基于宋词生成的大容量构造式信息隐藏算法

秦 川[1)] 李蓉受[1)] 钱振兴[2)] 张新鹏[2)]

[1)](上海理工大学光电信息与计算机工程学院 上海 200093)

[2)](复旦大学计算机科学技术学院 上海 200433)

摘 要 在基于文本生成的信息隐藏算法研究中,如何在保证生成文本质量的同时提高隐藏容量是主要存在的挑战. 为此本文提出一种基于宋词生成的构造式信息隐藏算法. 首先对宋词文本数据进行预训练，然后基于自回归语言模型搭建宋词生成模型；其次根据宋词词牌固有的格式信息设计格律模块，在宋词生成阶段，需要向生成模型输入该格律模块，并通过符号集设计、编码等综合作用，生成宋词诗句. 在利用宋词生成模型进行秘密信息隐藏的过程中，对格律模块进行重构，通过平仄韵词牌、词牌格式模板、关键字、韵律及押韵字符的不同选择，有效实现秘密信息的隐藏. 信息提取是隐藏的逆过程，且提取过程不需要利用宋词生成模型，仅需根据模板和词典库来进行索引即可，提高了信息提取的效率. 实验结果表明，本文提出的算法能够生成格式严格、韵律清晰、句子完整性高的宋词，且生成的宋词文本的信息隐藏容量均值可达 21 比特/句、安全性高，整体性能优于已报道的主流算法.

关键词 文本生成；构造式信息隐藏；宋词；格律控制；隐藏容量

中图法分类号 TP309 **DOI 号** 10.11897/SP.J.1016.2023.00017

Large-Capacity Constructive Information Hiding Based on Song Ci Generation

QIN Chuan[1)] LI Rong-Shou[1)] QIAN Zhen-Xing[2)] ZHANG Xin-Peng[2)]

[1)](*School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai* 200093)

[2)](*School of Computer Science, Fudan University, Shanghai* 200433)

Abstract As the typical media of information dissemination, text data is one kind of the commonly-used cover data in the field of information hiding research, which has attracted extensive attentions in the community of information security. At present, the technique of text generation based on natural language processing (NLP) becomes more and more popular, but its application in the field of information hiding is not very satisfactory, which is still in the elementary stage. As for the current research of information hiding algorithms based on text generation, how to improve the hiding capacity while ensuring the quality of the generated texts is the main challenge. Therefore, in this paper, we propose a novel constructive information hiding algorithm based on Song Ci generation. Firstly, the text data of Song Ci is pre-trained, and a set of customized indicators are introduced to improve the modeling performance, especially the robustness in format, rhythm and sentence integrity. Then, the Song Ci generation model in our algorithm is established based on the attention mechanism, whose backbone is the transformer-based auto-regressive language model. Secondly, the metrical module is designed according to the inherent format information of Song Ci tunes. In the stage of Song Ci generation, the metrical module is inputted into the Song Ci generation model,

收稿日期：2021-10-22；在线发布日期：2022-06-07. 本课题受国家自然科学基金面上项目(No.62172280)、国家自然科学基金重点项目(No. U20B2051)、上海市自然科学基金项目(No. 21ZR1444600)的资助.**秦 川**(通信作者)，博士，教授，博士生导师，中国计算机学会(CCF)高级会员，主要研究领域为多媒体信息安全、AI安全. E-mail: qin@usst.edu.cn. **李蓉受**，硕士研究生，主要研究领域为文本信息隐藏. **钱振兴**，博士，教授，主要研究领域为信息隐藏、AI安全. **张新鹏**，博士，教授，主要研究领域为多媒体信安全、AI安全.

and then the Song Ci poetries can be generated through the comprehensive action of symbol set design and coding. In the process of information hiding by using the Song Ci generation model, the metrical module is reconstructed through choosing different tunes, tune templates, keywords, rhythms and rhyming characters according to secret information. Information extraction is the inverse process of information hiding, and the extraction process does not need to use the Song Ci generation model, but only needs to use the index mapping according to the metrical template and dictionary library, which can improve the efficiency of information extraction. Experimental results show that the proposed algorithm can generate Song Ci with strict format, clear rhythm and high sentence integrity, and the generated Song Ci have high security and the information hiding capacity that reaches 21 bits/sentence in average. The overall performance of the proposed algorithm is significantly better than some reported algorithms. We also utilized the same set of training data on our model with information hiding and the current typical poetry generation model without information hiding, respectively, and the trained results show that the performance indices of the two models, such as the perplexity, are close. Then, by comparing a series of Song Ci poetries randomly generated by these two models, we can find that the performances of semantic quality for the generated Song Ci poetries are quite good. Moreover, we also make an experimental analysis on the anti-semantic similarity detection of our algorithm, and experimental results demonstrate that the Song Ci with information hiding and the Song Ci without information hiding can hardly be distinguished in the semantic space. In addition, the time complexity, tampering detection analysis and security of the algorithm are also discussed. Finally, the future research directions of this work are given.

Keywords text generation; constructive information hiding; Song Ci; metric control; hiding capacity

1 引 言

在互联网时代，信息安全越发引人关注. 相较于传统的加密技术，信息隐藏是将秘密信息隐藏于公开的各种载体数据中，除收发双方而不被人所察觉的技术，隐蔽性强. 信息隐藏如今已经成为一个热门的研究课题，已被广泛应用于多个领域，如军事、法律和知识产权保护等. 在诸多信息载体如视频[1]、音频[2]、图像[3]及文本当中，文本作为人们最常用于获取信息的媒介之一，具有传输效率高、处理方便等优点，故基于文本载体的信息隐藏研究具有重要意义. 基于文本载体的信息隐藏技术主要可以分为两类，即基于载体格式的方案和基于载体内容的方案[4].

基于文本载体格式的信息隐藏方案，一般是通过修改文本字符之间的格式特征来隐藏秘密信息，如调整字符的间距、行距、大小、颜色、字体风格等[5-9]. 这些格式的细微调整，一般来说人类视觉系统难以直接发现异常之处. 文献[5]通过改变单词之间的间距来隐藏信息. 文献[6]通过在 HTML 文本中嵌入特殊的空格字符进而实现信息隐藏. 文献[7]通过在 HTML 文本中控制字体的颜色编码，从而达到秘密信息通信的目的. 文献[8-9]通过不可见字符来实现秘密信息隐藏，其中文献[8]通过改变字体的类型和样式，将秘密信息隐藏于空格中；文献[9]利用 Unicode 零度宽字符实现信息隐藏. 基于文本载体格式修改的信息隐藏方案不会改变文本的内容语义，但是如果对含密文本进行轻微的格式修改即会导致信息提取的失败. 所以，基于格式修改的方案鲁棒性不强，安全性不高，且易于被检测出异常.

基于文本载体内容的信息隐藏方案，通常是基于词汇、句法或语义的操作来进行信息隐藏，根据操作方式的不同大致可分为两类. 第一类是基于文本语义修改的方案. 这类方案通常是修改文本载体的语义内容来隐藏信息. 如文献[10-11]使用同义词替换策略及改变语句结构等. 该类方案通过一般的阅读基本无法感知异常，隐蔽性较高，但难免会出现因为替换修改而导致的语义失真问题；另外由于可供替换的词汇量是相对有限的，容易受到隐写分析攻击[12]，且隐藏容量不高；第二类是基于文本生成的信息隐藏方案，这类方案通常是由秘密信息的驱动以直接生成含密文本，不需对文本内容或格式进行任何修改，隐蔽性更强，且隐藏容量较高，故逐渐成为文本信息隐藏研究的热点. 下面对基于文

本生成的信息隐藏的一些代表性方法进行介绍.

文献[13]提出一种基于马尔可夫模型和霍夫曼编码的文本信息隐藏算法，利用状态转移图中的条件概率进行编码，但隐藏容量不理想. 随着深度学习在机器翻译中的广泛应用，基于机器翻译模型进行文本创作的技术日渐成熟. 文献[14]提出一种基于循环神经网络(Recurrent Neural Network, RNN)编解码结构生成绝句的信息隐藏算法，通过对候选库的候选字进行编码从而达到信息隐藏的目的，取得了较大的隐藏容量. 但文献[14]因选择候选字组合成句时会产生较大的损失，导致生成的绝句质量一般，且该算法在提取时需要遍历模板，因此提取秘密信息的计算复杂度较高. 文献[15]对[14]进行了改进，提出了一种基于绝句生成的构造式信息隐藏算法. 该算法在工作记忆模型的基础上，于生成绝句时增加了主题信息，因此生成的绝句质量较高，但同时因为每次迭代受到关键词的限制，隐藏容量没有显著的提高. 文献[16]采用 GPT-2 语言模型作为文本生成模型，提出了基于自适应算术编码的文本信息隐藏算法，提升了隐藏容量，但增加了时间复杂度. 为了解决感知不可察觉性与统计不可察觉性的冲突问题，文献[17]采用 BERT(Bidirectional Encoder Representations from Transformers)模型作为编码器，提出一种 VAE-Stega 方案，进一步改善了生成的含密文本的不可感知性，但并没有提高隐藏容量.

目前学术界在基于中国古诗词机器化的研究方面已取得了一系列初步成果. 宋词作为古诗词中流传广泛的一种诗歌载体，其词牌对应确定的格式模板，具有韵律清晰，格式工整的特点. 与一般性文本相比，宋词的语法明确，机器化可实现度更高. 在古诗词中，宋词的长度相对更长，可隐藏的秘密信息容量更大. 文献[18]提出一种基于宋词生成的信息隐藏算法，根据宋词固有的词牌模板选取不同的词汇来构造新的宋词文本以实现秘密信息的隐藏，但该算法只考虑了宋词的音律规则，并没有利用词汇的语义联系，生成的宋词缺乏中心思想，可读性较差，鲁棒性较弱，且隐藏容量较低. 文献[19]针对文献[18]的不足进行了改进，隐藏容量有所提高，但在词汇间的语义联系没有提升，生成的宋词质量不高. 文献[20]提出一种基于混合加密的宋词生成模型，提高了安全性，但在隐藏容量方面表现一般，且生成的宋词质量不理想. 文献[21]利用宋词韵律“中平仄”出现频率的奇偶性进行信息隐藏，但其隐藏容量相对于其它文献方法没有提升.

针对上述文本信息隐藏算法所存在的问题和不足，本文提出一种基于宋词生成的构造式信息隐藏算法，在生成具有较高质量宋词的同时，实现了较大的信息隐藏容量. 本文算法通过添加字符的韵律信息，使得生成宋词的可读性更高. 在隐藏容量方面，本文算法根据生成宋词的词牌、词牌格式模板、关键字、韵律和押韵字符进行秘密信息的隐藏，具有较高的隐藏容量. 另外，本文算法还根据宋词的格律和曲调信息，设计词牌的选择，这样可在不影响隐藏容量的情况下，进一步保证宋词的质量. 实际上，本文提出的构造式信息隐藏算法除宋词之外，还可拓展应用于其它具有严格模板要求的文本生成与信息隐藏中，如元曲、对联、歌词、十四行诗等(需注意是，针对不同类型的文本需使用对应文本的数据集进行训练，同时使用相应的模板和韵律规则). 具体应用场景包括可以安全隐蔽通信，及其对相关文学作品的所有权保护和完整性认证等.

本文工作的贡献主要在以下三个方面：(1)设计了宋词生成的格律模块，通过利用词牌、格式模板、关键字、韵律以及押韵字符的选择信息实现秘密信息的隐藏；(2)增加了韵律编码，通过韵律的选择设计来提高生成诗歌韵律的准确性；(3)本文算法在隐藏容量方面有着较好的表现，且可对提取出的秘密信息进行校验，判断含密文本是否受到篡改. 相比较于载体选择式的信息隐藏方案，本文提出的基于宋词生成的构造式信息隐藏算法在算法安全性方面表现更加可靠. 构造式信息隐藏算法在文本生成的过程中没有确定的词库，通过模型计算来进行字与字之间的语义联系继而进行选择和匹配，所以在没有完全相同数据集的情况下，攻击者想攻击和分析本文算法的难度非常高. 在生成文本的质量方面，构造式信息隐藏算法在生成文本时更加注重字与字之间的语义联系以及整体句子的完整性，因而本文算法在生成文本的质量方面表现更良好.

结构内容安排如下：第 2 节给出本文算法的整体框架；第 3 节介绍基于格律控制的宋词生成模型；第 4 节为信息隐藏和提取方法；第 5 节为实验结果及分析；第 6 节是全文总结与进一步的工作.

2　算法整体框架

本文提出的基于宋词生成的构造式信息隐藏算法主要分为四个部分：模型训练、构建数据集、信息隐藏和信息提取. 算法的整体框架如图 1 所示.

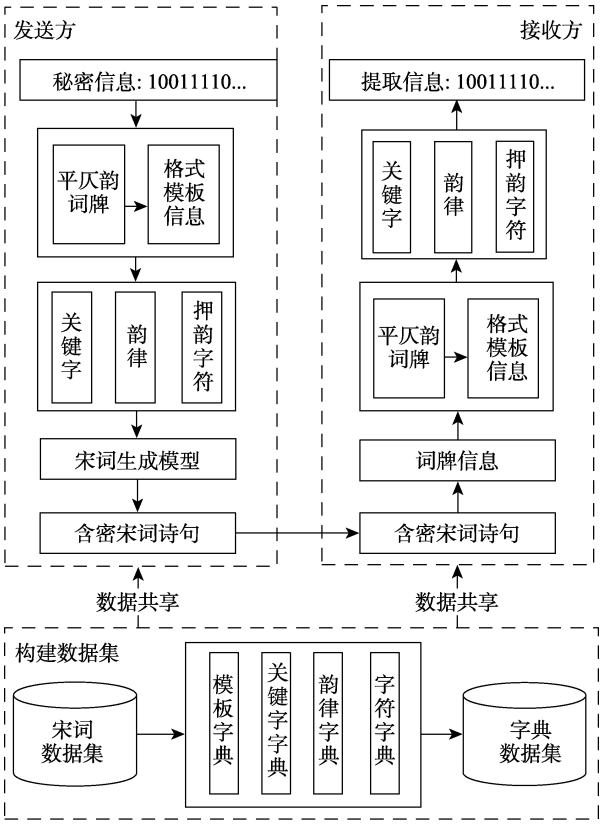

图 1 整体算法框架图

在模型训练的过程中，根据本文提出的方案进行训练，并经过微调抛光，最终获得符合秘密信息隐藏要求的宋词生成模型. 在数据库构建的过程中，根据模型训练获取的关键字等相关参数，同时根据词牌格式模板信息、韵律等得到符合要求的映射字典. 收发双方需要共享相关参数，确保双方训练得到的字典词库一致.

在信息隐藏的过程中，首先根据秘密信息进行平仄韵词牌和词牌格式模板的选择，然后依据词牌格式模板和隐藏算法来判定模板所能隐藏的关键字和押韵字符的数量信息，从而根据秘密信息选择对应的关键字、韵律及押韵字符，最后输入模型，生成含密的宋词诗句. 在信息提取的过程中，首先依据接收到的含有秘密信息的宋词诗句，利用模型提取平仄韵词牌和词牌格式模板的信息，再利用收发双方共享的相关参数，提取含密宋词诗句中的关键字、韵律和押韵字符的信息.

3 宋词生成模型

本文的宋词生成模型的主干是基于 SongNet 模型，而 SongNet 模型是启发于 BERT 改进的自回归语言模型. 具体地，本文模型利用 SongNet 对具有格式要求的文本的生成能力，设计了宋词的格律生成模块，并利用该模块的格律信息实现了秘密信息的隐藏，从而生成含密宋词. 因此我们先对 BERT 模型和 SongNet 模型进行相关介绍，接着于第 4 节介绍本文提出的宋词的格律生成模块.

3.1 BERT模型

BERT 是双向的 Transformer 模型[22]，其输入可是单个句子或句子对，BERT 模型的框架如图 2 所示. 字符编码是将输入的自然语言文本转换成词向量表示的过程，以字为单位进行分割，结果仅包含文本的字符序列信息. 分段编码是为记住输入的字符所属句段的位置；位置编码是为解决自注意力机制难以识别字符位置信息的问题，以提高模型对单个字符的位置敏感度，并能增强模型的建模能力.

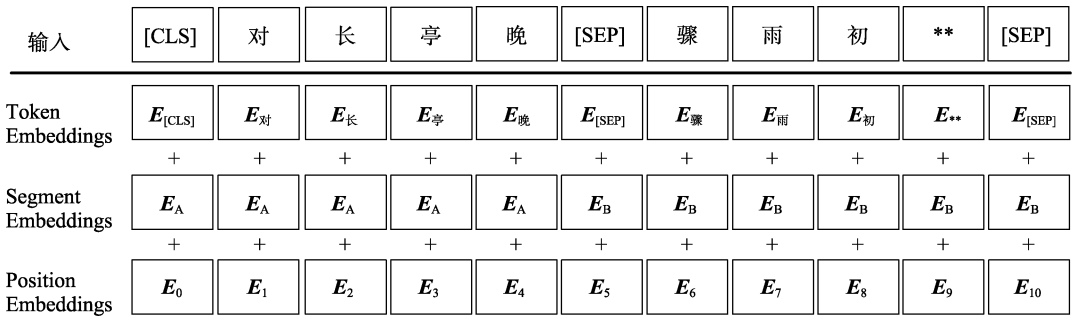

图 2 BERT 模型框架图

由于 BERT 本身并不具有生成文本的能力，为使其转换成文本生成模型，即将 BERT 变为自回归语言模型，可采用遮掩策略，其核心是遮蔽的注意力(Masked Multi-Head Attention)算法[23]. 下面以一个例子进行说明，如图 3 所示，图中的三个模块从左到右分别是自回归语言模型在五个时间步的信息、用矩阵所表示的 Mask 策略和 BERT 在五个时间步的信息. BERT 的五个时间步的信息均相同，即“明月几时有”；自回归语言模型在各个时间步的信息为：第一个时间步为“明”，第二个时间步为“明

月”, 第三个时间步为“明月几”, 第四个时间步为“明月几时”, 第五个时间步为“明月几时有”, 即 BERT 通过遮掩策略可转换为自回归语言模型.

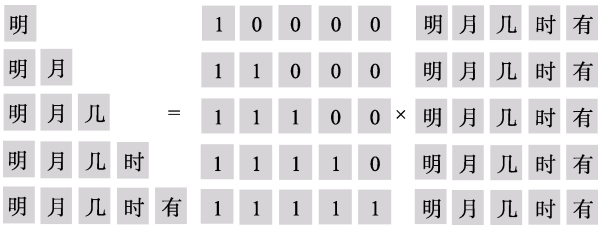

图 3　BERT 转换示例图

3.2　SongNet模型基本原理

在将 BERT 模型转换成自回归语言模型的基础上，文献[24]提出 SongNet 模型，其模型框架如图 4 所示，其中例句摘自柳永的《西平乐》. 该模型的主要任务是解决严格格式控制的文本生成问题，模型的输入有作者信息，词牌名和模板信息. 特别地，模板的格式设定是任意的，可以是与词牌名对应模板信息，也可以是自行创建的新模板. SongNet 模型包含符号集设计的模块，注意力机制修改模块以及预训练与微调模块.

SongNet 模型运用了一系列符号来提高格宋词句子的完整性、格式和韵律的准确性. 具体来说，格式和韵律符号 $C=\{c_i\}$，用于格式和押韵的建模；内部位置符号 $P=\{p_i\}$，用于捕捉句子的动态信息，获取每个句子结束的位置信息，从而提高句子的完整性，其中 p_i 表示同一子句中标记的本地位置；分段符号 $S=\{s_i\}$，用于识别句子的边界，使模型能够捕捉到押韵的句对，其中 s_i 表示第 i 个句子的序列索引信息.

在训练的过程中，对输入至模型中的所有文本字符向量和符号进行编码并求和（见图 4 中实线框部分）:

$$\boldsymbol{H}_t^{(0)}=\boldsymbol{E}_{w_t}+\boldsymbol{E}_{c_t}+\boldsymbol{E}_{p_t}+\boldsymbol{E}_{s_t}+\boldsymbol{E}_{g_t} \tag{1}$$

其中，t 是状态索引，$\boldsymbol{E}$ 是输入的 Embedding 向量，$\boldsymbol{H}$ 表示输入的 Embedding 向量之和,(0)是层索引(索引可从 0，1 迭代至模型的总层数)，w_t 是位置 t 的 token 标记，c、p 和 s 是三个预设的符号，g 是全局位置索引. t 时刻的状态需要在自回归的动态过程中获得未来的信息，进而才能掌握序列的全局动态信息. 因此，这里引入 $\boldsymbol{F}^{(0)}$来表示针对设计的符号集进行求和运算（见图 4 中虚线框部分）:

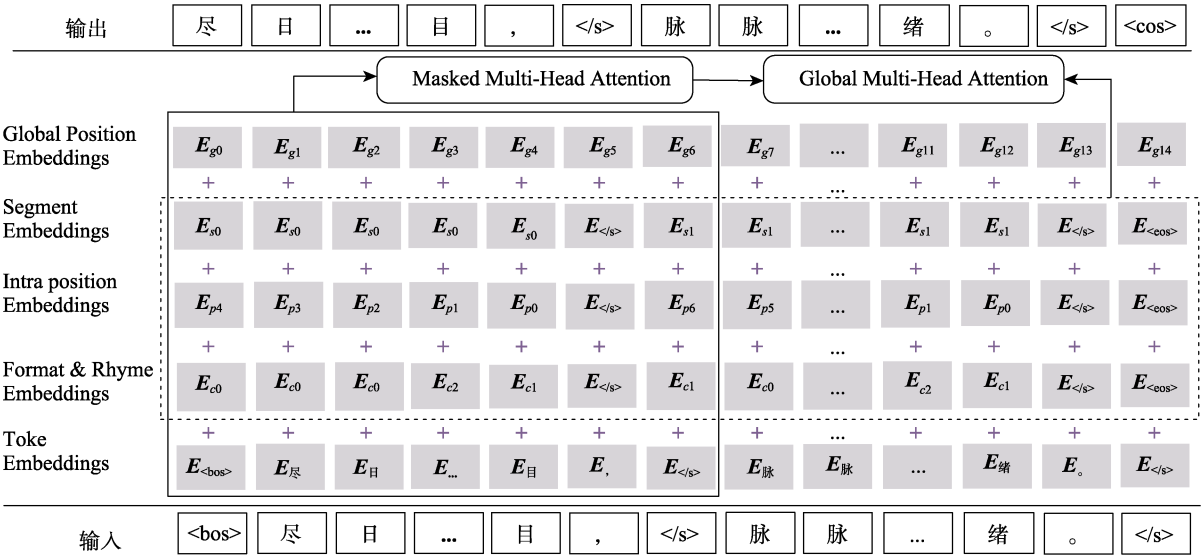

图 4　SongNet 模型框架图

$$\boldsymbol{F}_t^{(0)}=\boldsymbol{E}_{c_t}+\boldsymbol{E}_{p_t}+\boldsymbol{E}_{s_t} \tag{2}$$

在对输入数据进行处理之后，引入两个注意力机制模块进行特征学习，注意力机制模块使用索引小于和等于 t 的状态. 第一个模块是 Masked Multi-Head Attention，其计算公式如下：

$$\left[\boldsymbol{Q}^{(0)},\boldsymbol{K}^{(0)},\boldsymbol{V}^{(0)}\right]=\left[\boldsymbol{H}^{(0)}\boldsymbol{W}^{Q},\boldsymbol{H}^{(0)}\boldsymbol{W}^{K},\boldsymbol{H}^{(0)}\boldsymbol{W}^{V}\right] \tag{3}$$

$$\boldsymbol{C}_t^{(1)}=\mathrm{L_N}\left(\mathrm{A_S}\left(\boldsymbol{Q}_t^{(0)},\boldsymbol{K}_t^{(0)},\boldsymbol{V}_t^{(0)}\right)+\boldsymbol{H}_t^{(0)}\right) \tag{4}$$

$$\boldsymbol{C}_t^{(1)}=\mathrm{L_N}\left(\mathrm{F_N}\left(\boldsymbol{C}_t^{(1)}\right)+\boldsymbol{C}_t^{(1)}\right) \tag{5}$$

其中 $\mathrm{A_S}(\cdot)$表示自注意力机制，$\mathrm{L_N}(\cdot)$表示层的归一化，$\mathrm{F_N}(\cdot)$表示前馈网络. $\boldsymbol{C}$ 为 Masked Multi-Head Attention 模块获得的内容向量，$\boldsymbol{Q}$、$\boldsymbol{K}$、$\boldsymbol{V}$ 分别是注意力机制中的 query 集合矩阵、key 集合矩阵、value

集合矩阵，$\boldsymbol{W}$ 为注意力机制模块中权重表示，$\boldsymbol{W}^{Q}$、$\boldsymbol{W}^{K}$ 和 $\boldsymbol{W}^{V}$ 为 $\boldsymbol{Q}$、$\boldsymbol{K}$、$\boldsymbol{V}$ 对应的权重变换参数. 第二个模块是 Global Multi-Head Attention，用 $\mathrm{A_G}()$表示. 第一个注意力机制模块捕获 $\boldsymbol{C}_t^{(1)}$，将 $\boldsymbol{C}_t^{(1)}$ 输入第二个注意力机制模块同时该模块从 $\boldsymbol{F}^{(0)}$中获取信息，继而获取全局动态信息，计算公式如下：

$$\left[\boldsymbol{Q}^{(1)},\boldsymbol{K}^{(1)},\boldsymbol{V}^{(1)}\right]=\left[\boldsymbol{C}^{(1)}\boldsymbol{W}^{Q},\boldsymbol{F}^{(0)}\boldsymbol{W}^{K},\boldsymbol{F}^{(0)}\boldsymbol{W}^{V}\right] \tag{6}$$

$$\boldsymbol{H}_t^{(1)}=\mathrm{L_N}\left(\mathrm{A_G}\left(\boldsymbol{Q}_t^{(1)},\boldsymbol{K}_t^{(1)},\boldsymbol{V}_t^{(1)}\right)+\boldsymbol{C}_t^{(1)}\right) \tag{7}$$

$$\boldsymbol{H}_t^{(1)}=\mathrm{L_N}\left(\mathrm{F_N}\left(\boldsymbol{H}_t^{(1)}\right)+\boldsymbol{H}_t^{(1)}\right) \tag{8}$$

4 基于宋词生成的信息隐藏和提取

4.1 生成格律模块

在保证生成文本质量的同时，如何进行高效的秘密信息隐藏是需要解决的问题. SongNet 模型可以针对歌词、宋词和十四行诗三种具有严格格式的载体进行填词生成文本. 通过对 SongNet 模型生成文本的原理进行分析，不难发现文本的生成严格受控制于输入的模板信息，生成的文本与给定的格式等有紧密的关系. 因此，如果格式模板在生成的过程中带有确定的格律信息，而格律信息与秘密信息存在对应关系，即可利用生成宋词实现秘密信息的隐藏和提取. 在宋词文本生成的过程中，格式模板中的格律信息不会在诗句生成的过程中丢失，也不依赖于后续诗句的生成，这样可在有效保证生成文本质量的同时，进行秘密信息的隐藏. 基于以上分析，下面给出基于宋词生成的信息隐藏方案.

本文的信息隐藏方案基于 SongNet 模型进行格律模块的设计，即严格控制格律的生成过程. 通过调整训练阶段的参数以及训练方法，保证格律信息在宋词生成的过程中保持不变. 利用输入格律在自动生成宋词文本的过程中隐藏秘密信息. 在信息隐藏方案中，我们增加了韵律的惩戒因子，确保每个韵律更加精确地表达信息，并且使得模型生成的文本更加符合格律要求. 在信息提取时，本文方案相比于其他文献可以更加精准判定秘密信息，有效保证了提取的秘密信息的正确率.

模型训练过程中，利用设计的符号集，能够学习到词牌格式模板中的位置信息，确保关键字、韵律及押韵字符信息能够自然地隐藏到诗句的预设位置，同时保证生成文本的质量不受影响. 格律信息所能出现的位置是发送方和接收方双方的共享信息，且可根据需要调整变化.

4.2 信息隐藏

宋词的生成通常需先确定词牌，每个词牌都有固定的声调、韵律和句子长度. 依据《平水韵》，声调一般有两种类型：“平声”和“仄声”；对应地，词牌韵律有“平韵”和“仄韵”两类. 在宋词的韵律情况中，叠韵的情况为极少数，且其对文本算法的影响甚微，故暂不考虑叠韵的情况. 基于以上分析，在进行秘密信息隐藏前我们将秘密信息分为两段，将平仄韵词牌和词牌格式模板所对应的秘密信息定义为第一段秘密信息；将关键字、韵律和押韵字符所对应的秘密信息定义为第二段秘密信息. 信息隐藏的具体过程可分为两步：(1)根据第一段秘密信息确定平仄韵词牌和词牌格式模板；(2)根据第二段秘密信息确定关键字、韵律和押韵字符，生成基于词牌格式控制的格律模块. 以词牌《少年游》为例，图 5 给出了秘密信息的隐藏的流程图.

在第一步中，主要工作是根据第一段秘密信息确定词牌信息. 词牌信息主要分为两个部分：平仄韵词牌(Tune)与词牌格式模板(Pattern). 因为词牌有平韵词牌、仄韵词牌两类，故 Tune 的数值设置为 $T_{\mathrm{num}}=2$. 宋词中词牌名都有其对应的格式模板，且模板一般与词牌名的选择相对应，故词牌名与模板具有映射关系. 为避免词牌名和模板的信息重复，可采用二取一的方式来隐藏信息. 本文算法选取模板来隐藏信息. 训练集中共有词牌 1740 种，我们选取其中 768 种常用词牌用于宋词生成及信息隐藏. 因此这里设置的 Pattern 有两种：一种对应平韵的 Pattern，个数设置为 $P_{\text{p-num}}=256$；另一种对应仄韵的 Pattern，个数设置为 $P_{\text{z-num}}=512$. 设第一段秘密信息为 L' 比特，则 L' 根据词牌选择分两种情况，即 L_{p}' 和 L_{z}'. 对应平韵词牌的 L_{p}' 的计算公式为

$$L_{\mathrm{p}}'=\log_2\left(T_{\mathrm{num}}\right)+\log_2\left(P_{\text{p-num}}\right)=\log_2\left(T_{\mathrm{num}}\cdot P_{\text{p-num}}\right) \tag{9}$$

对应仄韵词牌的 L_{z}'的计算公式：

$$L_{\mathrm{z}}'=\log_2\left(T_{\mathrm{num}}\right)+\log_2\left(P_{\text{z-num}}\right)=\log_2\left(T_{\mathrm{num}}\cdot P_{\text{z-num}}\right) \tag{10}$$

在信息隐藏过程的第二步中，依据第一步选择的词牌格式模板进行. 假设关键字的数量为 K_{num}，模板所能隐藏的关键字个数由在词牌格式模板中选择的预设符号个数来确定. 其中预设符号是根据模板中带有符号所设计，一般词牌格式模板当中的符号有 5 种：逗号、顿号、句号、问号和感叹号，因此预设符号包含有 5 种. 本文选取逗号(Comma Symbol)为隐藏关键字的预设符号，设每个词牌格式模板隐藏关键字的数量为 $S_{\text{c-num}}$. 关键字所隐藏的位置可以是任意设置的，如可设置为距离逗号前 1 或 2 个位

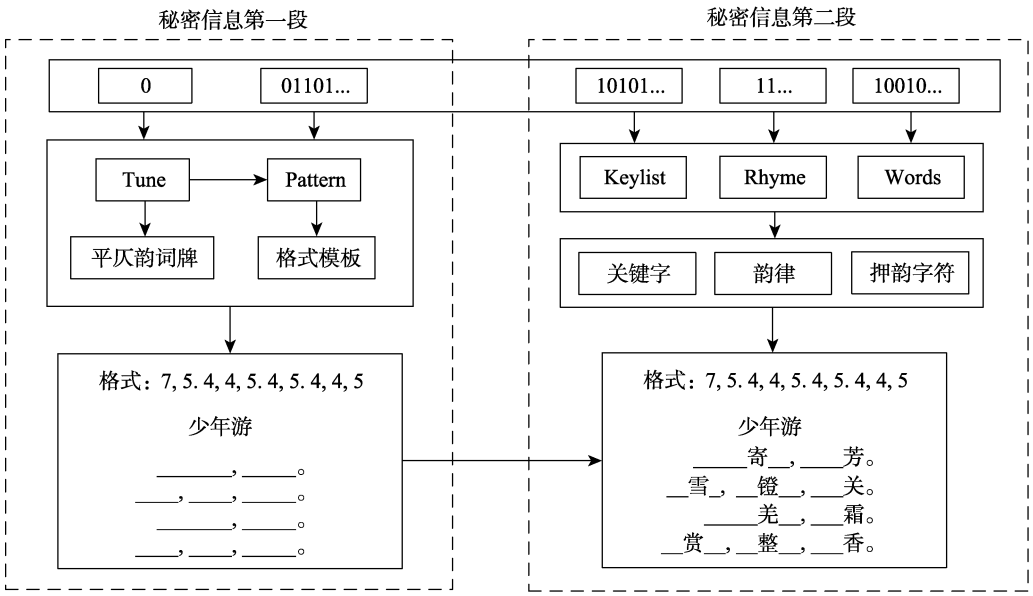

图 5　信息隐藏流程图

置为隐藏关键字的位置，并可根据需要调整模型. 在词牌韵律(Rhyme)方面，选取平韵的种类数为 $R_{\text{p-num}}=16$，仄韵的种类数为 $R_{\text{z-num}}=64$. 押韵字符(Rhyme word)根据选择韵律的不同而数值不一，这里假设每种押韵字符数值为 W_{num}. 宋词中押韵字符极大多数出现在每句话的末尾，故选取句号(Full Stop)为隐藏押韵字符的预设符号，即词牌格式模板隐藏押韵字符的数量和句号(Full Stop)的个数一致，设为 $S_{\text{f-num}}$. 设第二段秘密信息为 L''比特，则 L''根据平韵和仄韵词牌分为两种情况，即 L_{p}''和 L_{z}''. 对应平韵词牌的 L_{p}''的计算公式为

$$\begin{aligned}L_{\text{p}}''=&S_{\text{c-num}}\cdot\log_2\left(K_{\text{num}}\right)+\log_2\left(R_{\text{p-num}}\right)\\&+S_{\text{f-num}}\cdot\log_2\left(W_{\text{num}}\right)\end{aligned}\tag{11}$$

对应仄韵词牌的 L_{z}''的计算公式为

$$\begin{aligned}L_{\text{z}}''=&S_{\text{c-num}}\cdot\log_2\left(K_{\text{num}}\right)+\log_2\left(R_{\text{z-num}}\right)\\&+S_{\text{f-num}}\cdot\log_2\left(W_{\text{num}}\right)\end{aligned}\tag{12}$$

因此，一首宋词隐藏秘密信息的总比特数 L 根据选择平韵或仄韵词牌分为两种情况，即 L_{p} 和 L_{z}. 对应平韵词牌的 L_{p} 的计算公式为

$$\begin{aligned}L_{\text{p}}&=L_{\text{p}}'+L_{\text{p}}''\\&=\log_2\left(T_{\text{num}}\cdot P_{\text{p-num}}\right)+S_{\text{c-num}}\cdot\log_2\left(K_{\text{num}}\right)\\&\quad+\log_2\left(R_{\text{p-num}}\right)+S_{\text{f-num}}\cdot\log_2\left(W_{\text{num}}\right)\\&=\log_2\left(T_{\text{num}}\cdot P_{\text{p-num}}\cdot R_{\text{p-num}}\right)+S_{\text{c-num}}\cdot\log_2\left(K_{\text{num}}\right)\\&\quad+S_{\text{f-num}}\cdot\log_2\left(W_{\text{num}}\right)\end{aligned}\tag{13}$$

对应仄韵词牌的 L_{z} 的计算公式为

$$\begin{aligned}L_{\text{z}}&=L_{\text{z}}'+L_{\text{z}}''\\&=\log_2\left(T_{\text{num}}\cdot P_{\text{z-num}}\right)+S_{\text{c-num}}\cdot\log_2\left(K_{\text{num}}\right)\\&\quad+\log_2\left(R_{\text{z-num}}\right)+S_{\text{f-num}}\cdot\log_2\left(W_{\text{num}}\right)\\&=\log_2\left(T_{\text{num}}\cdot P_{\text{z-num}}\cdot R_{\text{z-num}}\right)+S_{\text{c-num}}\cdot\log_2\left(K_{\text{num}}\right)\\&\quad+S_{\text{f-num}}\cdot\log_2\left(W_{\text{num}}\right)\end{aligned}\tag{14}$$

由公式(13)和(14)可以看出，每首宋词能隐藏的秘密信息长度与词牌的格律模块密切相关. 影响本文算法信息隐藏容量的主要因素包括：平仄韵词牌 $T_{\text{num}}=2$；词牌格式模板（平韵词牌模板 $P_{\text{p-num}}=256$、仄韵词牌模板 $P_{\text{z-num}}=512$）；韵律种类（平声韵律 $R_{\text{p-num}}=16$、仄声韵律 $R_{\text{z-num}}=64$）；关键字的个数 K_{num}；每种押韵字符的个数 W_{num}，其中关键字和押韵字符的参数均可根据实际情况进行设置.

接下来以词牌《渔家傲》为例来介绍本文算法过程，如图 6 所示，图中将对应词牌模板、韵律、押韵字符和关键字信息的比特信息划分. 该实例依据秘密信息对应选择平仄韵词牌、词牌格式模板、韵律、押韵字符和关键字，首先秘密信息“1”对应选择平仄韵词牌《渔家傲》，秘密信息“1011…”对应选择词牌格式模板“_____，_____，_____. ___，_____. _____，_____，_____. ___，_____.”. 在该实例中，算法设定隐藏押韵字符的预设符号为“句号”，隐藏位置为句末，统计词牌模板信息可知能隐藏押韵字符 4 个. 故秘密信息“110…1”对应选择韵律信息，即仄韵类别和押韵字符“晓、好、笑、了”. 接着，算法设定隐藏关键字的预设符号为“逗号”，隐藏位置为逗号前两个位置，统计模板

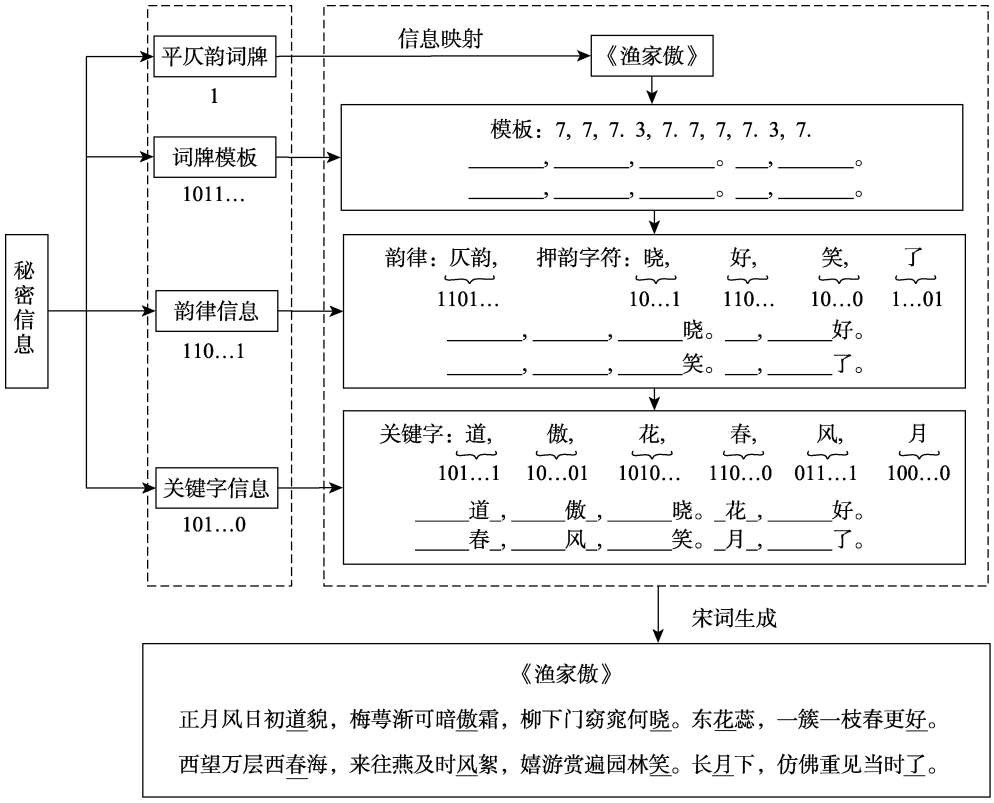

图 6 宋词实例图

信息可知能隐藏关键字 6 个. 故秘密信息“101…0”对应选择关键字“道、傲、花、春、风、月”. 至此，构建完成格律模块《渔家傲》“_____道_, _____傲_, _____晓. _花_, _____好. _____春_, _____风_, _____笑. _月_, _____了. ”,最终生成含密宋词《渔家傲》“正月风日初道貌，梅萼渐可暗傲霜，柳下门窈窕何晓. 东花蕊,一簇一枝春更好. 西望万层西春海，来往燕及时风絮，嬉游赏遍园林笑. 长月下，仿佛重见当时了. ”

4.3 信息提取

信息提取过程基本是信息隐藏的逆过程. 接收方的信息提取具体可分为三个步骤：(1)从宋词获取平仄韵词牌(Tune)的信息；(2)根据第 1 步获取的平仄韵词牌信息，进一步获取词牌信息下的词牌格式模板(Pattern)，即提取第一段秘密信息；(3)根据第 2 步获取的词牌格式模板信息，获取关键字(K_{num})、韵律(Rhyme)和押韵字符(W_{num})信息，完成第二段秘密信息的提取. 秘密信息提取的流程图如图 7 所示.

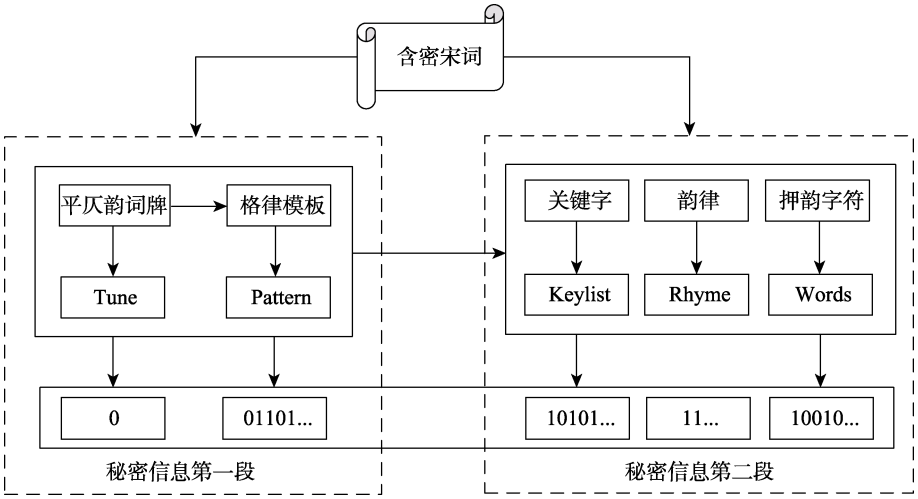

图 7 信息提取流程图

具体来说，在第 1 步中，首先根据平仄韵词牌信息的隐藏规则，在含密宋词中提取平仄韵词牌 Tune 的信息. 在第 2 步中，根据第 1 步获取的词牌信息，根据对应的隐藏规则判断宋词诗句的词牌是平韵词牌还是仄韵词牌. 若是平韵词牌，则对照平韵词牌信息表，查找词牌格式模板 Pattern 的信息. 若是仄韵词牌，同理可获得词牌格式模板 Pattern 的信息. 通过以上步骤，即可提取出第一段秘密信息. 在第 3 步中，根据第 2 步提取的词牌格式模板信息，统计模板中逗号的数量 $S_{\text{c-num}}$ 和句号的数量 $S_{\text{f-num}}$，再根据关键字的隐藏规则，对照关键字信息表，提取关键字对应的秘密信息. 然后根据韵律的隐藏位置提取韵律，在判断韵律类型 Rhyme 的过程中，我们对多音字进行区分，确保 Rhyme 类型判定的唯一性. 对应词牌的平仄韵信息，韵律表可分为平声韵律表和仄声韵律表，依据韵律和押韵字符的隐藏规则，对照信息表可提取出韵律和押韵字符所对应的秘密信息. 通过以上步骤即可完成第二段秘密信息的提取. 最后，本文算法所隐藏的全部秘密信息可通过对提取的两段秘密信息进行级联得到.

5　实验结果及分析

本节将从生成宋词的质量、秘密信息的隐藏容量、时间复杂度、篡改检测分析及安全性五个方面进行实验结果分析和比较.

5.1　实验数据集及参数设置

本文实验中共使用宋词 21053 首，包括训练数据 19905 首，测试数据 661 首，验证数据 487 首. 在预训练阶段，只使用字符来构建词汇，且字符数为 27681，训练时使用 Adam 算法进行优化. 本文算法模型在预训练阶段和微调阶段的参数设置都是确定的. 在训练过程中，模型的具体参数设置如下：网络层数为 12，隐藏网络层数为 768，遮蔽的自注意力机制的级数为 12，全局的自注意力机制的级数为 12；训练时每次输入的 Batch 为 8；为防止模型出现过拟合的问题，Dropout 参数设置为 0.2. 整个模型搭建使用的是 Pytorch 框架，训练过程中使用 NVIDIA GeForce RTX 2080 Ti GPU 进行计算.

5.2　生成宋词质量分析

SongNet 模型生成的宋词在句子完整性、格式以及韵律方面的性能比文献[18, 20]更为理想. 因此，本文算法模型生成宋词的质量主要与 SongNet 模型生成宋词的质量进行比较分析. 在自然语言处理领域(Natural Language Processing, NLP)中，Perplexity (PPL)是衡量语言模型优劣的重要指标. PPL 计算的基本思想是：在一个语言模型生成文本时，会预测下一个字的生成有多少种合理的选择，当未来可选字的数量越少，我们就认为模型越准确，其计算公式如下：

$$\mathrm{PPL}(S_{\mathrm{t}}) = P(w_1 w_2 ... w_M)^{-\frac{1}{M}} \tag{15}$$

其中 S_{t} 代表句子，M 为句子长度，$p(w_j)$表示第 j 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. 为保证实验比较的公平性，我们使用相同的参数设置和训练数据库分别对 SongNet 模型和本文的宋词生成模型进行训练. 通过计算得到：SongNet 模型在训练时的 PPL 值为 12.64；本文模型在训练时的 PPL 值为 12.92. 从理论上来讲，本文模型与 SongNet 模型生成的宋词质量基本处于同一水平. 此时本文模型是隐藏有秘密信息的，而 SongNet 模型没有隐藏信息的功能.

图 8 给出了本文模型与 SongNet 模型生成的宋词文本的一些实例，展示的宋词的 6 种词牌格式模板分别是：如梦令、清平乐、点绛唇、浣溪沙、卜算子、少年游. 通过对生成宋词的阅读和格律分析，本文模型所生成的宋词符合词牌格式模板的要求，与 SongNet 模型生成宋词的质量基本相当. 由于本文算法设计了韵律编码，故生成宋词的韵律更加明确，而 SongNet 模型生成的宋词并没有全部符合韵律要求，如《如梦令》中末句的‘院’与‘婴’,《清平乐》中的‘日’与‘平’,《点绛唇》中的‘照’与‘烛’等. 另外，与其它仅考虑词向量的算法[18, 20]相比，本文算法生成的宋词在韵律方面更加自然，在句子完整性方面表现更好.

5.3　隐藏容量分析

针对隐藏容量性能，我们将本文算法与几种典型的基于文本生成的信息隐藏算法[13-15, 18, 20, 21]进行了比较. 文献[18]是较早提出使用宋词为载体进行信息隐藏的算法之一，该算法利用词牌模板，根据平仄音调划分词语块组. 文献[18, 20]都是基于宋词格律模板的词语分块进行秘密信息隐藏，且隐藏容量的大小随词典数量的增加而提高. 文献[21]是利用宋词曲调中的“中平仄”出现频率的奇偶性进行信息隐藏. 文献[13]以英文为载体，基于马尔可夫链模型和霍夫曼编码实现信息隐藏，生成文本语句质量较一般. 文献[14-15]均为使用绝句为载体的信息隐藏算法，文献[14]以单个字为单位进行信息隐藏，因候选字库较大故在隐藏容量方面表现突出，但生

本文算法	SongNet[24]	本文算法	SongNet[24]
词牌：如梦令 水国芙蓉少女， 三径独行争夺. 认杨花应是， 乱点嫩黄金剩. 催晚，定燕， 闲趁东风到达.	词牌：如梦令 小逐马蹄骄骢， 不乱莺啼庭院. 青人卖花声， 去买些材料担. 施纤，施纤， 尚可左手为婴.	词牌：浣溪沙 薄晚林枫子贤路， 柳绵逐马东疏雨， 江南水陆离潮营. 卖花声过独立阑， 魂断人归霞曾降， 夜寒犹暖洞府蛮.	词牌：浣溪沙 冰雪初拆碎琉璃， 玉碾朱扉暗锁月， 沈磨斗薰风殿阁. 沈沈夜色迷人处， 斜月暗风来到地， 高桐阴下白头啼.
词牌：清平乐 红豆元宵， 小车载柳橡. 东风犹带柳忆惜， 还共故人车辆. 踏青斗草衣闹， 谁道轻狂太爽. 还是那番令月， 小窗依旧多潸.	词牌：清平乐 梅黄柳紫， 绿暗门庭日. 午槐初浴东风软， 正嘉节重阳好. 笙酣度曲初醉， 清跸百万人共. 雍容四倚阑干， 太平渐看升平.	词牌：卜算子 凉夜向犹小， 拂拂长虹憭. 渺渺烟波南穷远， 烟树晚来缥. 谁共丽照明， 花影斜阳渺. 斜倒盏面红暗倒， 花影倒回赵.	词牌：卜算子 风荷叶满丛， 又见新荷叶. 长见绿庭与叶艳， 淡伫映娉婷. 暮霭迷空里， 潇洒避春丛. 人去绿阴深处游， 尽日栏干倚
词牌：点绛唇 一番愁绝， 一番雨过西郊幻. 正西情少， 几点红楼铲. 暮山寄信， 问塞管吹裯. 暮才去， 明朝忽到， 又是在西洵.	词牌：点绛唇 嫩凉新霁， 月华如水初相照. 正西风破， 香点破茶汐. 翠帷深处， 烛影动金莲. 盛宴中， 夜凉声里， 烛影动银烛.	词牌：少年游 林叶青青归客舍， 风蝉摇金蜗. 蒲柳使君， 天与黄品， 西洛与他偕. 小莲王国王吟咏， 小风池边乖. 绿窗柳色， 曲堤邻里， 灯火烂银.	词牌：少年游 秋老庭牲夕阳红， 风生杨柳外. 莺啼日暮， 池亭深掩， 门艾萧骚少. 驱车入江南太平， 好是恁欢娱. 恨少年少， 恼人情重， 惟有梦魂.

图 8 生成的宋词质量对比

成的绝句文本也因候选字过多而质量一般. 文献[15]本质上也是以单个字为基础进行信息隐藏，故也需控制候选字的数量，以保证生成绝句的文本质量. 本文算法是以词牌格式为基础进行信息隐藏，且因不同词牌所对应的格式模板不同，使得可隐藏的秘密信息量也不同. 本文算法本质上是以词牌格式中每个句子为单位进行信息隐藏.

在实验中，候选字共有 5310 个，因此生成诗句的字理论上有 5130^{l} 种可能（l 为宋词的长度），这表明生成诗歌的字有相对较大的选择范围. 为保证生成宋词的质量，本文算法采用 Top-k sampling[25] 来进行候选字的筛选. Top-k sampling 的核心是对类型相同的数据组按照某种标准进行比较排序，找出其中前 k 个最优的数据元素. 通过 Top-k sampling 对候选字进行动态选择，可以显著增加结果的多样性. 由公式(13)和(14)可知，本文算法的隐藏容量主要与 T_{num}，$S_{\text{c-num}}$，$S_{\text{f-num}}$，$P_{\text{p-num}}$，$P_{\text{z-num}}$，$R_{\text{p-num}}$，$R_{\text{z-num}}$，K_{num}，W_{num} 有关. 存在有平仄韵词牌的选择，因此对应平、仄韵词牌的格式模板和韵律等参数并不会相互干扰. 最终得出生成的含密宋词中平均每个句子的信息隐藏容量根据词牌选择有两种情况，即 C_{p} 和 C_{z}. 对应平韵词牌的句子信息隐藏容量 C_{p} 为

$$C_{\text{p}}=\frac{L_{\text{p}}}{S_{\text{f-num}}}=\frac{\log_2\left(T_{\text{num}}\cdot P_{\text{p-num}}\cdot R_{\text{p-num}}\right)}{S_{\text{f-num}}}+\frac{S_{\text{c-num}}\cdot\log_2\left(K_{\text{num}}\right)+S_{\text{f-num}}\cdot\log_2\left(W_{\text{num}}\right)}{S_{\text{f-num}}} \tag{16}$$

对应仄韵词牌的句子信息隐藏容量 C_{z} 为

$$C_{\text{z}}=\frac{L_{\text{z}}}{S_{\text{f-num}}}=\frac{\log_2\left(T_{\text{num}}\cdot P_{\text{z-num}}\cdot R_{\text{z-num}}\right)}{S_{\text{f-num}}}+\frac{S_{\text{c-num}}\cdot\log_2\left(K_{\text{num}}\right)+S_{\text{f-num}}\cdot\log_2\left(W_{\text{num}}\right)}{S_{\text{f-num}}} \tag{17}$$

实验训练数据候选字共有 5,310 个，这里选取 4,096 个用于宋词生成并隐藏信息，即 K_{num} = 4,096. 表 1 给出了不同算法的隐藏容量的比较结果.

表 1 信息隐藏容量比较结果

算法	隐藏容量
文献[13]	3 bits/单词
文献[14]	16.98 bits/句
文献[15]	18.23 bits/句
文献[18]	1.32 bits/字节
文献[20]	1.6-2.3 bits/字节
文献[21]	3 bits/词牌
本文算法	18-21 bits/句

从表 1 可以看出，本文算法表现最佳；虽然文献[14]的算法在隐藏容量上也有较好的表现，但生成的诗句不符合严格的韵律要求；文献[18, 20, 21]都是利用宋词文本作为载体，但文献[18]和[20]生成的宋词质量不够理想，文献[21]的隐藏容量较低；文献[15]的隐藏容量与本文算法最为接近，但二者用于隐藏的载体不同，且文献[15]在时间复杂度上要高于本文算法.

5.4　时间复杂度分析

在本文实验中，信息隐藏和提取所使用的全部字典码本都是预生成的，因此本文算法的复杂度分析主要针对信息隐藏和提取过程中引入的计算操作进行统计. 本文算法的信息隐藏过程包含的主要计算操作有：针对平仄韵词牌 Tune 和词牌格式模板 Pattern 进行选择，Tune 和 Pattern 的数值确定，故对应的时间复杂度为 $O(1)$；接下来根据词牌格式模板 Pattern 来隐藏关键词 Keyword，韵律 Rhyme 和押韵字符 Rhyme Word，在这个过程中时间复杂度主要集中在关键字、韵律和押韵字符的选择，Keyword，Rhyme 和 Rhyme Word 的数值确定，时间复杂度为 $O(N)$，N 为格式模板长度，故对应的时间复杂度为 $O(n)$；最后依据参数对应的固定格式模板来生成宋词，时间复杂度为 $O(N\cdot size)$，其中 $size$ 为候选字的个数，因此信息隐藏过程的时间复杂度为 $O(n^2)$. 信息提取过程仅需根据平仄韵词牌 Tune 字典，词牌格式模板 Pattern 字典、关键词 Keyword 字典，韵律 Rhyme 字典和押韵字符 Rhyme word 字典来进行索引，便可提取出秘密信息，故对应的时间复杂度为 $O(n)$.

表 2　时间复杂度比较结果

算法	信息隐藏	信息提取
文献[13]	$O(n^2)$	$O(n^2)$
文献[14]	$O(n^2)$	$O(n^4)$
文献[15]	$O(n^2)$	$O(n^2)$
文献[18]	$O(1)$	$O(n)$
文献[20]	$O(1)$	$O(n)$
文献[21]	$O(n)$	$O(n)$
本文算法	$O(n^2)$	$O(n)$

本文算法与其它文献的时间复杂度比较结果如表 2 所示. 文献[13, 14, 15, 18, 20, 21]中，只有[15]的隐藏容量和本文算法相近，但在时间复杂度方面，本文算法的信息提取效率要明显优于文献[15]. 由于本文算法追求更加高质量的生成文本，增加了模型的预训练和抛光，故时间复杂度上并没有达到最优. 文献[15]在提取信息时，需要将从诗歌中提取到的第一部分信息（如关键字、格式模板及韵律等）输入到模型中，在生成可选择诗句后才能提取出第二部分信息，故时间复杂度为 $O(N\cdot size)$，即 $O(n^2)$，远高于本文算法提取过程的复杂度 $O(n)$.

5.5　篡改检测分析

目前，对于文本的篡改形式有两种：一是针对文本格式的篡改，二是针对文本内容的篡改. 算法可以针对以下三种篡改情况进行校验：(1)对词牌名和词牌格式模板的篡改. 在本文算法当中，词牌名与词牌格式模板有对应关系，如果词牌名和模板遭受到篡改，将无法正确提取信息，即可以校验出文本被篡改. 如原文：“《满江红》”，篡改后：“《上江虹》”，即篡改后破坏了词牌名与模板的对应关系，算法可以进行检测分析. (2)押韵字符的篡改. 设计算法时，将韵律和押韵字符进行对应关系编码，故可通过押韵字符是否属于同一韵律来进行篡改检测. 以平声韵词牌“长相思”为例，有两组平声韵字符，1 组字符“桐、东、丛、笼”归属一类韵律，2 组字符“松、冬、从、茏”归属另一类韵律. 原文：“《长相思》___，__桐. _____，____东. ___，__丛. _____，____笼.”，篡改后：“《长相思》___，__桐. _____，____冬. ___，__丛. _____，____茏.”，即篡改后的押韵字符虽仍同为平声韵字符，但归属不同韵律类别，故可以校验出押韵字符被篡改过. (3)对除格律模块之外的内容篡改. 算法利用格律模块进行信息隐藏，若被篡改之处位于格律模块之外，则不影响正确信息的提取. 正确信息提取后，可以重构格律模块，在相同的模型和数据集的情况下，同一个格律模块输入可以生成的相同宋词诗句，通过比对即可进行篡改检测. 以词牌《卜算子》为例，格律模块：“___雨_，____笑. _____春_，____到. ___风_，____报. _____花_，____俏.”，原文：“凉月破雨碧，又见一番笑. 鸦与鹤寿俱春少，欢喜为春到. 人少多风流，谁使华鬓报. 十二玉楼无花样，小字金钩俏.”，篡改后：“凉风破雨碧，又是一番笑. 莺与鹤寿俱春少，欢乐为春到. 人少几风流，谁让华鬓报. 十二玉楼有花样，小字金勾俏.”. 篡改后没有破坏格律模块，即可正确提取信息，并重构格律模块：“___雨_，____笑. _____春_，____到. ___风_，____报. _____花_，____俏.”，继而生成上述原文，通过二者比对可检测到被篡改的内容，即为下划线部分.

5.6 安全性分析

本文算法中秘密信息的隐藏是基于平仄韵词牌、词牌格式模板、关键字、韵律及押韵字符的选择信息进行编码的. 表 3 给出了词牌格式模板、关键字和韵律数据集的构成及其排列组合的可能数. 另外由于每种韵律对应的押韵字符各不相同且数量较多，故押韵字符数据的构成没有在表中列出.

表 3 候选库构成

候选库	容量	可能数
词牌模板	768	$C_{1740}^{768}\times 768!$
关键字	4096	$C_{5310}^{4096}\times 4096!$
韵律-平韵	16	$C_{30}^{16}\times 16!$
韵律-仄韵	64	$C_{70}^{64}\times 64!$

攻击者若想提取秘密信息，必须做到以下几点：(1)分析提取生成宋词的关键字，包含关键字的隐藏位置和数量，而每首宋词中关键字的数量根据格式模板的不同而不同；(2)分析关键字所对应的二进制码，候选字有 5310，本文算法从中选取关键字数量为 4096，即有 C_{5310}^{4096} 种可能，且经排列组合后又有 4096！种可能；(3)分析词牌格式模板所对应的二进制码，词牌有 1740 种，从中选取 768 种，即有 C_{1740}^{768}种可能，且选取后的 768 种模板排列组合后又有 768！种可能；(4)分析宋词的韵律信息，包含韵律的种类以及对应的二进制码. 平声韵有 30 种，从中选取 16 种，即 C_{30}^{16} = 3.04e^21 种可能，且经排列组合后又有 16！种可能，仄声韵有 70 种，从中选取 64 种，即 C_{70}^{64} = 1.66e^97 种可能，且经排列组合后又有 64！种可能；(5)分析每种押韵字符所对应的二进制码，且押韵字符是在韵律选择的基础上进行的，需要分析出韵律选择后才能进行，押韵字符的编码也可进行排列组合，可能数设为 D；(6)分析提取平仄韵词牌、词牌格式模板、关键字、韵律及押韵字符在秘密信息隐藏过程中的顺序，经排列组合后有 5！种可能. 所以，本文算法信息隐藏过程对应的所有可能情况数是为上述六个方面可能情况数的乘积，即 $C_{5310}^{4096}\times 4096!\times C_{1740}^{768}\times 768!\times C_{30}^{16}\times 16!\times 5!\times D$（平韵情况）或者 $C_{5310}^{4096}\times 4096!\times C_{1740}^{768}\times 768!\times C_{70}^{64}\times 64!\times 5!\times D$（仄韵情况）. 基于以上分析，攻击者想通过暴力破解获取本文算法秘密信息的难度极高.

另外，我们还对本文算法在抗语义相似性检测方面进行了实验分析. 对不同信息隐藏容量下本文算法的含密载体文本和非含密载体文本进行区分实验，并用可视化的方法展示了统计上区分的难易程度，如图 9 所示.

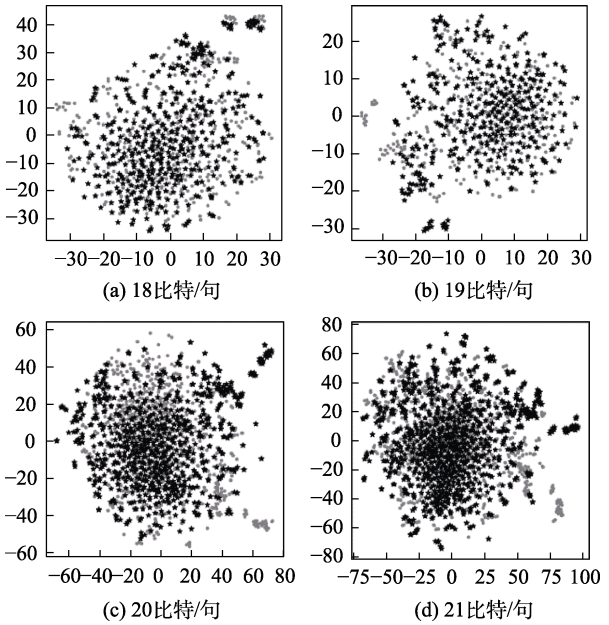

图 9 含密文本（深色星点）和不含密文本（浅色圆点）在语义空间中的分布差异

实验运用了文献[26]中的方法分别对含密文载体文本和非含密载体文本进行语义空间映射，然后使用 t-SNE 算法来可视化结果，其中浅色圆点和深色星点分别对应非含密文本和含密文本. 由图 9 可以看出，含密文本和非含密文本的语义空间非常相近；且随着隐藏容量的增加，含密文本的分布更加集中，与不含密文本的分布虽存在些许偏差，但总体仍保持在同一区域内. 该结果表明了本文算法生成的隐藏有秘密信息的宋词文本与未隐藏信息的宋词文本在语义空间上几乎是无法区分的.

6 总 结

本文提出一种基于宋词生成的构造式信息隐藏算法，在保证生成宋词质量的同时，实现了高容量的秘密信息隐藏，且算法的安全性较高. 通过加入符号集的设计来确保生成宋词格式、韵律和句子完整性. 在秘密信息隐藏过程中，通过选择平仄韵词牌、词牌格式模板、关键字、韵律以及押韵字符来构建格律模块，输入模型生成宋词的同时进行秘密信息的隐藏. 在信息提取过程中，关键字、韵律等信息的编码设计有效保证了信息提取的正确性. 在未来的工作中，将研究生成多风格的含密文本载体，如对联、歌词、小说等，以增加用于秘密信息隐藏的文本种类；同时通过增加关键字和押韵字符的数量，以及考虑文本的情感主题分类等，来进一步提

高信息隐藏容量和生成含密文本的质量.

致　谢　在此，我们向对论文提出宝贵意见的审稿专家们表示衷心的感谢!

参　考　文　献

[1] Chen W B, Yang G B, Chen R C, et al. Digital video passive forensics for its authenticity and source. 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. 通信学报, 2011, 32(6): 177-183)

[2] Zhang X P, Qian Z X, Li S. Prospects of information hiding research. Journal of Applied Sciences, 2016, 34(05): 475-489 (in Chinese)
(张新鹏,钱振兴,李晟.信息隐藏研究展望. 应用科学学报, 2016, 34(05): 475-489)

[3] Wu Y Q, Guo Y T, Tang J, Luo B, Yin Z X. Reversible data hiding in encrypted images using adaptive huffman encoding strategy. Chinese Journal of Computers, 2021,44(04):846-858(in Chinese)
(吴友情, 郭玉堂, 汤进, 罗斌, 殷赵霞. 基于自适应哈夫曼编码的密文可逆信息隐藏算法.计算机学报, 2021, 44(04): 846-858)

[4] Wu G H, Gong L C, Yuan L F, Yao Y. Review of information hiding on Chinese text. Journal on Communications, 2019, 40(09): 145-156(in Chinese)
(吴国华, 龚礼春, 袁理锋, 姚晔. 中文文本信息隐藏研究进展. 通信学报, 2019, 40(09): 145-156)

[5] Tan Y. Comparative study of information hiding method based on text typesetting format. Computer and Modernization, 2013(06): 52-56 (in Chinese)
(谭瑛.基于文本排版格式的信息隐藏方法比较研究.计算机与现代化, 2013(06): 52-56)

[6] IS Lee, W H Tsai. 2008. Secret communication through web pages using special space codes in HTML files. International Journal of Applied Science and Engineering 6, 2 (2008), 141-149

[7] Lee C, Chen H. Data concealment scheme for HTML documents based on color code//Proceedings of the Fifth International Conference on Intelligent Information Hiding & Multimedia Signal Processing. Kyoto, Japan, 2009: 632-635

[8] R Kumar, A Malik, S Singh, B Kumar, S Chand. A space based reversible high capacity text steganography scheme using font type and style// 2016 International Conference on Computing, Communication and Automation (ICCCA), Greater Noida, India, 2016, 1090-1094

[9] Chen Y N, Li Q M, Lv C X, Sang X N, Dong X. Research on text security hiding algorithm for invisible characters. Cyberspace Security, 2019, 10(05): 88-96(in Chinese)
(陈旖旎, 李千目, 吕超贤, 桑笑楠, 董潇. 不可见字符的文本安全隐藏算法研究. 网络空间安全, 2019, 10(05):88-96)

[10] Su B, Ding X Y, Liu G S, et al. An information hiding method for text by substituted conception//Proceedings of the Fourth International Symposium on Information Science and Engineering. Shanghai, China, 2012: 131-135

[11] Xiang L Y, Li Y, Hao W. Reversible natural language watermarking using synonym substitution and arithmetic coding. Computers Materials & Continua, 2018, 55(3): 541-559

[12] Xiang L, Y u J M, Y ang C F, et al. A word-embedding-based steganalysis method for linguistic steganography via synonym substitution. IEEE Access, 2018, 6: 64131-64141

[13] Yang Z L, Jin S Y, Huang Y F, Zhang Y J, Li H. 2018. Automatically generate steganographic text based on Markov model and Huffman coding. 2018-11-12

[14] Luo Y, Huang Y. Text steganography with high embedding rate: using recurrent neural networks to generate Chinese classic poetry//Proceedings of the Seventh ACM Workshop on Information Hiding & Multimedia Security. New York, USA, 2017: 99-104

[15] Qin C, Wang M, Si G W, Yao H. Constructive information hiding with chinese quatrain generation. Chinese Journal of Computers, 2021, 44(04): 773-785(in Chinese)
(秦川, 王萌, 司广文, 姚恒. 基于绝句生成的构造式信息隐藏算法. 计算机学报, 2021, 44(04): 773-785)

[16] Shen J M, Ji H and Han J W. 2020. Near-imperceptible neural linguistic steganography via self-adjusting arithmetic coding. 2020-10-01

[17] Yang Z L, Zhang S Y, Hu Y T, Hu Z W, Huang Y F. 2020.VAE-Stega: Linguistic steganography based on variational auto-encoder. IEEE Transactions on Information Forensics and Security, 16: 880-895

[18] Yu Z S, Huang L S, Chen Z L, et al. High embedding ratio text steganography by ci-poetry of the Song dynasty. 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. 中文信息学报, 2009, 23(4): 55-62)

[19] Liu Y C, Wang J, W ang Z B, et al. A technique of high embedding rate text steganography based on whole poetry of song dynasty//Proceedings of the Second International Conference on Cloud Computing and Security. Nanjing, China, 2016: 178-189

[20] Liu Y C, Wang J, Qu Q F. Text information hiding technique by carrier of song poetry based on hybrid encryption. Computer Technology and Development, 2018, 28(01): 138-143(in Chinese)
(刘彦辰, 王箭, 屈琪锋. 混合加密的宋词载体文本信息隐藏技术. 计算机技术与发展, 2018, 28(01): 138-143)

[21] Zou Z Y, Wang K X. Method of text steganography without carrier based on rhythm in chinese song poems. Journal of Qingdao University (Natural Science Edition), 2021, 34(01): 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. 青岛大学学报(自然科学版), 2021, 34(01): 7-12)

[22] Devlin J, Chang M W, Lee K, et al. BERT：pretraining of deep bidirectional transformers for language understanding. arXiv: 1810.04805,2018

[23] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need. //Proceedings of the 31st Conference on Neural Information Processing Systems, Long Beach, USA, 2017:6000-6010.

[24] Li P, Zhang H, Liu X, et al. Rigid formats controlled text generation. //Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. Washington , USA, 2020: 742-751

[25] Angela F, Mike L, Yann D. Hierarchical neural story generation. arXiv:1805.04833,2018

[26] Le Q. V., Mikolov T. Distributed representations of sentences and documents. JMLR Workshop and Conference//Proceedings of the 31th International Conference on Machine Learning (ICML). Beijing, China: ACM, 2014, 32: 1188-1196

QIN Chuan, Ph.D., professor, Ph.D. supervisor. His research interests include multimedia security and AI security.

LI Rong-Shou, M.S.. Her research interest is text information hiding.

QIAN Zhen-Xing, Ph.D., professor, Ph.D. supervisor. His research interests include information hiding and AI security.

ZHANG Xin-Peng, Ph.D., professor, Ph.D. supervisor. His research interests include multimedia security and AI security.

Background

The research in this work belongs to the field of information hiding. In the Internet age, information security has attracted more and more attention. With the continuous improvement of computer capabilities, the possibility of traditional information hiding algorithms being detected and cracked continues to increase. Therefore, there occurs the coverless information hiding technique, which can also be called as the constructive information hiding. The characteristic of constructive information hiding is that there is no modification on cover data, which can reduce the possibility of being discovered by the attacker, and the risk of being cracked is also greatly reduced. So far, there are two main types of coverless information hiding algorithms: one is based on the cover selection strategy, and the other is based on the cover generation strategy. At present, how to improve the hiding capacity while ensuring the quality of the generated text is the main challenge in the design of information hiding algorithm based on text generation. This work focuses on constructive information hiding with Song Ci generation. In the process of Song Ci generation, our algorithm uses pre-training and fine-tuning, and pays more attention to the template and rhythm information of format. Therefore, the Song Ci generated by our algorithm has clear rhythm and higher sentence integrity. In addition, information hiding capacity of our algorithm is greater than that of some typical reported algorithms.

This work is supported in part by the National Natural Science Foundation of China (NSFC) project “Research on Watermarking and Hashing Algorithms towards Copyright Protection and Tampering Authentication for Neural Network Models” under Grant No. 62172280, which aims at studying theory and methodology on digital watermarking and hashing for neural network models and realizing copyright protection and tampering authentication for neural network models, in part by the NSFC project “On Detection and Recognition of Fake Media Content in Online Social Networks” under Grant No. U20B2051, which focuses on developing theories and methods for detecting and recognizing fake media contents (FMC) in online social networks and building a complete scheme of FMC forensics, FMC recognition and FMC blocking, and in part by the Natural Science Foundation of Shanghai project “Reversible Data Hiding for JPEG Images with RAW Reconstruction Capability” under Grant No. 21ZR1444600, which focuses on establishing the nonlinear relationship between RAW data and RGB data in digital images, simulating the approximate reconstruction process of RAW image with few data and developing reversible data hiding methods for JPEG images with RAW reconstruction capability.