

文章编号:1003-0077(2005)01-0056-07

## 自然语言文本水印

张宇,刘挺,陈毅恒,赵世奇,李生

(哈尔滨工业大学 计算机学院信息检索实验室,黑龙江 哈尔滨 150001)

**摘要:**本文主要介绍了基于自然语言处理的文本水印技术,也即自然语言文本水印技术。该技术是在不改变文本原意的前提下,将需要隐藏的文本信息(水印信息)插入到原始文本中的一种信息隐藏技术。这种技术对于确认信息来源和信息的秘密传送,以及版权维护等方面都有着很大的应用价值。本文首先给出了基于自然语言处理技术的文本水印的概念、特点及攻击模型,并对文本水印的研究现状进行了分析。通过分析可以看出,自然语言文本水印技术有着更好的灵活性,并且在适度的攻击下,不会破坏水印信息。本文详细介绍了文本水印系统的设计过程,包括该技术的基础数学理论-二次余数理论。最后详细介绍了两种自然语言文本水印嵌入方法,分别是基于句法分析和基于语义的水印嵌入方法。

**关键词:**人工智能;自然语言处理;文本水印;二次余数;本体语义

**中图分类号:**TP391.2

**文献标识码:**A

## Natural Language Watermarking

ZHANG Yu, LIU Ting, CHEN Yi-heng, ZHAO Shi-qi, LI Sheng

(Information Retrieval laboratory, School of computer science and technology, Harbin institute of technology, Harbin, Heilongjiang 150001, China)

**Abstract:** A new technique based on natural language processing was proposed in this paper, that is, natural language watermarking. It is a novel technique for information hiding. The meaning of the original text can not be changed after embedding the hiding information (watermark message) in it using this technique. Firstly, the concept, characteristic and the adversary model of natural language watermarking were presented in this paper. This paper also investigated some related research works in this field. This method was more flexible than traditional methods, and the watermark can not be damaged under moderate attacks. Secondly, the design of watermarking system, including the theory of quadratic residue which is the basis theory of this method were described in detail in this paper. Finally, two marking schemes were described in detail, the syntactic watermarking approach and the semantic watermarking approach.

**Key words:** artificial intelligence; natural language processing; natural language watermarking; quadratic residue; ontological semantics

## 1 引言

随着计算机技术的不断发展以及互联网的普及,现代社会进入了信息爆炸的时代,信息在网络上大量的发布、传递。信息的接收者要注意信息的来源是否可靠,信息的发布者要考虑到新闻或者文章的版权问题,军方在信息传递的过程中要确保信息的保密性,这就需要有有效的方

收稿日期:2004-02-19

基金项目:国家自然科学基金资助(60435020);哈尔滨工业大学校科学研究基金资助(hit.2002.73)

作者简介:张宇(1972—),男,博士,副教授,研究方向为自然语言处理,信息检索,信息安全。

法来对信息进行保护。

信息主要由两种基本保护方法:一种是利用密码术对明文实施各种变化,使它不为局外人所理解。这种利用随机性来对抗密码攻击的技术,在防止他人从中得到信息具体内容的同时,也暴露了消息的重要性。在实际应用中,这一点常常是需要极力避免的;另一种方法就是信息隐藏技术,利用载体信息中具有随机特征的冗余部分,将重要信息潜入载体信息中,使其不被其他人发现。在实际应用中,存在冗余信息的载体非常丰富,这一点也在客观上增强了信息隐藏技术的隐蔽性和可行性。同时,通过把信息存在本身隐藏起来的技术使得攻击者无从获取秘密信息的位置,从而增强了安全性<sup>[1]</sup>。

基于自然语言处理的文本水印技术是近几年来被广泛关注的研究问题,国外的很多学者对此进行了大量的研究,美国普渡大学的 Victor Raskin 教授、日本塾庆应义大学的 Kosin Chamrongsathai 教授等学者对基于自然语言处理的文本水印技术及其基础的理论进行了大量系统的研究。目前,我国在数字水印(图象水印、音频水印)方向上进行了大量的研究,但是在自然语言处理领域尚未开展对水印技术的研究。

## 2 自然语言文本水印及其攻击模型

### 2.1 自然语言文本水印的概念

什么是自然语言文本水印?目前仍然没有一个对它的准确定义。根据数字水印技术以及自然语言文本水印自身的特性。我们认为自然语言文本水印是指(1)通过对语言的理解,在给定文本里,利用等价信息替换,语态转换等办法把水印信息嵌入文本;(2)被嵌入的信息通常是不可见或不可察的,但是通过特定算法可以被检测或者被提取;(3)水印信息可以经历一些不破坏源文本使用价值或商用价值的操作而存活下来。

上述定义中,(1)说明了自然语言文本水印在加入水印过程中所采用的方法与数字水印不同,我们根据文本本身所固有的特点,采用了适合于文本转换和表达的方法,已达到将水印信息加入到文本中的目的。(2)中说明了自然语言文本水印的最根本的特征,在加入水印的文本中,水印信息并不是直接体现在文本中,读者只能通过特定的算法来读取出水印信息,这样体现了水印信息的隐蔽性。(3)说明了我们加入的水印信息应该是不易被破坏的,也就是说,只要其它人在不破坏原意的情况下对文本进行攻击,我们加入的水印信息仍然能够读取出来。

### 2.2 文本水印的研究现状

自然语言文本水印的设计思想同样遵循图像水印设计思想:除了文本的作者或者拥有者,其他任何人都不能从中检测出水印信息。但是,在文本中加入水印信息更加困难,原因在于:“与图像中的噪声数据不同,文本中不包含用于秘密信息传递的冗余信息”<sup>[2]</sup>。对于文本水印进行的最初的尝试,是在文本图像中加入水印<sup>[3~5]</sup>,或者水印的加入是在结构化文本的基础上进行的,如 LaTeX、HTML 或 PostScript<sup>[2]</sup>。格式化文本中,采用标签来实现文本的格式化,如 HTML 中的 head、title 等等。这种方式下,水印的嵌入主要是针对标签来实现的。由于在标签中添加一些附加的信息对文本并不产生影响,所以水印可以加在标签中,如标签 head,可以在其中嵌入一个水印信息“1”,变成 head 1。对于用户来讲,他看到的还是原来的 HTML 网页,但是网页中的内容却嵌入了水印信息。但是,如果用户查看源代码的话,隐藏的信息就会被用户发现,不能实现很好的隐藏水印信息。

由于上述的通过在文本图像和格式化文本中加入水印信息的方法的安全性较差,因此一些学者们试图采用插入拼写字母、词的变换、标点符号甚至一些错误的内容等方法来实现这个

目的。包括还有同义词替换的方法<sup>[6]</sup>,但是这种方法中,如果我们要提取水印信息的话,一定要有原始文本作为参照,否则水印信息无法提取出来,这样限制了水印系统的灵活性。以上这几种方法都不能够实现一个健壮的文本水印系统,这些系统很容易被攻击者攻击,同时也会降低文本的质量。

自然语言处理是为了在一些特定的应用中来自动地处理自然语言书写的文本,如机器翻译、信息检索、文摘、信息抽取、数据挖掘和智能搜索引擎等。自然语言处理技术经过多年的发展,积累了大量的经验和技巧,如分词、句法分析、重述(Paraphrasing)<sup>[7]</sup>、词义消歧等等。这些都为自然语言文本水印技术奠定了基础。和以往的方法相比较,通过采用各种自然语言处理技术嵌入的水印信息更加安全、可靠。本文的主要指导思想就是在不改变文本原意的情况下通过对句子的变换以实现水印的加入和提取。

### 2.3 自然语言文本水印的特点

给定一个自然语言文本  $T$ ,令  $W$  是要插入的水印信息,且  $W < T$ 。在文本中插入水印信息后,生成含有水印信息的文本  $T'$ ,我们希望<sup>[2]</sup>:

- ①  $T$  和  $T'$  表达相同的意思;
- ②  $T'$  中含有水印信息  $W$ ,该水印信息可以作为识别信息来源或者处理版权纠纷问题的依据等;
- ③ 水印信息  $W$  在  $T'$  中不可以被直接读出,只能通过一个预先设定的密钥读出该信息;
- ④ 如果某人有预先设定的密钥,那么他不需要任何与文本  $T$  相关的知识,可以直接从  $T'$  中读出水印信息;
- ⑤ 除非知道这个预先设定的密钥,否则,水印信息  $W$  在不改变文章原意的情况下很难从  $T'$  中获得;
- ⑥ 向文本  $T$  中加入水印的过程和从文本  $T'$  中提取水印的过程不是保密的,而是预先设定的密钥保证了该方案的安全性;
- ⑦ 如果两个人  $A$  和  $B$  在同一个文本里面加入了两个不同的水印,那么  $A$  不能读出或删除  $B$  所加入的水印信息,对  $B$  也一样。

以上就是自然语言文本水印的特点,在构造水印系统的时候,应该满足以上的七个特点,这样所构造的系统才是完整的、健壮的、安全的。

### 2.4 攻击模型

攻击方的目的是为了在不大量改变句子原意的条件下破坏文本中所包含的水印信息。为了达到这个目的,攻击方通常采用以下措施来对含有水印的文本进行攻击<sup>[2,3]</sup>:

- ① 对文本中的句子进行保留意义的转换不能够破坏水印信息;
- ② 如果将句子的意义进行了改变,那么有可能会破坏文本中的水印信息;
- ③ 在文本中插入一个新的句子,也有可能破坏文本中的水印信息;
- ④ 将文本中一块连续的部分从一个地方移动到另外一个地方,也有可能破坏文本中的水印信息。

水印系统在实现的过程中一定要综合考虑到攻击方所采用的上述几种方法,以便接受方能够正确地从文本中提取出水印信息。

## 3 在文本中加入水印的过程

本文中所提到的自然语言文本水印,是将水印加入到纯文本的电子文档中,而不是经过扫

描之后的文本图像中或者打印出来的文档中。在文本中加入水印的过程通常都是相同的,第三部分主要介绍文本水印的嵌入过程。

### 3.1 二次余数理论

在文本中加入水印的过程中采用二次余数理论的目的是为了使 0,1 在句子中能够近似均匀地分布。这样使得文本通过适当地转换后,能够准确地将水印信息加入到文本中去。

设  $P$  是一个很大的非偶质数,  $X$  是任意一个自然数,如果  $X \times X = KP + N$  ( $K$  是自然数,  $N < P$ ),则  $N$  是质数  $P$  的一个二次余数(0 既不是二次余数,也不是非二次余数)<sup>[8]</sup>。现在我们再举两个例子来说明这个理论。

例 1:  $P=5$ , 求  $P$  中的二次余数和非二次余数

$X=1, X \times X=0 \times 5+1$ , 此时  $N=1$ 。  $X=2, X \times X=0 \times 5+4$ , 此时  $N=4$ 。

则我们称 1,4 是对于质数 5 的二次余数,而 2,3 为非二次余数

例 2:  $P=7$ , 求  $P$  中的二次余数和非二次余数

$X=1, X \times X=0 \times 7+1$ , 此时  $N=1$ 。  $X=2, X \times X=0 \times 7+4$ , 此时  $N=4$ 。

$X=3, X \times X=1 \times 7+2$ , 此时  $N=2$ 。

则我们称 1,2,4 是对于质数 7 的二次余数,而 3,5,6 为非二次余数。

我们可以发现对于某一个质数,当一个数(模  $P$ )为二次余数,则用“1”表示,否则用“0”表示,使用二次余数理论可以近似把节点平均分配成 0 或 1。

### 3.2 在文本中加入水印的过程

有了上述的二次余数理论,下面我们就可以进行水印的加入和提取了。水印的嵌入过程分为以下几个步骤<sup>[9,10]</sup>:

第一步:我们要插入的水印信息是文本形式的,但是直接把文本插入文章中会很容易被其他人发现。所以,我们首先把文本形式的水印信息转换成二进制代码形式的信息。这样把文本水印用二进制形式按位插入到文本中(水印信息转换为二进制代码后,字符串的长度假设为  $N$ )。同时给定一个预先设定的密钥  $P$ (很大的非偶质数,20 位的十进制数);

第二步:我们把一篇文章看成一组句子的组合。并给每个句子从 1 开始按顺序进行编号;

第三步:对每个句子进行句法分析。找到句子中的核心词,并分析句子中词和词之间的依存关系。根据词之间的依存关系,将句子用树型结构进行表示;

第四步:对于每个句子的树型表示,首先按照先根的顺序对树中的节点进行编号(从 1 开始)。然后构造散列函数  $H(p)$ ,计算  $w = i + H(p)$  ( $i$  为树中节点的编号)。如果  $w$  是二次余数(模  $p$ )的话,就用“1”替代该节点的编号;否则,用“0”替代该节点的编号。这样,对每一个节点都进行处理后,树的节点就会得到新的编号(为 0 或者 1)。然后按照后根遍历的顺序可以得到表示该句子的 0,1 字符串;

第五步:由于句子长短不同,得到的字符串长短也不一致,按首对齐降序排列(或升序排列)的办法把这些串排序。而排在这个序列里的前  $N$  个串(后  $N$  个串)所表示的句子,称之为“标识句”。插入水印的句子是在文本中标识句的下一个句子,该句子称之为“水印句”。一定要注意的是,水印的信息是嵌入在水印句中的;

第六步:在第五步中,标识句确定了水印信息要加入到哪个句子中。同时,标识句也确定了水印信息要嵌入到水印句的一位上。位置的确定与所选取的散列函数有关;

第七步:按顺序选择标识句,确定了要嵌入水印的句子和位置之后,下一步我们要看水印句中指定的位置上是不是我们所要嵌入的水印信息。例如,我们要把“1”作为水印信息嵌入到

水印句的第三位上,如果这时候水印句的第三位恰好是“1”,则不做任何变换,完成该水印的加入过程;如果这时候水印句的第三位为“0”,则需要对原始文本中的水印句进行变换(嫁接、剪枝、等价信息替换等),使水印句的第三位变换为“1”,以完成水印的嵌入过程。重复该步骤,直到所有的水印信息都加入完毕。

水印信息的提取过程是插入的一个逆过程,在得到嵌有水印信息的文本后,我们按第二步、第三步、第四步和第五步的过程得到排序后的字符串。按照升序(或降序)的顺序确定标识句,然后从水印句的指定位置上将水印信息依次提取出来。

由上述的过程可以看出,20位质数的密钥决定了水印信息要嵌入到文本中的哪一个句子中,而且这些句子是通过特定的方法(对所选择的要插入水印信息的句子进行句法分析,对句法分析的结果进行先根遍历或后根遍历)随机选取的,并不是顺序的。有些句子在嵌入水印位的时候产生了变化,但是有些句子嵌入水印位过程中并没有产生任何变化。这样即便攻击方得到了原文,看到了哪些句子产生了变化,也没有办法知道里面真正隐含的信息是什么,因为某些水印位可能嵌入在没有发生变化的句子中。所以,只要攻击方得不到密钥,就没有办法得到文本中嵌入的水印信息。这对于军事领域中保密情报的传递有着很重要的应用价值。

## 4 自然语言文本水印设计方法

由于以往的水印技术局限性比较大,而且极大程度地降低了文本的质量。所以,本文主要讨论基于自然语言处理技术的文本水印的设计方法。在基于自然语言处理技术的文本水印系统的设计过程中,通常采用两种方法来实现水印的嵌入:基于句子结构的和基于语义的方法<sup>[9,10]</sup>。

### 4.1 基于句法结构的自然语言文本水印设计方法

该方法主要是对句子的句法结构进行转换,以达到在文本中加入水印的目的。在这种方法中,公认的最常用的变换方式有以下四种:

#### 1 移动附加语的位置

与补语不同,句子中附加语(如前置短语、状语等)的位置是可以作适当移动的。比如下面的例句:

® 孔令辉最后夺得了金牌。

句中的状语“最后”可以移至句首。这样上面的例句被转换为:

® 最后孔令辉夺得了金牌。

#### 2 加入形式主语

这种变换方式对于英文句子来讲是很容易实现的,但由于汉语语法中没有“形式主语”,因此这种变换方式无法直接应用于汉语的句法结构变换中。然而我们可以使用与之相似的变换方式。比如对于上面的例句,可以变换为:

® 是孔令辉最终夺得了金牌。

#### 3 主动式变被动式

无论是在汉语中还是在英语中,任何含有及物动词的句子都可以由主动式变换成被动式。由于对汉语句子的主动-被动转换中不用考虑人称、时态和语态的问题,因此相比较而言,这种变换方式更容易在汉语中实现。同样是对上面的例句,对其进行主动-被动转换之后将变为:

® 金牌最后被孔令辉夺得了。

#### 4 在句子中插入“透明短语”

这里所谓的“透明短语”是指一些几乎不含语义信息的习惯表达,比如:“众所周知……”、“正如我们看到的……”、“值得一提的是……”等等。这些词或短语的加入不会影响句子的语义。利用这种方式对上面的例句进行变换后,我们将得到以下结果:

正如我们看到的,孔令辉最后夺得了金牌。

以上种种变换方式虽然各不相同,但却存在着几个共同的特点:

1) 都会使句子的句法结构发生变化,亦即句子句法树的形状发生变化,进而使得变换前后句子的二进制编码变的不同。

2) 都存在可逆变换:以上四种变换方式均是可逆的。例如:“主动变被动”的逆变换即是“被动变主动”,“插入‘透明短语’”的逆变换是“删除‘透明短语’”等。

3) 几种方法可以同时使用:有时仅使用一种方式对句子进行变换后可能仍然无法插入水印信息,这是便可以综合使用上面介绍的各种方法。例如,综合运用上面四种变换方式,我们的例句变为:

® 正如我们看到的,最后金牌是被孔令辉夺得了。

#### 4.2 基于语义的自然语言文本水印的设计方法

该方法中,主要是在基于对句子进行深层的理解的基础上对句子进行变换,以达到在文本中加入水印的目的。

美国普渡大学的 Victor Raskin 等人在本体语义 (Ontological Semantics) 的基础上,采用 TMR (Text Meaning Representation) 树的方式对文本中的句子进行表达<sup>[11]</sup>。通过对 TMR 树的操作来实现对文本中句子进行修改。对 TMR 树进行操作主要有以下三种方式:嫁接 (Grafting)、剪枝 (Pruning) 和等价信息替换 (Adding/ Substitution)<sup>[9]</sup>。

嫁接的方法,主要是根据上下文的有关信息来进行操作的。例如,在文章中其中一个句子的 TMR 树的表示如下:

® assault-1 —| —agent —nation-1 —“United States”  
| —theme —nation-4 —“Afghanistan”

这时可以看到句子中表示被攻击的对象是 Afghanistan,而实施者是 United States。如果文章中出现下面这句话:

® The Pentagon ordered two new spy planes to the region to start flying over Afghanistan.

这时我们可以使用嫁接的方法来对句子进行修改:

® The Pentagon ordered two new spy planes to the region to start flying over Afghanistan, which they are attacking.

剪枝主要是对上下文中一些重复的信息来对句子进行修改,如果某个概念在文章中的其它的地方出现了,那么可以将其进行剪枝处理,这样的处理也不会影响文章的意思。例如,文章中 Washington 出现了五次,第一次出现的时候不可以进行剪枝的处理,但是后面出现的 Washington 都可以进行处理。再比如下面的例子中,文章中多次出现了 Afghanistan,那么在第一个 Afghanistan 后面出现的词都可以进行剪枝处理(如果在文中,前面已经出现过了关于美国攻打阿富汗的信息,那么这些例句中的斜体部分都可以从句子中去除):

® With no visible victory so far in Afghanistan, President Bush asserted that the campaign he launched in reprisal for September 's mass killings on U. S. soil was going well, and he urged Americans to remain patient.

® In Pakistan ,which is backing U. S. strikes on Afghanistan ,a minister said official tests confirmed that at least one suspicious letter received there contained anthrax spores.

® The Pentagon ordered two new spy planes ,including the unmanned “ Global Hawk ”,to the region to start flying over Afghanistan.

等价信息替换的方法不同于同义词替换。这种方法中的等价信息主要是来源于事实数据库(Fact Database) ,该数据库是本体语义中的一个静态资源。例如 ,在数据库中有这样一段记录:

Afghanistan (nation - 1)

borders - on China ,Iran ,Pakistan ,Tajikistan ,Uzbekistan

has - currency afghani

has - member Pashtun ,Tajik ,Hazara ,Uzbek

has - representative Mullah Mohammad Omar

.....

通过在数据库中查寻记录 ,可以发现 ,目前 Afghanistan 的领导人是 Mullah Mohammad Omar。因此 ,我们可以通过数据库中的信息来对原文中的信息进行替换 ,以实现句子的修改。该句子可以变换如下:

® The United States are attacking the country ruled by Mullah Mohammed Omar.

以上介绍的三种基于语义的对句子进行修改的方法 ,是在水印嵌入过程中常用的方法。通过对句子在进一步理解的基础上进行修改、变换 ,实现水印的嵌入。

## 5 结论

综上所述 ,通过自然语言处理的方法实现在文本中加入水印信息 ,既提高了系统的灵活性 ,又增加了系统承受攻击的能力 ,弥补了同义词替换等方法的不足。由于该技术的研究刚刚起步 ,目前还存在很多尚待解决的问题 ,如在一个篇幅较短的文本中如何加入水印信息 ,如何对文本中的句子进行变换以便有效地向文本中插入水印信息 ,如何对文本水印系统进行评价 ,如何进一步增强系统的防攻击能力等等 ,这些都要借助于自然语言处理技术。基于自然语言处理的文本水印技术为在信息传播的过程中提供了更大的安全性、保密性 ,在国防、国民经济、生活等领域有着广泛的应用前景和重要的应用价值。目前 ,对于自然语言文本水印技术的研究在国内尚属空白 ,哈工大信息检索实验室在已有的中文信息处理技术的基础上 ,开展了对该技术的研究 ,目前已经取得初步的成效 ,已成功实现了一个自然语言文本水印的演示系统。

## 参 考 文 献:

- [1] 刘振华 ,尹萍编著. 信息隐藏技术及其应用[M]. 北京 :科学出版社 ,2002.
- [2] Katzenbeisser ,S. C. Principles of Steganography. In :S. Katzenbeisser and F. A. P. Petitcolas (eds. ) ,Information Hiding [C]. Techniques for Steganography and Digital Watermarking. Boston :Artech ,2000 ,17 - 41.
- [3] Brassil J. ,N. F. Maxemchuk ,and L. O 'Gorman. Electronic Marking and Identification Technique to Discourage Document Copying[C]. Proceedings of INFOCOM '94 ,1994 ,1278 - 1287.
- [4] Maxemchuk ,N. F. 1994. Electronic Document Distribution[J]. AT&T Technical Journal ,September/ October ,73 - 80.
- [5] Low ,S. H. ,N. F. Maxemchuk ,and A. M. Lapnoe. Document Identification for Copyright Protection Using Centroid Detection[J]. IEEE Transactions on Communication , 1998 ,46(3) ,372 - 383

(下转第 70 页)

- text[A]. In: Proc. of COLING- ACL '98[C]. 1998, 1098 - 1102.
- [4] 董振东. 语义关系的表达和知识系统的建造[J]. 语言文字应用, 1998, (3): 76 - 82.
- [5] 梅家驹, 等. 同义词词林[M]. 上海: 上海辞书出版社出版, 1983.
- [6] Pease, A., Niles, I., and Li, J. The Suggested Upper Merged Ontology: A Large Ontology for the Semantic Web and its Applications [A]. In: Working Notes of the AAAI - 2002 Workshop on Ontologies and the Semantic Web[C]. Edmonton, Canada: 2002.
- [7] Niles, I. and Pease, A. Linking Lexicons and Ontologies: Mapping WordNet to the Suggested Upper Merged Ontology [A]. In: Proceedings of the 2003 International Conference on Information and Knowledge Engineering[C]. Las Vegas, Nevada: 2003.
- [8] D. Yarowsky. Word Sense Disambiguation Using Statistical Models of Porter's Categories Trained on Large Corpora [A]. In: Proc. Of COLING '92[C]. Nantes, France: 1992, 454 - 460.
- [9] Ji Donghong, Chen Junqin, Huang Changning. Combining a Chinese Thesaurus with a Chinese Dictionary [A]. In: Proc. of COLING- ACL '98[C]. 1998, 600 - 606.
- [10] 杨尔弘, 黄昌宁, 李涪子. 《现代汉语词典》的义类标注[A]. 黄昌宁, 董振东主编. 见: 计算语言学文集[C]. 北京: 清华大学出版社, 1999, 167 - 173.
- [11] Marine Carpuat, Grace Ngai, Pascale Fung and et al. Creating a Bilingual Ontology: A Corpus - Based Approach for Aligning WordNet and HowNet [A]. In: Proceedings of the 1st Global WordNet Conference[C], 2002.
- [12] 陈祖舜, 周强, 赵强. 情境——组织/存放词汇语义知识的恰当框架[J]. Computational Linguistics and Chinese Language Processing, 2002, 7(2): 1 - 36.

## (上接第 62 页)

- [6] Atallah, M. J., C. J. McDonough, V. Raskin, and S. Nirenburg. Natural Language Processing for Information Assurance and Security: An Overview and Implementations[C]. In: M. Shaeffer (ed.), NSPW '00: Proceedings of Workshop on New Paradigms in Information Security, Cork, Ireland, September 2000. New York: ACM Press, 51 - 65.
- [7] 张玉洁, 山本和英. 汉语语句的自动改写[J]. 中文信息学报, 2003(6): 31 - 38.
- [8] Mikhail J., Atallah, Samuel S., Wagstaff, Jr.. Watermarking with Quadratic Residues[C]. Working Paper, Department of Computer Science, Purdue University, 1996.
- [9] Mikhail J. Atallah, Victor Raskin, Michael Cogan, Christian Hempelmann, Florian Kerschbaum, Dina Mohamed and Sanket Naik. Natural Language Watermarking: Design, Analysis, and a Proof - of - Concept Implementation[C]. Information Hiding 2001: 185 - 199
- [10] Mikhail J. Atallah, Craig J. McDonough, Victor Raskin. Natural Language for Information Assurance and Security: An Overview and Implementation[C]. Published in: M. Schaeffer (ed.), Proceedings. New Security Paradigm Workshop. September 18th - 22nd, Ballycotton, County, Cork Ireland. New York: ACM Press, 2000, pp. 51 - 65.
- [11] Nirenburg, S., and V. Raskin 2003. Ontological Semantics. Cambridge [M], MA: MIT Press (forthcoming). Pre - publication draft, <http://crl.nmsu.edu/Staff/Pages/Technical/sergei/book/index-book.html>.