文章编号:1007-5321(2018)06-0007-07 **DOI**:10.13190/j.jbupt.2018-032

基于图像描述的文本信息隐藏

薛一鸣[1], 周雪婧[1], 周小诗[1], 牛少彰[2], 文 娟[1]
(1. 中国农业大学 信息与电气工程学院, 北京 100083; 2. 北京邮电大学 计算机学院, 北京 100876)

摘要: 针对文本信息隐藏嵌入容量低和语义连贯性差的问题,提出了一种基于神经网络图像描述的文本信息隐藏模型. 将卷积神经网络与长短期记忆网络相结合,把图像特征和生成语句进行关联. 从收发双方能否共享图像及模型参数的不同应用前提出发,设计了多种概率采样方式,从而生成载密的图像描述文本. 实验结果表明,该算法具有较高的隐藏容量,载密描述句能较好地表达图像内容. 该模型归属于"无载体"自然语言生成式信息隐藏,具有较好的隐蔽性和安全性.

关 键 词: 文本信息隐藏; 图像描述; 卷积神经网络; 长短期记忆网络
中图分类号: TP309.2 **文献标志码**: A

Text Steganography Based on Image Caption

XUE Yi-ming[1], ZHOU Xue-jing[1], ZHOU Xiao-shi[1], NIU Shao-zhang[2], WEN Juan[1]
(1. College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China;
2. School of Computer Science, Beijing University of Posts and Telecommunication, Beijing 100876, China)

Abstract: Aiming at the problem of low embedding capacity and poor semantic coherence of text steganography, a text steganographic scheme based on neural image caption is proposed. An encode-decode structure with a combination of long short term memory and convolution neural network is used to model the joint probability distributions between image features and the descriptive sentences. Two methods with different sampling process are designed from the perspectives of sharing and non-sharing models. Experimental results show that the proposed model can achieve high embedding capacity and desirable text quality. This scheme belongs to "carrier-free" steganography and has good security.

Key words: text steganography; image caption; convolutional neural network; long short term memory

信息隐藏技术以隐藏机密信息的存在性为根本目的,利用载体的冗余,将秘密信息隐藏于视频、语音、图像或文本等载体中,以躲避人类视觉和隐写分析算法的检测,从而保证机密数据的安全.

文本是使用最为广泛的一类信息载体,它形式多样、编码简单、存储方便、传输快捷. 最初的文本隐藏是基于格式的,通过修改文本的字间距、格式文档的某些属性等来嵌入信息[1]. 这类方法具有较大的隐藏容量,但不能抵抗基于统计特性的隐写分析及重排版. 随着自然语言处理技术的发展,自然语言信息隐藏成为信息隐藏领域的一大研究热点. 从隐藏机制上看,自然语言信息隐藏可以分为嵌入隐藏法和生成隐藏法. 嵌入隐藏法通过对原始文本在句法或语义层的等价修改来嵌入信息,如同义词替换[2-3]、引入拼写错误[4]、句法变换[5]、语句复述[6]、机器翻译[7]等. 这类方法存在原始载体,攻击者可以通过对比分析来发现修改的位置,因此安全性不高. 生成隐藏法利用上下文无关文法等自然语言生

收稿日期: 2018-02-02
基金项目: 国家自然科学基金项目(61802410, 61872368)
作者简介: 薛一鸣(1968—), 男, 副教授; 文 娟(1982—), 女, 讲师, E-mail:wenjuan@cau.edu.cn.

成技术自动生成载密文本数据,如 Spammimic[8] 和 Nicetext[9]. 这类方法无需原始载体文本,无法通过比对分析检测,但难点在于如何解决上下文语义连贯性问题. 因此,有专家采用对语义连贯要求不太高的体裁如诗歌进行载密文本生成,取得了较好的效果[10-11].

随着深度学习技术的发展,自然语言生成技术有了极大提高. 在图像理解和描述领域,也有了许多新的进展. 基于神经网络的图像描述,借鉴了神经网络机器翻译中"编码-解码"的思想[12],将图像看作源语言,将图像描述当作目标语言模拟机器翻译的过程,生成性能接近于人工标注的描述语句.

受到图像描述方法的启发,笔者对于给定图像,基于图像描述框架,用卷积神经网络(CNN, convolutional neural network)提取图像特征向量[13],并结合长短时记忆网络(LSTM, long short term memory)[14],在充分理解图像内容的基础上,生成符合自然语言统计规律的载密描述语句. 这种隐藏方法无需修改图像本身,也没有可比对的原始文本,可视为一种"无载体"隐藏方法. 该方法可以为自然语言信息隐藏领域提供一种新的解决思路.

1 基于神经网络的图像描述生成

图像描述即根据图像内容生成描述性文字. 图像描述算法不仅要检测图像中的物体,理解物体的相互关系,还要用自然语言表达出来. 基于神经网络的图像描述融合了计算机视觉、深度学习和自然语言处理技术. Google 最先借用机器翻译算法中序列到序列的"编码-解码"结构,构造了一个图像描述生成模型——NIC(neural image caption)[15]. NIC 整体框架如图 1 所示.

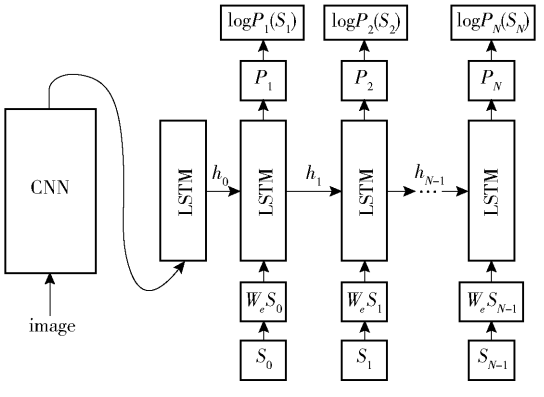

图 1 NIC 模型框架

NIC 生成图像描述的过程:首先,将一副图像用 CNN 经多个卷积层、池化层和激活层,得到一个固定维数的特征向量;然后,将该向量作为 LSTM 网络 t_{-1} 时刻的输入,并将 t_{-1} 时刻 LSTM 的隐藏层状态 $\boldsymbol{h}_0$ 输入至 t_0 时刻的 LSTM;对于 $t_{i+1}(i\geqslant 0)$ 时刻的 LSTM 网络,其输入为上一时刻 LSTM 的隐藏层状态 $\boldsymbol{h}_i$ 及预测的单词输出 $\boldsymbol{S}_i$,所有的 LSTM 共享参数. 训练时,所有的单词都通过 one - hot 词向量表示. 模型可计算出每一时刻所有单词的概率分布,取概率最大的单词作为输出并输入至下一层级的网络中,如此循环,直到生成的单词为结束符或句子长度超过一定的阈值. 数学描述如下:

$$\boldsymbol{x}_{-1}=\mathrm{CNN}(\boldsymbol{I}) \tag{1}$$

$$\boldsymbol{x}_t=\boldsymbol{W}_e\boldsymbol{S}_t,\ t\in\{0,1,\cdots,N-1\} \tag{2}$$

$$\boldsymbol{P}_{t+1}=\mathrm{LSTM}(\boldsymbol{x}_t),\ t\in\{0,1,\cdots,N-1\} \tag{3}$$

其中:$\boldsymbol{I}$ 表示图像,$\boldsymbol{W}_e$ 为词典的词向量矩阵,$\boldsymbol{W}_e\boldsymbol{S}_t$ 为抽取 t 时刻输入词的词向量. 在 NIC 框架基础上,Kelvin 等[16]在网络中加入注意力机制,自动从输入序列中选取对应特征,提高了模型性能; Wu 等[17]尝试用多标签分类方法将高层语义特征加入模型; Andrei 等[18]提出用视觉语义对齐模型以及 Multimodal RNN 模型将图像与语句片段关联起来,生成图像各个不同区域的文本描述.

2 基于图像描述的文本信息隐藏算法

在图像描述任务的测试阶段,将前一时刻预测输出词与状态传入 LSTM 结构后,通常有 2 种搜索算法来获取当前时刻的图像描述词. 第 1 种为直接采样,即每次选取概率最大的单词作为当前时刻的输出与下一时刻的输入. 第 2 种方法为集束搜索(beam search),假设集束参数 beam size 大小为 b,在 t 时刻,模型会选择 LSTM 输出中前 b 个概率最大的单词,将其全部作为下一时刻的输入,依次进行扩展. 笔者所提的自然语言隐藏算法是在改动 beam search 搜索模式的基础上提出的. 从接收方是否能共享神经网络模型和测试图像的角度出发,设计了不同的隐藏方案. 在收发双方共享模型、参数和图像的情况下,设计了基于句子的隐藏算法(SSH, sentence by sentence hiding)和基于单词的隐藏算法(WWH, word by word hiding). 针对接收方未共享神经网络模型的情况,设计了一种基于散列函数的信息隐藏模型(HH, Hash hiding). 以下将详细介绍这几种信息隐藏模型.

2.1　收发方共享模型参数的信息隐藏算法

当接收方能共享神经网络模型参数并已知载体图像集时，可以对照生成的文本，按照生成规则还原出文本的生成路径，从而提取秘密信息. 值得注意的是，这种情况虽然需要已知载体图像，但由于无需在载体图像上做任何修改，所以实际应用时可以分享图片链接或图片相关位置，无须发送图像集，从而节省传输带宽.

2.1.1　基于句子的嵌入算法 SSH

基于句子的嵌入算法主要思想是针对每一幅图像，根据最后一个时刻生成的 beam size 个生成句，按照概率大小进行排序和编码，然后通过匹配秘密信息比特来选择对应编码的句子作为最终输出. 例如，设置 beam size = 4，则运行完最后一层 LSTM，会生成 4 个待选句子. 利用等长编码，依次编码成 00，01，10 和 11. 选取当前待嵌入的 2 bit 密文，选择对应编码的句子作为该图像的输出描述句. 为了获取密文比特长度信息，嵌入之前，需在密文比特前加入 16 bit 数据表示密文长度. SSH 算法描述如下：

算法 1　SSH 隐藏算法

输入：图像集 CNN 特征矩阵；beam size = 2^n；密文二进制比特流.

输出：图像载密描述语句集.

1）获取密文比特流，在前端加入 16 bit 表示密文长度的头数据. 将整合后的密文数据进行分组，每组 n bit，若不能被 n 整除，则在末尾补零.

2）从图像特征矩阵中提取当前图像特征向量.

3）$t = 0$ 时刻，输入图像特征向量和 START 符，对 LSTM 的输出 $\boldsymbol{P}_1$ 采样，取前 2^n 的词作为 $t = 1$ 的待选词集 $W_1 = \{w_{11}, w_{12}, \cdots, w_{12^n}\}$.

4）在 $t = 1$ 时刻，分别将 $\{w_{11}, w_{12}, \cdots, w_{12^n}\}$ 输入 LSTM，得到概率分布集合 $\{\boldsymbol{P}_{2i}^1, \boldsymbol{P}_{2i}^2, \cdots, \boldsymbol{P}_{2i}^{2^n}\}$，$i \in \{1, 2, \cdots, 2^n\}$，取概率前 2^n 的词作为当前输出与下一时刻的输入. 重复该过程，直到生成的单词为 end 符或超过句子长度的阈值. 最后生成 2^n 个候选句.

5）将 2^n 个候选句按照概率进行等长编码，根据当前分组密文 n bit 内容，选择相应候选句作为当前图像的载密描述句.

6）从图像集 CNN 特征矩阵中提取下一个图像特征向量，重复 3）~5），直至所有密文比特嵌入完毕.

密文提取时，接收方需要事先已知图像特征矩阵、神经网络模型参数和 beam size 等信息. 收到自然语言图像描述集后，根据 beam size 大小，将图像特征矩阵输入并生成图像描述，对生成的 beam size 个候选句进行排序和编码，再与接收到的图像描述句进行比对，获取当前描述句的相应编码，即嵌入的秘密信息. 重复直到完成所有描述句的密文提取. 取前 16 位密文比特计算实际密文长度，截取对应长度的密文比特即可.

2.1.2　基于单词的嵌入算法 WWH

基于单词的嵌入算法主要思想是针对每一幅图像，对每一个时刻的 LSTM，统一固定 beam size = 1 进行采样，因此每一时刻 LSTM 都会有一个单词输出. 若当前密文比特为 1，则选择概率大的那个单词作为输出；若密文比特为 0，则选概率小的作为输出. WWH 算法描述如下：

算法 2　WWH 隐藏算法

输入：图像集 CNN 特征矩阵；密文二进制比特流.

输出：图像载密描述语句集.

1）获取密文比特流，在其前端加入 16 bit 表示密文长度的头数据.

2）从图像特征矩阵中提取当前图像特征向量.

3）$t = 0$ 时刻，输入图像特征向量和 START 符，若当前密文比特为 1，则选择 LSTM 输出词向量中概率最大的词作为当前时刻的输出与下一时刻的输入；若密文比特为 0，则选择概率次大的单词作为当前输出与下一时刻的输入. 重复此过程，直到生成的单词为 end 符或超过句子长度的阈值.

4）将每个时刻生成的词连接起来，生成当前图像的载密描述句.

5）从图像集 CNN 特征矩阵中提取下一个图像特征向量，重复 3）和 4），直至所有密文比特嵌入完毕.

密文提取时，将图像 CNN 特征向量依次输入至 LSTM 网络，在每一时刻的概率采样时，计算当前描述词的概率，通过查询字典的方式，若当前生成的词对应的概率最大，则该时刻隐藏的密文比特为 1，反之为 0，并将该单词作为下一时刻的输入，重复 LSTM 的生成过程，直到所有的描述句提取完毕. 取前 16 位密文比特计算实际密文长度，截取对应长度的密文比特即可.

SSH 和 WWH 的相同点是，两者都需要接收方事先知道发送方生成描述句所用的图像特征矩阵、网络结构和网络参数；不同点是 SSH 是在最后时刻

所有词都已生成完毕之后,对句子进行排序和编码,以选择对应的描述句,而 WWH 是在每一时刻均做一次排序和编码,选择出符合密文比特的词,最后将词连接起来组成图像载密描述句.

2.2 基于散列函数的信息隐藏算法

由于共享参数型隐藏算法需要共享神经网络模型和图像,使用上具有一定的局限性,笔者也设计了一种不需要共享参数的图像描述隐写算法. 使用该算法,接收方根据密钥,就可以直接通过生成的文字提取出密文比特信息.

为了将词与二进制密文相对应,采用一个密钥型 md5 散列函数将词转换成固定长度的数据,然后转换成十进制数据并根据奇偶性给出该词对应的二进制比特位,公式如下:

$$v(w,\text{key}) = (\text{md5}(w+\text{key}))\bmod 2 \quad (4)$$

其中:w 为词,key 为密钥. 函数 $v(w,\text{key})$ 即信息位获取函数,可以通过修改单词 w,使其转换为 0 或者 1. 首先将词和密钥进行字符串连接,然后通过 md5 散列算法转换成十六进制的 md5 数据摘要,转换成十进制并判断其奇偶性,奇数则结果为 1,偶数为 0. 具体算法描述如下:

算法 3 HH 算法

输入:图像集 CNN 特征矩阵;beam size 为 n;密文二进制比特流;散列密钥 key.

输出:图像载密描述语句集.

1) 获取密文比特流,在其前端加入 16 bit 表示密文长度的头数据. 2) 对于训练样本词典中的每一个词 w,用散列映射 $v(w,\text{key})$,得到词的对应信息位.

3) 从图像特征矩阵中提取当前图像特征向量.

4) $t=0$ 时刻,获取当前一个密文比特 m,输入图像特征向量和 START 符,通过 LSTM 网络得到词向量 $\boldsymbol{P}_1$,依据概率大小选出 n 个信息位等于 m 的词作为当前的输出与下一时刻的输入. 重复此过程,直到生成的单词为 end 符或超过句子长度的阈值.

5) 选取概率最大的那个输出结果作为当前图像的载密描述句.

6) 从图像集 CNN 特征矩阵中提取下一个图像特征向量,重复 4)、5),直至所有密文比特嵌入完毕.

提取时,依次将图像描述的每个单词 w 与密钥 key 做 Hash 运算 $v(w,\text{key})$,一个 w 能得到 1 bit 数据,所有的单词得到的比特流中前 16 bit 为密文长度信息,截取相应长度可得到秘密信息.

3 实验结果及分析

3.1 数据准备

采用了来源于雅虎相册网站的 Flickr8k[19] 图像集,一共 8 000 张图片,每张图像对应了 5 句人工描述参考句. 实验中训练集、验证集和测试集的图片数目分别为 6 000、1 000 和 1 000 张.

所提取图像特征向量的 CNN 结构如图 2 所示.

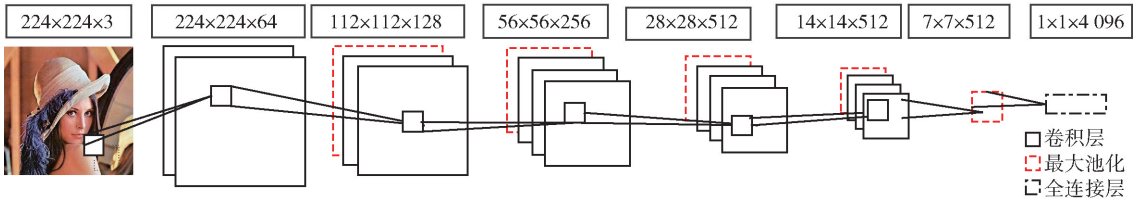

图 2 提取图像特征的 CNN 网络结构

通过 CNN 网络全连接层,每幅图片将生成一个 4 096 维的特征向量,将所有图片的特征向量组合起来,便得到了实验中的图像集 CNN 特征矩阵.

3.2 评价指标

笔者用 BLEU[20] (bilingual evaluation understudy)评价生成文本的质量. BLEU 由 IBM 于 2002 年提出,是目前流行的机器翻译评价指标,用于评价生成译文和参考译文中 n 元组相同的程度. 评价本文算法的性能可以将生成的句子与人工描述的参考句进行对比. 依据 n 元组计算的 BLEU 计算公式如下:

$$\text{BLEU}_n = b(C,S)\exp\left(\sum_{n=1}^{N}\omega_n \lg \text{CP}_n(C,S)\right) \quad (5)$$

$b(C,S)$ 是惩罚因子:

$$b(C,S)=\begin{cases}1, & \text{若 } l_c \geqslant l_s\\ \mathrm{e}^{1-\frac{l_s}{l_c}}, & \text{若 } l_c \leqslant l_s\end{cases} \quad (6)$$

其中:l_c 为待评价句的长度,l_s 为参考句的有效长度(多个参考句时选择与 l_c 最接近的长度).

$\text{CP}_n(C,S)$ 是生成的图像描述与语料库中的参考句的重合精度:

$$\mathrm{CP}_n(C,S)=\frac{\sum_i\sum_k \min(h_k(c_i),\max_{j\in m}h_k(s_{ij}))}{\sum_i\sum_k h_k(c_i)} \tag{7}$$

其中：c_i 为待评价的图像描述，对应的一组参考描述为 $S_i=\{s_{i1},s_{i2},\cdots,s_{im}\}$；$n$ 为句子中的 n 元组，即 n 个单词组成的词组；$h_k(c_i)$ 为第 k 个可能出现的 n 元组在待评价的图像描述 c_i 中出现的次数；$h_k(s_{ij})$ 为第 k 个可能出现的 n 元组在参考句 s_{ij} 中出现的次数.

3.3　实验结果及分析

1）生成的载密文本质量评价.

针对图像描述任务，除了 WWH 固定 beam size 为 1 以外，其他的都可以变动 beam size 的值. 实验所选取的 beam size 为 2^n，$n=(1,2,\cdots,5)$. NIC 为不做隐藏的标准图像描述模型. 采用 BLEU1 ~4 指标评价生成描述文本的质量，结果如表 1 所示.

表 1　标准模型与信息隐藏模型的 BLEU 评价指标

模型	Beam size	BLEU1	BLEU2	BLEU3	BLEU4
NIC	1	55.7	37.3	24.0	15.7
	2	57.2	37.3	25.4	16.8
	4	58.3	39.3	26.0	**17.3**
	8	58.7	39.4	25.7	16.8
	16	59.1	39.8	26.0	16.9
	32	**59.1**	**40.0**	**26.3**	17.0
SSH	2	55.5	37.4	24.4	16.0
	4	56.1	37.6	**24.8**	**16.6**
	8	56.4	37.8	24.7	16.1
	16	**56.6**	**37.9**	24.7	16.1
	32	56.0	37.2	24.1	15.6
WWH	1	49.5	28.8	15.7	8.2
HH	2	46.3	26.8	14.7	8.2
	4	49.5	29.1	16.1	8.8
	8	50.3	29.9	17.5	10.1
	16	**50.9**	**30.8**	**17.7**	**10.2**
	32	50.7	30.2	17.0	9.3

从表 1 中可以看出，几种不同隐藏方案的文本质量大都随着 beam size 的增加有所上升，但当 beam size 增大到 16 以后，文本质量上升的空间会越来越小，对于 SSH 和 HH 算法来说，甚至出现了下降的现象. 这是因为 beam size 如果设置过小可能会错过最优生成句，而 beam size 设置过大，选择范围过大，容易给模型带来噪声，从而导致生成文本质量不再增加甚至下降. 实验中的最优文本出现在 NIC 模型，beam size 为 32 的情况. 这是显而易见的，因为从理论上看，未进行信息隐藏的 NIC 模型的精度应该高于同样 beam size 情况下的其他隐写算法. HH 模型 beam size 为 2 时性能最低，这是因为 HH 需要接收方在没有神经网络模型和图像数据的基础上进行密文的正确提取，根据每一时刻词的信息位进行选择，所以所选词可能会偏离最大概率采样的结果，带来较大的数据误差，但可以通过增加 beam size 的大小来提高 HH 模型的生成句质量.

2）隐藏容量比较

为了考查隐藏容量性能，将本算法与 Nicetext 信息隐藏算法[9]、机器翻译信息隐藏系统（TBS，translation-based steganography）[7] 以及基于同义词替换的信息隐藏算法[2]进行了对比. 嵌入容量为文本每比特中所藏的密文比特数，计算公式如下：

$$\text{Embedding rate}=\frac{\text{嵌入密文比特数}}{\text{生成文本字符数}\times 8} \tag{8}$$

其中 8 表示英文字符的编码位数. 嵌入容量结果如表 2 所示.

表 2　几类隐藏算法嵌入容量对比

信息隐藏模型	嵌入容量/%	文献
Nicetext	0.29	[9]
TBS	0.52	[7]
同义词替换	0.68	[2]
SSH（beam size 为 4）	0.56	
SSH（beam size 为 32）	1.41	
WWH	**2.75**	基于本文算法
HH（beam size 为 4）	2.71	
HH（beam size 为 32）	2.67	

从表 2 中可以看出，基于图像描述生成的信息隐藏方法，相对于其他几种自然语言信息隐藏算法来说，在嵌入容量上具有较好的优势. 特别是利用基于词层面的隐藏 WWH 和 HH，每一个单词可以隐藏 1 bit 信息，比目前主流的 1 个句子隐藏 1 bit 的 TBS 算法和基于同义词替换的算法来说，具有明显的优势. SSH 算法从整体句子上做隐藏，但通过改变 beam size 的大小可以改变 1 个句子可载密的比特数，因此也可以通过扩大 beam size 的大小来得到理想的隐藏容量.

3）图像生成描述具体实例

下面给出具体的生成文本实例. NIC 生成的是不载密的文本,其他方法生成的都是载密文本. 嵌入的秘密信息内容为:Within the text of most pages. 为了嵌入该信息,WWH 和 HH 均用了 22 幅测试图像进行秘密信息生成,SSH(beam size 为 4)采用了 103 幅测试图像,SSH(beam size 为 32)采用了 45 幅测试图像. 对于每种算法,针对其中一幅测试图像(见图 3),给出了不同 beam size 下的具体生成句子.

图 3 测试图像

NIC(beam size = 32): two dogs play in the snow

WWH: two dogs are playing in the snow

SSH(beam size = 4): two dogs running in the snow

SSH(beam size = 32): two brown dogs are playing in the snow

HH(beam size = 4): two dogs fight in the air while running around

HH(beam size = 16): brown dogs run through the shallow snow

4 结束语

笔者提出了一种基于神经网络图像描述的自然语言信息隐藏模型. 将 CNN 与 LSTM 网络联合起来,建模图像特征和描述句之间的关系. 在 LSTM 网络基于最大似然概率进行采样预测输出的基础上,改动了 beam search 的搜索方式,提出了共享参数型信息隐藏模型和无需共享参数的 HH 模型. 共享参数型信息隐藏模型的 SSH 算法通过句子层面的采样编码输出载密描述句,WWH 算法基于单词层面进行采样编码输出载密描述句. HH 模型结合密钥散列函数获取单词的信息位,并在每一时刻 LSTM 采样时选择符合密文编码的单词进行隐写文本生成. 所提信息隐藏方法,属于一种“无载体”的自然语言生成法信息隐藏模型,具有较好的隐蔽性与安全性.

参考文献:

[1] Fu Zhangjie, Sun Xingming, Shu Jiangang, et al. New forensic methods for ooxml format documents[J]. Lecture Notes in Computer Science, 2014(8389): 503-513.

[2] Yajam H, Mousavi A. A new linguistic steganography scheme based on lexical substitution[C]// International Isc Conference on Information Security and Cryptology. Tehran: IEEE Press, 2014: 155-160.

[3] Cao Qi, Sun Xingming, Xiang Lingyun. A secure text steganography based on synonym substitution[C]// IEEE Conference Anthology. China: IEEE Press, 2013: 1-3.

[4] Topkara M, Topkara U, Atallah M J. Information hiding through errors: a confusing approach[J]. SPIE Proceedings, 2007(6505): 1-12.

[5] 戴祖旭, 洪帆, 崔国华. 基于词性标记串统计特性的文本数字水印算法[J]. 通信学报, 2007, 28(4): 108-113.
Dai Zuxu, Hong Fan, Cui Guohua. Watermarking text document based on statistic property of part of speech string[J]. Journal on Communications, 2007, 28(4): 108-113.

[6] Jin C, Zhang D, Pan M. Chinese text information hiding based on paraphrasing technology[C]// International Conference of Information Science and Management Engineering. China: IEEE Press, 2010: 39-42.

[7] Grothoff C, Grothoff K, Stutsman R. Translation-based steganography[J]. Journal of Computer Security, 2009, 17(3): 269-303.

[8] Mckellar D. Spammimic[EB/OL]. 2017. http://www.spammimic.com.

[9] Chapman M, Davida G. Nicetext[EB/OL]. 2017. http://www.securityfocus.com/tools/1183.

[10] He J, Zhou M. Generating Chinese metrical poetry by a statistical MT approach[J]. Journal of Chinese Information Processing, 2010, 24(2): 96-103.

[11] Luo Yubo, Huang Yongfeng, Chang Chinchen, 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.

[12] Kyunghyun C, Dzmitry B, Fethi B, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation[C]// Conference on Empirical Methods in Natural Language Processing (EMNLP). Qatar: Association for Computational Linguistics, 2014: 1724-1734.

[13] Sermanet P, Eigen D, LeCun Y, et al. Overfeat: integrated recognition, localization and detection using convolutional networks[C]//International Conference on Learning Representations. Banff: arXiv preprint, 2013: 1312, 6229.

[14] Hochreiter S, Schmidhuber J. Long short-term memory [J]. Neural Computation, 1997, 9(8): 1735-1780.

[15] Oriol V, Alexander T, Samy B, et al. Show and tell: a neural image caption generator[C]//Computer Vision and Pattern Recognition. Boston: IEEE press, 2015: 3156-3164.

[16] Kelvin X, Jimmy L B, Ryan K, et al. Show, attend and tell: neural image caption generation with visual attention[C]//Proceeding of the 32nd International Conference on Machine Learning. France: ACM press, 2015: 2048-2057.

[17] Wu Q, Shen C S H, Liu L Q, et al. What value do explicit high level concepts have in vision to language problems[C]//Computer Vision and Pattern Recognition. Lasa Vegas: IEEE press, 2016: 203-212.

[18] Andrei K, Li F F. Deep visual-semantic alignments for generating image descriptions[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(4): 664-676.

[19] Rashtchian C, Young P, Hodosh M, et al. Collecting image annotations using amazon's mechanical turk[C]// In NAACL HLT Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk. California: ACM press, 2010: 139-147.

[20] Kishore P, Salim R, Todd W, et al. Bleu: a method for automatic evaluation of machine translation[C]//Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics. Philadelphia: ACM press, 2002: 311-318.

(上接第 6 页)

[13] Wang Minghua, Yin Heng, Bhaskar A V, et al. Binary code continent: finer-grained control flow integrity for stripped binaries[C]//Proceedings of the 31st Annual Computer Security Applications Conference. New York: ACM, 2015: 331-340.

[14] Ge Xinyang, Talele N, Payer M, et al. Fine-grained control-flow integrity for kernel software[C]//IEEE European Symposium on Security and Privacy (EuroS&P). New York: IEEE Press, 2016: 179-194.

[15] Davi L, Dmitrienko A, Egele M, et al. MoCFI: a framework to mitigate control-flow attacks on smartphones[C]//Annual Network and Distributed System Security Symposium, San Diego, February 2012.

[16] Pewny J, Holz T. Control-flow restrictor: compiler-based CFI for iOS[C]//Proceedings of the 29th Annual Computer Security Applications Conference. New York: ACM, 2013: 309-318.

[17] Tice C, Roeder T, Collingbourne P, et al. Enforcing forward-edge control-flow integrity in GCC & LLVM[C]//Proceedings of USENIX Conference on Security. Berkeley: USENIX Association, 2014, 26: 27-40.

[18] Payer M, Barresi A, Gross T R. Fine-grained control-flow integrity through binary hardening[C]//International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment. Berlin: Springer, 2015: 144-164.