性. 与选择式方法不同, 生成式方法没有唯一词库, 字与字之间的语义匹配由模型计算, 这种特殊的生成方式使得攻击者难以分析和攻击生成式方法. 此外, 生成式方法更注重文本的语义联系, 因此能够生成更高质量的文本.

1 方法整体框架

基于对联载体的生成式信息隐藏方法框架如图 1 所示, 该框架由模型训练、信息隐藏和信息提取 3 个部分组成.

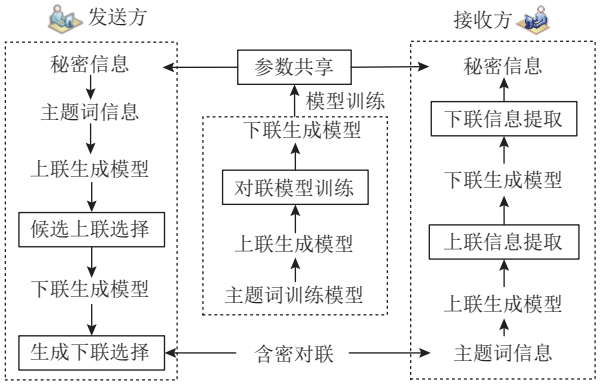


Fig. 1 Overall frame diagram of our method

图 1 本文方法整体框架图

1)模型训练. 该部分包括上联生成模型和下联生成模型的训练过程. 首先将对联数据集进行预处理, 训练完成后进行参数的微调操作, 确保本文模型生成的

对联符合秘密信息的隐藏和提取. 同时, 发送和接收双方需共享模型, 以确保双方的模型生成效果一致.

2)信息隐藏. 该部分是将秘密信息隐藏在对联中的过程. 首先以秘密信息选定相应主题词, 输入上联生成模型, 生成候选上联, 随后再依据秘密信息在候选上联中选择上联; 接下来, 将该上联作为下联生成模型的输入, 生成下联候选词, 通过设定候选词而逐一选择最终生成的下联, 上联与下联构成含密对联.

3)信息提取. 该部分是隐藏信息的逆过程, 用于从接收到的对联中提取隐藏的秘密信息. 根据接收到的对联提取主题词, 将主题词输入上联生成模型, 生成候选上联. 接下来, 根据接收到的对联对照确定上联隐藏的秘密信息. 将上联输入下联生成模型, 生成候选词, 根据含密对联进行对照选择, 逐一提取下联隐藏的秘密信息.

2 对联生成模型

本文的对联生成模型主要基于 ERNIE-GEN 模型, 因此本节重点介绍该模型.

2.1 ERNIE-GEN 模型基本结构

ERNIE-GEN 模型采用多流注意力(multi-flow attention)结构的 Seq2Seq 训练框架[20], 如图 2 所示. 其中 S 和 T 分别代表源序列和目标序列, Y 为生成序列. 损失系数 λ 通过逐字符和逐片段生成以及引入噪声的目标序列 T' 计算, 成功避免了重复率问题.

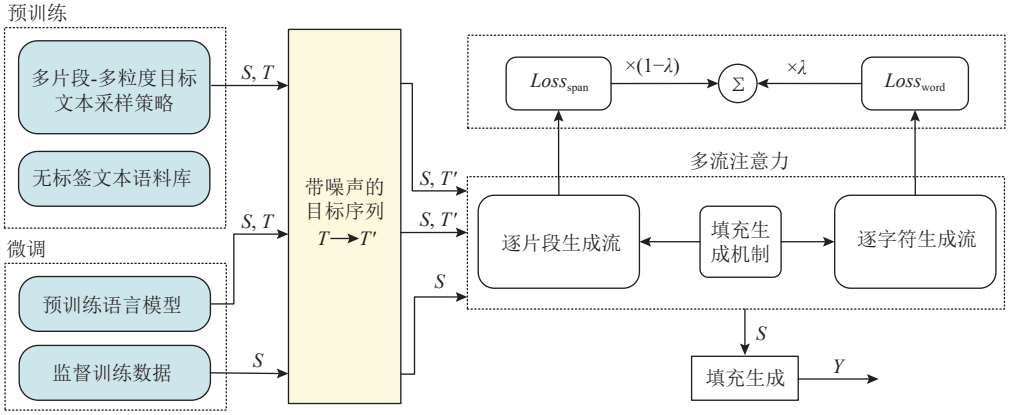


Fig. 2 Frame diagram of ERNIE-GEN model

图 2 ERNIE-GEN 模型框架图

在 Seq2Seq 框架中, 进行训练和预测解码生成字符 y_{i+1} 的公式分别为

$$y_{i+1} \leftarrow Multihead(Q=t_i, KV=[S, t_{\leqslant i}]), \quad (1)$$

$$y_{i+1} \leftarrow Multihead\left(Q=y_i, KV=\left[S, y_{\leqslant i}\right]\right), \quad (2)$$

其中 t_i 表示模型学习训练过程中第 i 个字符使用的真实值, y_i 表示模型解码过程中第 i 个字符的预测值, S 代表编码端的输入表达. 在这些公式中, Q 是包含上文环境的信息表示, 它对预测 y_{i+1} 产生直接影响. 训练和解码过程中, 式(1)(2)揭示了 Q 的形成差异, 这也是曝光偏差的主要原因.

ERNIE-GEN 模型采用了填充生成(infilling generation)机制来解决曝光偏差的问题. 在训练和解码过程中, 对于解码器输入句子序列的每个字符, 通过填充特殊符号 [A] 来接收 Q, 统一方式汇聚上下文信息, 可消除训练和解码时存在的差异, 并更有效地利用上下文信息进行预测. 图 3 展示了这种方法的优点, 它提高了模型的泛化能力和性能, 并增强了文本生成任务的表现力.

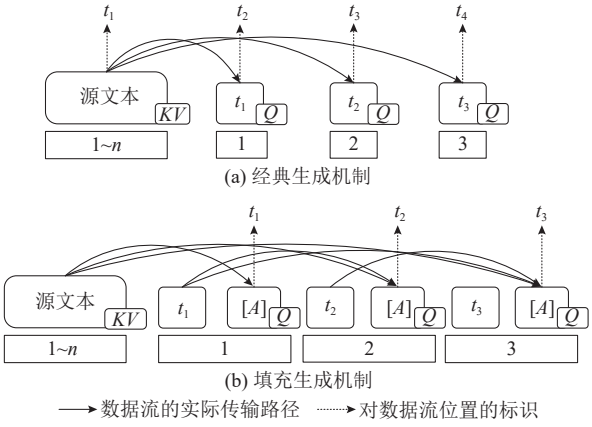


Fig. 3 Infilling generation mechanism

图 3 填充生成机制

2.2 ERNIE-GEN 模型的逐段学习方式

传统自回归 Seq2Seq 训练将模型逐字符(word-by-word)输入并生成句子序列, 即只利用单侧信息进

行学习. 然而, 这种方式容易导致模型过度依赖和局部过拟合, 因为模型的注意力主要集中在当前位置, 对上一个字符的信息高度依赖. 与人类写作不同, 人们通常会先构思整体框架, 同时构思好实体、短语或句子, 而非逐字思考.

ERNIE-GEN 提出的逐片段(span-by-span)学习[21]方式能够解决上述问题. span-by-span 学习方法通常用于解决类似的序列生成任务, 如自然语言生成、机器翻译等. 与逐片段学习的任务相比, 我们的方法专注于对联生成, 这是一项具有独特挑战性的任务, 需要更多的语言创意和修辞技巧. 这种学习方式更符合自然语言的表达方式, 可以提高生成文本的连贯性和逻辑性, 并降低过拟合的风险. 图 4 展示了逐片段学习过程.

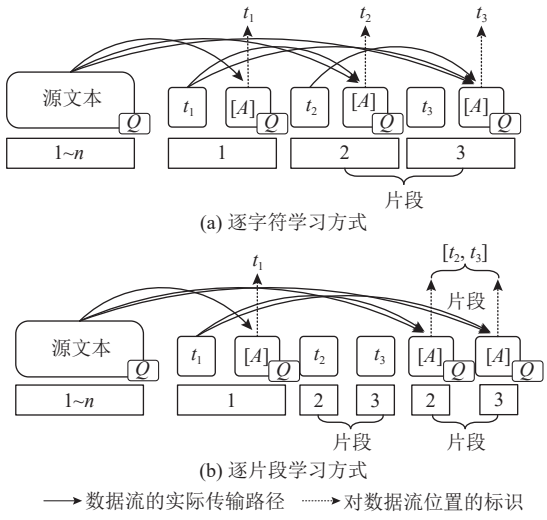


Fig. 4 Illustration of span-by-span learning

图 4 逐片段学习示意图

使用逐字符学习方式, 训练所得的目标序列的计算概率为

$$P(t_1,t_2,t_3 \mid S,\theta) = P(t_1 \mid S,\theta)\times P(t_2 \mid [S,t_1],\theta)\times P(t_3 \mid [S,t_1,t_2],\theta). \quad (3)$$

然而, 如果使用逐片段学习方式, 训练所得的目标序列的计算概率为

$$P(t_1,t_2,t_3 \mid S,\theta) = P(t_1 \mid S,\theta)\times P(t_2,t_3 \mid [S,t_1],\theta), \quad (4)$$

$$P(t_2,t_3 \mid [S,t_1],\theta) = P(t_2 \mid [S,t_1],\theta)\times P(t_3 \mid [S,t_1],\theta). \quad (5)$$

在处理 $[t_2, t_3]$ 过程中, 逐字符学习方式需要依赖上文语言环境信息 $[S, t_1, t_2]$ 来训练预测 t_3, 而逐片段学习方式在训练预测 t_3 的同时进行预测工作, 并且使用相同的上文语言环境信息来预测 t_3. 这个策略可以让模型在 t_1 时就开始准备生成下一个语义片段, 类似于人类思维机制中的“远见”能力. 因此, 这有助于提高模型生成的文本质量.

3 基于对联生成的信息隐藏和提取

3.1 信息隐藏

本文所提出的方法将二进制比特流作为 2 段秘密信息, 上联隐藏信息设为第 1 段, 下联隐藏信息设为第 2 段. 上联隐藏主题词和候选上联的选择信息, 下联隐藏逐字选择信息. 图 5 是以主题词“花”为示例的信息隐藏算法具体流程.

第 1 步. 将秘密信息进行分段, 依据分段信息对应查找主题词(subject words), 参数输入上联生成模型中进行候选上联的生成, 随后依据秘密信息选择上联. 在模型生成上联的过程中, 本文采用了集束搜索(beam-search)算法, 该算法通过设定集束大小和损失排序计算, 在候选序列筛选出近似最优解(而非最优解).beam-search 算法在选择候选解时, 类似于动态规划, 以较少的代价完成候选对联的选择. 图 5 中的候选上联个数 B_{num}=4, 即依据秘密信息可从 4 个候选上联中选择最终输出上联. 通过阅读可以看出, 4 个候选上联的语义都较为通顺且归属于同一个主题词: “花”. 假设主题词列表大小为 S_{num}, 则第 1 段秘密信息的长度 L_{bs} 的计算公式为

$$L_{\text{bs}} = \text{lb}\, B_{\text{num}} + \text{lb}\, S_{\text{num}}. \quad (6)$$

第 2 步. 需要依据秘密信息选择生成下联. 下联的生成过程中, 严格符合对联的对仗特性, 根据输入的上联长度(图 5 的上联长度为 7)可知下联字选择的次数. 在模型生成下联的选择过程中, 利用 Top-k Sampling[22] 技术筛选候选字. 在生成每个字时, 从预测出的可能性分布中选择可能性最大的前 k 个词, 然后按照它们的概率来进行排序选择.

将秘密信息划分为 D_{Length} 段, 依据秘密信息选择每个候选词包内的对应排序的字生成下联即可. 假设候选词包个数为 H_{dict}(图 5 中所设置的 H_{dict}=8), 则第 2 段秘密信息的长度 L_{dh} 为

$$L_{\text{dh}} = D_{\text{Length}}\text{lb}\, H_{\text{dict}}. \quad (7)$$

随着 B_{num} 和 H_{dict} 的增加, 对联的秘密信息隐藏容量能得到一定的提升. 由式(6)(7)可知, 本文算法中对联所能隐藏的秘密信息长度 L_{couplet} 为

$$L_{\text{couplet}} = L_{\text{bs}} + L_{\text{dh}} = \text{lb}\, B_{\text{num}} + \text{lb}\, S_{\text{num}} + D_{\text{Length}}\text{lb}\, H_{\text{dict}}. \quad (8)$$

结合式(8)可以看出, 本文的隐藏容量与 4 个参数因子相关, S_{num} 为主题词个数, B_{num} 为候选上联的

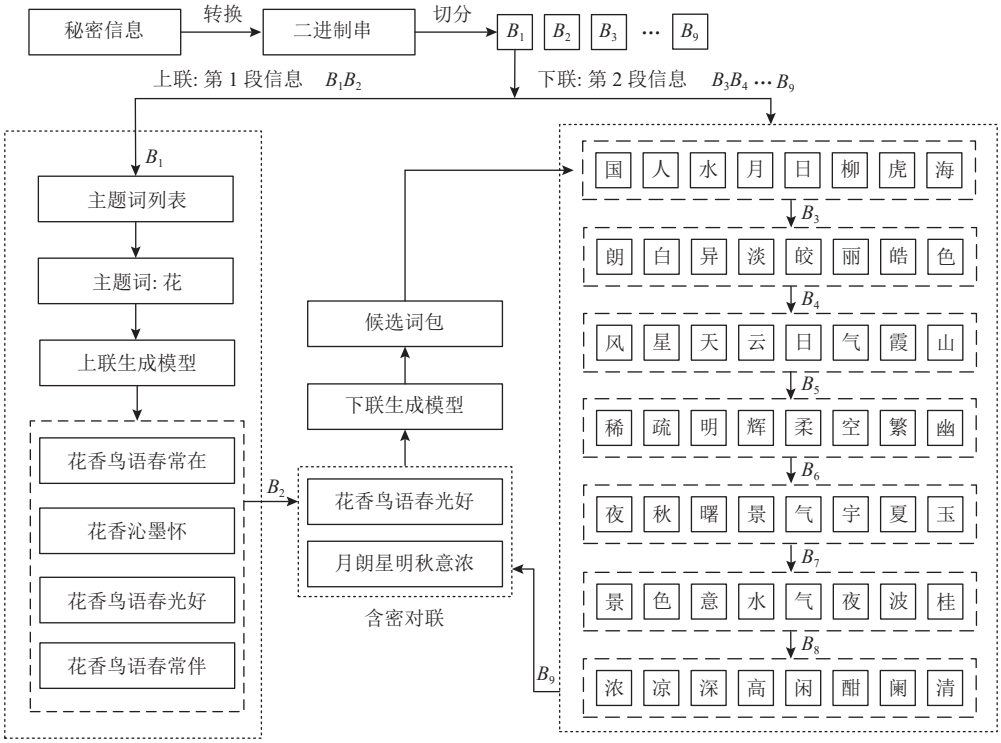


Fig. 5 Flow chart of information hiding

图 5 信息隐藏流程图

个数，H_{dict} 为生成下联时逐字选择的数量，在一定范围内可以自行设置；D_{Length} 是依据生成上联而定的字数长度.

3.2 信息提取

信息提取可视为隐藏算法的逆过程. 第 1 步. 从接收到的上联中提取主题词并查找对应的秘密信息，将主题词输入上联生成模型生成候选上联，依据接收到的含密上联对照提取选择信息，2 个部分的信息级联得到第 1 段秘密信息. 第 2 步. 将上联输入下联生成模型，依据接收到的含密下联，对照候选词包，提取隐藏在下联中的第 2 段秘密信息. 在我们提出的基于对联载体的生成式信息隐藏算法中，发送方和接收方共享相同的模型，从而确保信息生成和提取的一致性. 这种设计允许接收方通过与发送方相同的输入（如主题词）来验证生成的上联和下联是否一致. 接收方可以根据发送方的上联来生成候选下联，并通过比较这些候选下联与实际接收到的下联来验证信息的正确性. 若下联在语义、节奏或对仗方面与上联不符，这可能指示信息提取过程中的错误. 接收方还可通过对比候选上联和实际上联，以及检查下联中每个字的语境适宜性，来进一步确认信息的准确性.

基于对联生成的信息提取方法流程如图 6 所示. 依据信息隐藏方法查找主题词位置以方便提取，如图 6 中对联的主题词为“花”，随即查找“花”所隐藏的秘密信息；随后将主题词输入上联生成模型，依据共享参数生成候选上联，根据含密对联中的上联提取其对应的选择信息. 将上联中 2 部分信息拼接为提取的第 1 段秘密信息. 之后将上联输入到下联生成模型，生成候选词包，依据输入上联读取长度 D_{Length}. 对照含密对联，查找候选字的排序信息，所查找“月朗星明秋意浓”的排序信息即为映射的秘密信息，逐一查找直到生成结束. 将 D_{Length} 段排序信息级联得到第 2 段秘密信息. 随后 2 段秘密信息级联，可获取含密对联的全部秘密信息.

4 实验结果及比较

4.1 实验设置

本文所提出的模型采用 PaddlePaddle 框架. 在实验中，对联模型的训练数据集包含 770 494 副对联，预训练和微调阶段的参数设置都是确定的. 模型相关参数设置如下：每次训练文本数量为 10，最大训练步数为 400，学习率为 5E−5，最长编码长度为 250，最长解码长度为 250，噪声概率为 0.2，标签平滑权重为 0.1 以及模型调优 Dropout 参数设为 0.2.

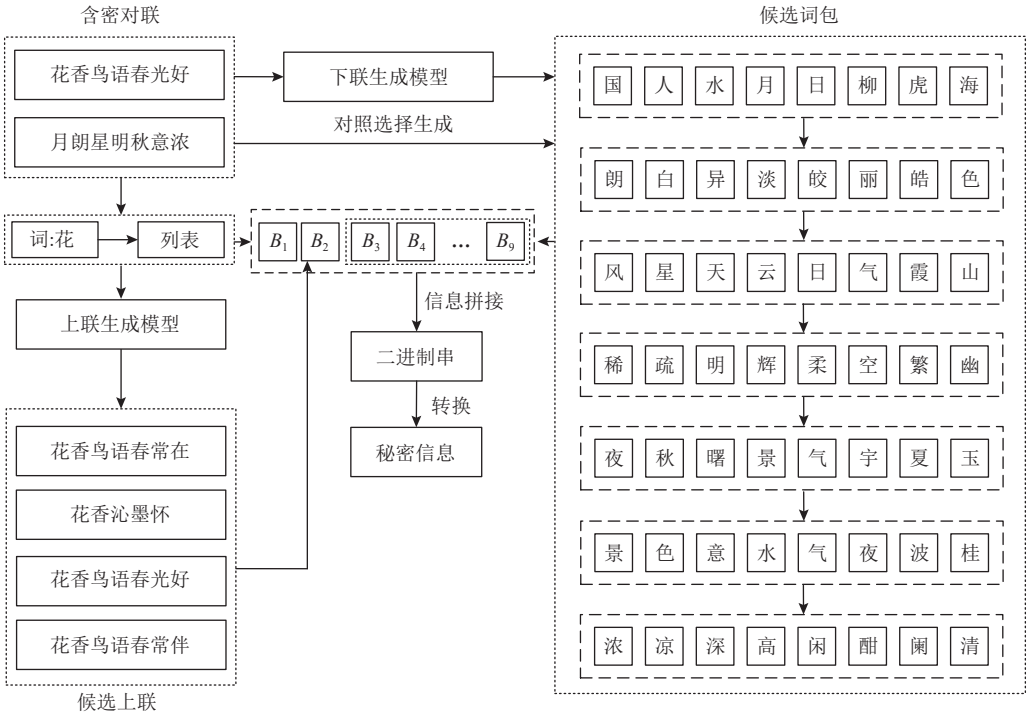


Fig. 6 Flow chart of information extraction

图 6 信息提取流程图

4.2 生成对联质量比较

中国传统对联是一个挑战性的文本生成任务, 文献 [23−26] 采用了不同的深度学习模型尝试解决对联生成. 利用 encoder-decoder 框架、Attention 机制、Transformer 模型以及双层注意力机制等技术优化对联的生成质量. 然而, 在现有的对联生成方法中仍然存在一些问题, 如仅使用字嵌入等方法导致词语语义的丧失和对联内部对应程度不佳等问题. ERNIE-GEN 模型能够生成语义完成、目标文本和原文本之间有较强关联性的文本, 使基于 ERNIE-GEN 模型生成的对联质量在当前的对联生成任务中处于较好水平. 因此, 本文所提方法主要与 ERNIE-GEN 模型进行生成对联的质量对比.

在自然语言处理领域中, 困惑度 (perplexity, *PPL*) 是衡量语言模型优劣的重要指标.*PPL* 的计算基于这样一个思想, 即当一个语言模型生成文本时, 它需要预测下一个字的生成可能性, 而当可选字的数量越少时, 便认为该模型越准确. 因此, *PPL* 计算为

$$PPL(S)=p(w_1w_2\cdots w_M)^{-\frac{1}{M}}, \tag{9}$$

其中 S 代表句子, M 为句子长度, $p(w_i)$ 表示第 i 个词的概率. 因此, 一般 *PPL* 的值越小, 则模型生成文本的质量越好. 分别对基于 ERNIE-GEN 的对联生成模型 (ernie_gen_couplet) 和本文模型进行训练. 在本次实验中使用 *PPL* 来衡量语言模型的性能, ernie_gen_couplet 模型的 *PPL*=33.71、损失值为 3.51, 本文模型的 *PPL*=33.16、损失值为 3.50. 从性能指标层面, 本文模型与 ernie_gen_couplet 生成的对联质量基本处于同一水平. 表 1 中展示了 2 种模型生成的对联.

本文方法所使用的对联训练样本是严格按照平仄对仗的标准对联进行选择的. 由于训练数据的高质量, 生成的对联符合对联的艺术形式要求. 表 1 中的 10 幅对联在对仗准确性、平仄协调性, 以及诗意和意境的深度等方面均符合生成对联的标准. 如上联“一支妙笔, 写尽人间幽默事”与下联“半壶浊醪, 消除世上寂寥心”满足字数相等, 上下联字数皆为 11 个字; 词性相当, 如数词“一支”对“半壶”, 形容词“妙”对“浊”, “幽默”对“寂寥”, 名词“笔”对“醪”, “事”对“心”等; 从阅读可以感知对联结构相称、节奏相应、平仄相协, 即“一支/妙笔, 写尽/人间/幽默事”对“半壶/浊醪, 消除/世上/寂寥心”. 从生成下联对比分析, 生成效果基本相同, 内容上或有不同, 但对联要求的对仗格律是基本一致的. 如“华章美景颂祥时”与“小康大道展宏图”, 对应同一个输入上联“盛世丰年交好运”, 在词性、结构、节奏和平仄等方面都是基本相当的.

4.3 隐藏容量比较

隐藏容量是隐藏方法的重要性能指标之一, 图 7 描述了本文方法与其他典型的基于古诗词生成的隐

Table 1 Display of Generated Couplets

表 1 生成对联展示

上联	本文方法生成的下联	ERNIE-GEN 模型生成的下联
一院梅花堪养眼	半帘明影最撩心	半窗竹影可清心
春度青山燕预报	花香碧海风传媒	花开碧野蝶先知
人生岂是方程式	岁月无非应用题	岁月如为应考关
花好终须明月伴	情真自是美酒喝	果真后面暗分居
盛世丰年交好运	华章美景颂祥时	小康大道展宏图
三杯浊酒，酒香留醉	一盏清茶，茶韵留芳	一曲高山，山色添愁
花影入帘，松翠在壁	云根作柱，风骨为梁	云气满室，清光映阶
春霖降大地，有情迎盛世	善政暖黎民，着意送和谐	红日升神州，无处不东风
风气清和，岁月温馨堪寄意	山河秀美，乾坤宁定好吟春	人文厚重，山河壮丽好抒怀
一支妙笔，写尽人间幽默事	半壶浊醪，消除世上寂寥心	几首歪诗，吟出世上乐忧情

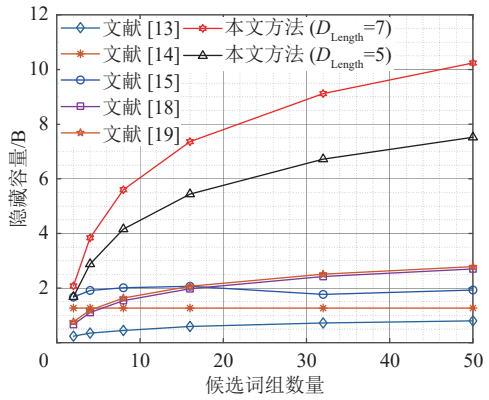


Fig. 7 Comparison results of hidden capacity of different methods

图 7 不同方法隐藏容量比较结果

藏方法的对比结果. 文献 [13−15] 基于宋词模板的词语分组, 当文本数据库规模相对固定时, 通过增加候选词组的数量可以提高方法的隐藏容量. 但需要注意的是, 词组的选择必须考虑文本的题材、风格等相关因素, 否则可能会影响信息的隐藏效果. 文献 [18−19] 基于绝句生成, 其本质上都属于基于单个候选字的方法来隐藏秘密信息的方法. 该方法的优势在于, 它拥有一个较大的候选字库, 并且能够在单个字的级别上实现信息隐藏, 提高了隐藏容量和隐蔽效果, 但需要控制候选字的数量, 以兼顾隐藏容量和生成诗词的质量. 本文的对联生成式信息隐藏方法之所以具有高隐藏容量, 归功于其精心的网络设计. 关键因素包括主题词的多样性和候选上联数量的增加, 以及下联生成中 Top-k Sampling 技术的应用. 逐片段学习方式和集束搜索方法的使用, 不仅提高了对联的生成质量, 还扩大了隐藏信息的复杂度和范围.

在本文实验中, 根据数据集训练生成的字表数量为 17 963, 经过低频词处理后从中选择 8 500 个作为候选字表. 由 4.1 节可知, 生成对联的长度为 D_{Length}, 那么对联的候选生成理论上共有 $8\,500^{D_{\text{Length}}}$ 种选择, 可选择的范围极大, 也就意味着本文算法可以实现高容量的信息隐藏. 本文实验通过 beam-search 算法、Top-k 算法对候选字进行筛选, 确保了从 $8\,500^{D_{\text{Length}}}$ 种选择中生成对联的质量.

从式(8)看出, 本文方法的信息隐藏容量主要与参数 S_{num}, B_{num}, H_{dict}, D_{Length} 相关, 即本文方法隐藏容量的计算本质上是以单个字为基本隐藏单位. 因此, 为能更好地与其他文献进行比较分析, 将 2 个候选参数进行统一并转换成以单字为基本隐藏单元的嵌入表达式, 即 $B_{\text{num}}=H_{\text{dict}}$, 候选词组 N 的计算方式满足

$$N^{D_{\text{Length}}}=B_{\text{num}}=H_{\text{dict}}. \tag{10}$$

在计算机中存储 1 个文字需 2 B, 因此本文方法的隐藏容量 C_{rate} 的计算公式为

$$C_{\text{rate}}=\frac{L_{\text{couplet}}}{2(D_{\text{Length}}+1)\times 2}=\frac{D_{\text{Length}}(D_{\text{Length}}+1)\text{lb}N+\text{lb}S_{\text{num}}}{2(D_{\text{Length}}+1)\times 2}, \tag{11}$$

其中 D_{Length} 的取值在理论上是任意的, 本文实验中选取 $D_{\text{Length}}\geqslant 5$. 本次实验中, 训练集的主题词通过方法进行筛选, 选取对联数据集中出现频率不少于 100 的主题信息用于生成上联且隐藏信息的主题词的 $S_{\text{num}}=1\,748$; 进行对比实验前, 为确保实验的公平性, 对方法的相关参数进行统一. 此时, 选取 D_{Length} 为 5 和 7 以便与其他文献的信息隐藏容量作对比. 由图 7 可知, 本文方法在隐藏容量方面表现极好, 远高于其他文献方法. 需要注意的是, 由式(11)可知本文方法的隐藏容量随着 D_{Length} 的增加而增加, 因此本文方法在隐藏容量方面拥有极高的隐藏能力.

4.4 时间复杂度分析

信息隐藏和提取过程中使用的字典码原本是预先生成的,引入的计算操作即为时间复杂度.针对选取主题词,该参量是依据实际情况而设定的固定值,即时间复杂度为 $O(1)$;随后基于选取的主题词生成上联,计算时间主要花费在选字生成过程,因此对应的时间复杂度为 $O(D_{\text{Length}}\times size)$, D_{Length} 为生成上联长度, $size$ 为候选字的数量,此时时间复杂度为 $O(n^2)$;接下来是输入上联生成下联的过程,计算也是集中在选字生成,故同理可求对应时间复杂度为 $O(n^2)$.因此,本文信息隐藏方法时间复杂度为 $O(n^2)$.在提取过程中,只需要依据主题词字典和候选上联进行索引,随后根据下联候选字进行索引的时间为 $O(n^2)$.本文方法的时间复杂度与文献 [13−15, 18−19] 比较结果列于表 2.

Table 2 Comparative Results of Time Complexity

表 2 时间复杂度的比较结果

方法	信息隐藏	信息提取
文献 [13]	$O(1)$	$O(n)$
文献 [14]	$O(1)$	$O(n)$
文献 [15]	$O(1)$	$O(n)$
文献 [18]	$O(n^2)$	$O(n^4)$
文献 [19]	$O(n^2)$	$O(n^2)$
本文方法	$O(n^2)$	$O(n^2)$

本文所提方法的时间复杂度与文献 [19] 相同,但隐藏容量远高于文献 [19].由于本文方法增加了模型的预训练以及微调能够生成更加高质量的文本,因此在时间复杂度上并没有达到最优.但从隐藏容量和时间复杂度 2 方面综合分析,所提方法综合表现最佳,方法更具灵活性.

4.5 安全性分析

秘密信息的隐藏主要基于主题词、候选上联、生成下联的候选词的选择信息进行编码.提取秘密信息时应具备:

1)生成上联的主题词、主题词的隐藏位置以及主题词映射的二进制秘密信息.主题词的候选数量依据数据集中的 770 494 副对联,通过统计全部主题词共有 4 791 个候选数量,本文方法从中选取 1 749 个主题词生成上联,即共计选择的数量为 $C_{4\,791}^{1\,749}$,选取后的主题词经过排列组合有 1 749! 种可能.

2)分析提取生成上联的候选信息,即当前上联在候选上联中的位置信息及其对应的映射关系.候选上联的输出依赖于候选参数 B_{num},当 B_{num} 设置的值不同时,生成的候选上联也会产生变化,因此还需要分析候选参数 B_{num} 的取值.

3)分析提取下联生成的候选信息,包含候选词的隐藏位置和候选词对应的映射关系.候选词的词库数量为 17 963,本文方法从中选取 8 500 个候选字以生成下联,需要分析出该候选词表.同时,候选词的输出依赖于候选参数 H_{dict},当参数 H_{dict} 取值不同时,候选词的可选择范围及其映射关系都会发生变化,因此也需要分析候选参数 H_{dict}.当前候选词库的生成依赖于上一个候选词的选择信息,当上一个候选词的选择信息发生变化时,当前候选词库也会产生变化.因此下联生成的候选信息分析需要做到每一步都能正确分析提取,否则秘密信息破解将会失败.

4)分析获取模型相关的训练数据,包含上联生成模型和下联生成模型的训练数据.所提信息隐藏算法的设计依赖于对联的生成模型,只有分析出相关的训练数据才能构建相应的生成模型,才能分析相应的映射关系.

5)即使攻击者拥有训练数据,在双方的模型拟合程度不同的情况下,即便使用相同的参数,生成的对联也可能不完全相同.

基于以上分析,每一步的分析破解代价相当高,攻击者很难获取秘密信息.

在该实验中,通过方法[27]对含密文本和非含密文本进行语义空间映射,并使用 t-SNE 方法进行可视化,观察它们在统计上的区分难易程度.实验结果显示在图 8 中.在不同的信息隐藏容量下,即 D_{Length}=5, 7, 11, 13 时,2 类文本在语义空间上基本一致.随着对联长度的增加,即信息隐藏容量增加,含密文本与不含密文本的分布仍然保持基本一致.因此,根据这些分析,本文方法生成的隐藏秘密信息的对联文本与未隐藏信息的对联文本在语义空间上几乎无法区分.

5 总 结

本文提出一种基于对联载体的生成式信息隐藏方法,该方法能够生成高质量对联,且保证高容量的秘密信息隐藏,提供相对较高的安全性.通过在预训练阶段加入逐片段学习方式来确保模型生成对联的语义相对性.通过主题词、候选上联和生成下联的候选词的不同选择,实现生成对联的同时隐藏秘密信息.主题词编码与生成模型有效保证信息提取的正确性.本文提出的基于对联的生成式信息隐藏方法在自然语言处理和信息安全领域展现了创新性,因

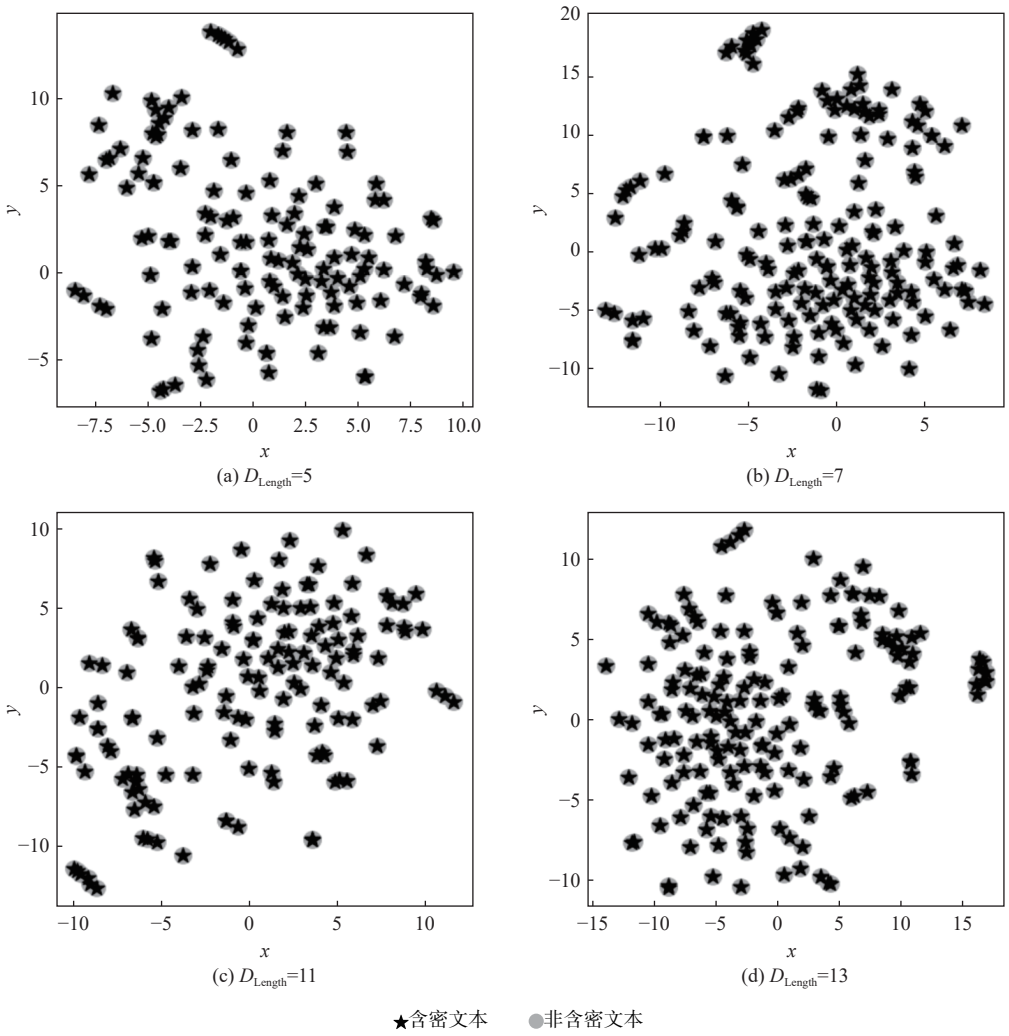


Fig. 8 Distribution differences of different D_{Length} values in semantic space

图 8 在语义空间中不同 D_{Length} 取值下的分布差异

其深植于中国文化，面临着在国际或非中文语境中推广的挑战. 未来的研究也将探索更多样化的文本载体，如元曲、歌词或小说，以适应不同文化背景. 此外，提高模型的语言处理能力和安全性，以及在增加隐藏容量的同时保持对联的自然和艺术性.

作者贡献声明：李欣然提出了算法思路和实验方案；李蓉受完成实验及数据整理；秦川撰写论文；钱振兴修改论文；张新鹏提出指导意见.

参 考 文 献

[1] Chen Weibing, Yang Gaobo, Chen Richao, et al. Digital video passive forensics for its authenticity and source[J]. Journal on Communications, 2011, 32(6): 177−183 (in Chinese)
(陈威兵, 杨高波, 陈日超, 等. 数字视频真实性和来源的被动取证[J]. 通信学报, 2011, 32(6): 177−183)

[2] Yin Zhaoxia, Guo Hongnian, Du Yang, et al. Multi-domain reversible data hiding in JPEG images and payload distribution algorithm[J]. Journal of Computer Research and Development, 2022, 59(8): 1831−1840 (in Chinese)
(殷赵霞, 郭红念, 杜洋, 等. JPEG 图像多域可逆信息隐藏及载荷分配算法[J]. 计算机研究与发展, 2022, 59(8): 1831−1840)

[3] Huang Taoyu, Huang Fangjun. Adaptive JPEG reversible data hiding method based on pairwise coefficients[J]. Journal of Software, 2022, 33(2): 725−737 (in Chinese)
(吴桃宇, 黄方军. 基于系数配对的自适应 JPEG 可逆信息隐藏方法[J]. 软件学报, 2022, 33(2): 725−737)

[4] Yan Shu, Chen Fan, He Hongjie. Reversible data hiding in encrypted image based on neighborhood prediction using XOR-permutation encryption[J]. Journal of Computer Research and Development, 2018, 55(6): 1211−1221 (in Chinese)
(鄢舒, 陈帆, 和红杰. 异或-置乱框架下邻域预测加密域可逆信息隐藏[J]. 计算机研究与发展, 2018, 55(6): 1211−1221)

[5] Kumar R, Malik A, Singh S, et al. A space based reversible high capacity text steganography scheme using font type and style [C] // Proc of the Int Conf on Computing, Communication and Automation

(ICCCA). Piscataway, NJ: IEEE, 2016: 1090−1094

[6] Lee C, Chen H. Data concealment scheme for HTML documents based on color code [C] // Proc of the 5th Int Conf on Intelligent Information Hiding & Multimedia Signal Processing. Piscataway, NJ: IEEE, 2009: 632−635

[7] Cui Gangming, Hong Xing, Yuan Xiang, et al. Research on information hiding based on invisible character replacement[J]. Computer Applications and Software, 2016, 33(4): 283−286 (in Chinese)
(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): 283−286)

[8] Liu Feng, Luo Pengpeng, Ma Zhujuan, et al. Security secret information hiding based on hash function and invisible ASCII characters replacement [C] // Proc of the 15th IEEE Int Conf on Trust, Security and Privacy in Comp and Commun/the 10th IEEE Int Conf on Big Data Science and Engineering/the 14th IEEE Int Symp on Parallel and Distributed Proc with Applicat (IEEE Trustcom/BigDataSE/ISPA). Piscataway, NJ: IEEE, 2016: 1963−1969

[9] Chen Yini, Li Qianmu, Lü Chaoxian, et al. Research on text security hiding algorithm for invisible characters[J]. Cyberspace Security, 2019, 10(5): 88−96(in Chinese)
(陈旖旎, 李千目, 吕超贤, 等. 不可见字符的文本安全隐藏算法研究[J]. 网络空间安全, 2019, 10(5): 88−96)

[10] Su Bo, Ding Xiaoyun, Liu Gongshen, et al. An information hiding method for text by substituted conception [C] // Proc of the 4th Int Symp on Information Science and Engineering. Piscataway, NJ: IEEE, 2012: 131−135

[11] Xiang Lingyun, Li Yan, Hao Wei, et al. Reversible natural language watermarking using synonym substitution and arithmetic coding[J]. Computers Materials & Continua, 2018, 55(3): 541−559

[12] Xiang Lingyun, Yu Jingmin, Yang Chunfang, et al. A word-embedding-based steganalysis method for linguistic steganography via synonym substitution[J]. IEEE Access, 2018, 6: 64131−64141

[13] Yu Zhenshan, Huang Liusheng, Chen Zhili, 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 (in Chinese)
(于振山, 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)

[14] Liu Yanchen, Wang Jian, Wang Zhibin, et al. A technique of high embedding rate text steganography based on whole poetry of song dynasty [C] // Proc of the 2nd Int Conf Cloud Computing and Security. Berlin: Springer, 2016: 178−189

[15] Luo Yubo, Huang Yongfeng, Li Fufang, 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

[16] Liu Yanchen. Text information hiding technique by the carrier of Song poetry [D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2017 (in Chinese)
(刘彦辰. 宋词载体的文本信息隐藏技术[D]. 南京: 南京航空航天大学, 2017)

[17] Zou Ziyi, Wang Kaixi. Method of text steganography without carrier based on rhythm in Chinese Song poems[J]. Journal of Qingdao University: Science and Technology, 2021, 34(1): 7−12 (in Chinese)
(邹孜逸, 王开西. 基于宋词韵律的无载体文本隐写方法[J]. 青岛大学学报: 自然科学版, 2021, 34(1): 7−12)

[18] Luo Yubo, Huang Yongfeng. Text steganography with high embedding rate: Using recurrent neural networks to generate Chinese classic poetry [C] // Proc of the 7th ACM Workshop on Information Hiding & Multimedia Security. New York: ACM, 2017: 99−104

[19] Qin Chuan, Wang Meng, Si Guangwen, et al. Generative information hiding with Chinese quatrain generation[J]. Chinese Journal of Computers, 2021, 44(4): 773−785
(秦川, 王萌, 司广文, 等. 基于绝句生成的生成式信息隐藏算法[J]. 计算机学报, 2021, 44(4): 773−785)

[20] Xiao Dongling, Zhang Han, Li Yukun, et al. ERNIE-GEN: An enhanced multi-flow pre-training and fine-tuning framework for natural language generation [C] // Proc of the 29th Int Joint Conf on Artificial Intelligence. San Francisco, CA: Morgan Kaufmann, 2020: 3997−4003

[21] Xiao Dongling, Zhang Han, Li Yukun, et al. ERNIE-GEN: An enhanced multi-flow pre-training and fine-tuning framework for natural language generation [C] // Proc of the 29th Int Joint Conf on Artificial Intelligence. San Francisco, CA: Morgan Kaufmann, 2021: 3997−4003

[22] Angela F, Mike L, Yann D. Hierarchical neural story generation [C] // Proc of the 56th Annual Meeting of the Association for Computational Linguistics. Stroudsburg, PA: ACL, 2018: 889−898

[23] Zhu Yuanyuan. Research on generalized model of Chinese couplet based on recurrent neural networks [D]. Changchun: Jilin University, 2018 (in Chinese)
(朱媛媛. 基于循环神经网络的对联生成模型研究[D]. 长春: 吉林大学, 2018)

[24] Wei Wangcheng. Research on automatic generation of poetry and couplet based on neural network [D]. Guilin: Guilin University of Electronic Technology, 2019 (in Chinese)
(卫万成. 基于神经网络的诗歌与对联自动生成方法研究[D]. 桂林: 桂林电子科技大学, 2019)

[25] Zhang Jiang. Research and implementation of Chinese couplet generation system based on deep neural network [D]. Nanjing: Nanjing University of Posts and Telecommunications, 2020 (in Chinese)
(张江. 基于深度神经网络的对联生成系统的研究与实现[D]. 南京: 南京邮电大学, 2020)

[26] Bu Tian. Research on the mechanism of double-layer lower couplet generation [D]. Hangzhou: Zhejiang Sci-tech University, 2022 (in Chinese)
(卜天. 基于双层注意力机制的对联下联生成方法研究[D]. 杭州: 浙江理工大学, 2022)

[27] Le Q, Mikolov T. Distributed representations of sentences and documents [C] // Proc of the 31st Int Conf on Machine Learning (ICML). New York: PMLR, 2014: 1188−1196



Li Xinran, born in 1992. PhD. Her main research interests include multimedia information security and information hiding.

李欣然, 1992 年生. 博士. 主要研究方向为多媒体信息安全、信息隐藏.



Li Rongshou, born in 1997. Master. Her main research interest includes text information hiding. (202430383@st.usst.edu.cn)

李蓉受, 1997 年生. 硕士. 主要研究方向为文本信息隐藏.



Qin Chuan, born in 1980. PhD, professor, PhD supervisor. Senior member of CCF. His main research interests include multimedia information security and AI security.

秦　川, 1980 年生. 博士, 教授, 博士生导师. CCF 高级会员. 主要研究方向为多媒体信息安全、AI 安全.



Qian Zhenxing, born in 1981. PhD, professor, PhD supervisor. His main research interests include information hiding and AI security. (zxqian@fudan.edu.cn)

钱振兴, 1981 年生. 博士, 教授, 博士生导师. 主要研究方向为信息隐藏、AI 安全.



Zhang Xinpeng, born in 1975. PhD, professor, PhD supervisor. His main research interests include multimedia information security and AI security. (zhangxinpeng@fudan.edu.cn)

张新鹏, 1975 年生. 博士, 教授, 博士生导师. 主要研究方向为多媒体信息安全、AI 安全.