

# 安全的 LZW 编码算法及其在 GIF 图像加密中的应用

向涛\*, 王安

(重庆大学 计算机学院, 重庆 400030)

(\* 通信作者电子邮箱 txiang@cqu.edu.cn)

**摘要:**提出了一种安全的 LZW 编码算法——SLZW。该算法在改进的 LZW 编码过程中嵌入加密,从而能够同时完成压缩和加密。SLZW 编码利用动态 Huffman 树作为 LZW 的字典,并且通过耦合映像格子(CML)产生的密钥流对字典的构建和更新进行控制,编码输出进一步和密钥流进行异或后产生密文。并且,该算法被应用于 GIF 图像加密中,实验结果和分析表明,该算法不仅具有较好的安全性,同时也将标准 LZW 算法的压缩效率提高了 10% 左右,具有广泛的实用性。

**关键词:**数据压缩;图像加密;Huffman 编码;耦合映像格子

**中图分类号:** TP309 **文献标志码:** A

## Secure LZW coding algorithm and its application in GIF image encryption

XIANG Tao\*, WANG An

(School of Computer Science, Chongqing University, Chongqing 400030, China)

**Abstract:** This paper proposed a Secure LZW (SLZW) coding algorithm, where encryption was embedded into the improved LZW coding process, and SLZW can fulfill compression and encryption in a single step. In SLZW algorithm, dynamic Huffman tree was utilized to code the dictionary of LZW, and the initialization and updating of Huffman tree were controlled by a sequence of keystream generated by Coupled Map Lattice (CML). The code words were further XORed with the keystream to generate the ciphertext. The SLZW was applied to GIF image encryption. The experimental results and their analyses indicate that the proposed SLZW algorithm not only has good security, but can also improve the compression ratio by about 10%. Therefore, SLZW can find its wide applications in practice.

**Key words:** data compression; image encryption; Huffman coding; Coupled Map Lattice (CML)

## 0 引言

随着信息的高度数字化,计算机需要处理越来越多的海量数据。特别是多媒体数据的广泛应用和普及,使得压缩成为数据存储和传输过程中一项必不可少的技术。同时,在日益复杂的网络环境下,信息的安全性也面临着日益严峻的考验,而加密是解决信息安全的重要手段。因此,在实际应用场合往往需要同时考虑数据压缩和数据加密。但是由于压缩和加密原本是两项相对独立的数据处理技术,传统的做法是分两步进行,即先对数据进行压缩,然后对压缩后的数据进行加密。这样做的问题在于处理时间开销较大,而且压缩和加密之间的速度同步也是一个问题。最新的研究表明,将压缩和加密结合起来考虑,即在数据压缩的同时对数据进行加密,将两个独立的步骤合并成为一个步骤,可以有效地减少处理的时间,特别是在图像等多媒体数据处理和实时传输方面有着非常显著的效果<sup>[1-2]</sup>。

LZW 是由 Abraham Lempel, Jacob Ziv 与 Terry Welch 创造的一种优秀的无损数据压缩算法<sup>[3]</sup>,在 GIF 等文件格式中得到了广泛的应用。目前国际上已有学者针对 LZW 压缩算法在压缩编码的同时引入加密机制方面展开了研究。文献[4]提出了一种基于 LZ78 的加密方案,该方案采用随机字典表(Randomized Dictionary Table, RDT)的方式在 LZW 压缩过程中实现加密。文献[5]也对基于 LZW 的加密方案进行了研

究,通过利用随机字典插入、字典置乱和二进制异或操作,实现对明文的压缩和加密。然而,这两个方案在安全性和压缩率方面存在缺陷;文献[5]指出文献[4]所提出的方案易受选择明文攻击;文献[6]分析了文献[5]所提出的加密算法的安全性,分别给出了选择明文和选择密文攻击的方法,同时还指出文献[5]中的算法对 LZW 算法本身的压缩效果有影响。

本文基于 LZW 压缩提出了一种安全的 LZW (Secure LZW, SLZW) 编码算法,SLZW 编码算法利用动态增长的 Huffman 树<sup>[7]</sup>作为字典,有效地将 LZW 和动态 Huffman 编码结合为一个整体,能显著提高压缩率。在编码的过程,利用基于混沌的耦合映像格子(Coupled Map Lattice, CML)产生随机比特流控制 Huffman 树的编码,实现在压缩过程中进行加密,达到较好的安全性。

## 1 LZW 编码和动态 Huffman 编码

SLZW 编码的思想是建立在 LZW 编码和动态 Huffman 编码的基础上的,下面首先对这两种编码作简单介绍。

LZW 编码的主题思想是用较短的字符串代替较长的字符串实现数据压缩,具体做法为:用一定的规则选择一些字符串放进字符串表(字典)中,当字典中的字符串再次出现时,就可以用字典中的位置索引值代替该字符串。LZW 在解码时可以构建出同编码方同样的字典,因此,编码方构建的字典没有必要发送给解码方。LZW 的编码和解码过程请参考文

**收稿日期:** 2012-06-19; **修回日期:** 2012-07-24。 **基金项目:** 国家自然科学基金资助项目(61103211); 中国博士后科学基金特别资助项目(201104319); 中央高校基本科研业务经费面上项目(CDJZR10180020)。

**作者简介:** 向涛(1980-),男,湖北荆门人,副教授,博士,主要研究方向:混沌密码学、多媒体安全; 王安(1987-),男,山东菏泽人,硕士研究生,主要研究方向:数据压缩、信息安全。

献[3]。

动态 Huffman 编码在对数据进行编码时不必事先统计数据出现的概率信息,它采用的是根据数据输入动态构建 Huffman 树的技术。更为重要的是,解码方在解码时不必事先拥有编码方构建出的 Huffman 树,解码方可以根据得到的压缩码流动态构建出与编码方一样的 Huffman 树,从而恢复出被编码的消息。这也是它和 LZW 编码结合起来后解码方能完全恢复出编码方字典的关键。文献[8]给出了一种动态 Huffman 实现的技术。

研究表明,将 LZW 与 Huffman 编码结合能取得更好的压缩效果,文献[9]就介绍了一种将 LZW 与 Huffman 编码结合进行压缩的方法。然而,这些方法均存在一个显著缺陷,就是没有将两个压缩过程无缝地合并为一个压缩过程,它们采用的思想都是在 LZW 压缩过程中计算出输出数据的概率,然后根据这些概率再对结果数据进行 Huffman 编码。本文将动态 Huffman 树结构作为 LZW 的字典,无需统计数据概率信息,一次 LZW 编码后即实现了总体压缩的效果。

## 2 SLZW 算法

### 2.1 算法原理

SLZW 编码和解码的原理如图 1 所示,它由两部分组成:一个经过改进的基于动态 Huffman 树的 LZW 编码器(或者解码器)和一个密钥流产生器(SLZW 编码和解码过程采用相同的密钥流产生器)。编码过程为:明文  $M$  在密钥流  $K$  的控制下,通过改进的 LZW 编码器,产生最终的密文  $C$ 。解码过程为其逆过程。

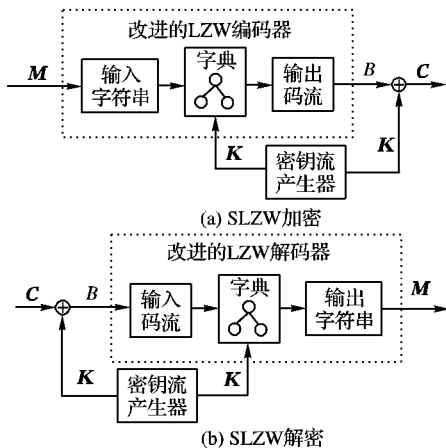


图 1 SLZW 算法原理

加密算法执行过程主要包含三个步骤:首先是产生密钥流,然后是基于动态 Huffman 树进行 LZW 编码,最后是异或产生密文输出。下面本文分别对这三个步骤进行详细的描述。解码过程为其逆过程,在此不再赘述。

### 2.2 密钥流的产生

密钥流产生分两步:首先,利用一个基于斜帐篷映射的二维 CML<sup>[10]</sup>混沌系统,CML 的定义为:

$$x_{n+1}^j = (1 - \varepsilon)f(x_n^j) + \varepsilon f(x_n^{j-1}) \quad (1)$$

其中  $x_n^j$  表示空间维为第  $j$  维( $j = 0, 1$ ),时间维为第  $n$  维( $n = 0, 1, \dots$ )的变量, $\varepsilon \in (0, 1)$ 。 $f(x)$  选择为斜帐篷混沌映射<sup>[10]</sup>函数,其定义为:

$$f(x) = \begin{cases} x/a, & 0 < x \leq a \\ (x-1)/(a-1), & a < x \leq 1 \end{cases} \quad (2)$$

其中  $a \in (0, 1)$ 。

然后,利用式(3)对产生的混沌变量进行二值化处理,生成二进制的密钥流  $K = k_0 k_1 \dots k_n \dots$ 。

$$k_n = \begin{cases} 1, & x_0^n > x_1^n \\ \text{无输出}, & x_0^n = x_1^n \\ 0, & x_0^n < x_1^n \end{cases} \quad (3)$$

通过该方法产生的二进制随机序列能够克服混沌系统在数字化过程中的精度退化问题,具有较好的密码学特性<sup>[11]</sup>。该密钥流用来对 Huffman 树的构建过程进行控制,以及与编码结果进行异或。

### 2.3 基于动态 Huffman 树的 LZW 压缩

在改进的 LZW 压缩算法中,我们将 LZW 中的字典采用动态 Huffman 树进行编码,根据字典中各词条出现的频率分配和调整其权值,从而达到提高 LZW 压缩效果的目的。

在编码的过程中,每当有新的符号(或者符号序列)加入到字典中或者字典中现有词条的频率改变时,就需要对这棵 Huffman 树进行更新<sup>[8]</sup>。Huffman 树中的节点包含 LZW 算法中的字典信息。在没有加入密钥的情况下,左子树的编码值为 0,右子树的编码值为 1。

构建初始字典的过程即为创建初始 Huffman 树的过程。将信源符号集中的每个符号(即 LZW 字典中的每个符号)赋初始权值 1,然后按照 Huffman 树的构建规则建立初始化的 Huffman 树。在将每个符号加入到 Huffman 树的叶子节点的过程中,我们依次取出密钥流中的一位,若该位的值为 1,则交换左右节点的编码值;否则不交换左右节点的编码值。

在利用 Huffman 树构建的字典进行 LZW 编码的过程中,每当一个符号(或者符号序列)被编码后,如果它在字典中已经存在,便更新其权值(将其权值加 1),然后根据新的权值调整 Huffman 树的结构;如果该符号(或者符号序列)在字典中不存在,则初始化其权值为 1,然后将其加入到 Huffman 树中。更新过程跟初始化 Huffman 树类似,由密钥流控制节点的编码值。

### 2.4 最终密文的生成

虽然本文在前面的步骤中已经通过密钥流混淆了初始化和 LZW 编码过程中产生的字典,为了掩盖密文的统计特性,进一步提高安全性,本文利用密钥流对产生的码字进行异或,产生最后的密文输出  $C = c_0 c_2 \dots c_n \dots$ ,即:

$$c_n = b_n \oplus k_n \quad (4)$$

其中  $b_n$  为改进的 LZW 编码器的输出。

## 3 实验结果与分析

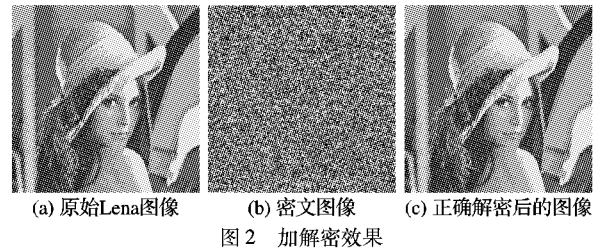
为了验证所提出的 SLZW 算法的正确性和安全性,将其应用于 GIF 图像加密中,并且采用国际上通用的 USC-SIPI 图像测试库<sup>[12]</sup>对算法进行了测试和验证(图片库中的图像默认是 tiff 格式,本文将其转换成 gif 格式后进行实验),实验结果表明 SLZW 算法在显著提高 LZW 算法压缩效率的同时达到了较好的安全性。但是由于篇幅限制,本文在此只以一幅大小为  $256 \times 256$  的 Lena 图像为例。

### 3.1 实验参数和加解密效果

产生密钥时需要设定一些初始参数,本实验中各参数值选取为: $a = 0.67899$ ,  $\varepsilon = 0.11228$ ,  $x_0^0 = 0.167862987$ ,  $x_0^1 = 0.870612439$ 。其中  $x_0^0$  和  $x_0^1$  为二维 CML 的两个初始值, $a$ ,  $x_0^0$  和  $x_0^1$  设定为算法的密钥。

图 2(a)是原始图像效果图,图 2(b)为利用 SLZW 算法加密后的结果,图 2(c)为正常解码后的图像。从图 2(b)可

以看出加密后的图片呈均匀噪声分布,不能分辨出任何关于明文图像的有意义的信息。而图 2(c)说明在给定正确密钥的情况下 SLZW 算法能够精确恢复出明文图像。



3.2 压缩效果分析

对于基于压缩的加密算法而言,一个基本的要求是不能因为引入加密而给压缩效果带来负面的影响。对于上述的 Lena 图像,SLZW 算法的压缩效果如表 1 所示。表 1 分别给出了原始图像数据的大小、经过 LZW 和 SLZW 压缩后图像数据的大小。通过计算得知,SLZW 的压缩率比 LZW 提高了 10.74%,可以明显看出 SLZW 算法在压缩效果上优于原始的 LZW 算法。

表 1 SLZW 压缩效果与 LZW 压缩效果对比 B			
图像名称	原始图像	LZW 压缩后图像	SLZW 压缩后图像
Lena.gif	65 536	63 686	56 849

3.3 安全性分析

3.3.1 密钥空间分析

一个好的加密算法必须具有足够大的密钥空间来抵抗穷举法的攻击。本文中,  $a$ 、 $x_1^1$  和  $x_1^2$  作为密钥,这三个值均是双精度浮点型。在计算机中,双精度浮点数采用 IEEE754 标准进行存储,这种标准下,浮点数的尾数值占 52 位,故本算法的密钥空间可以达到  $2^{52 \times 3} = 2^{156}$ , 具有较大的密钥空间,可以抵挡穷举法的攻击。

3.3.2 统计分析

统计分析是试图利用密文统计特性恢复出明文信息,为

了能够抵挡攻击者采用统计手段破解编码信息,加密后的图像应具有均匀的直方图,并且密文图像相邻两像素的相关性也必须尽可能地低。本实验分别对明文和密文图像的直方图,及其相邻像素的相关性进行了测试。

图 3 显示了明文图像和密文图像的直方图,通过对比,明文直方图的分布有多处峰值,而密文的直方图分布均匀,有效地掩盖了密文的统计信息,符合安全性要求。

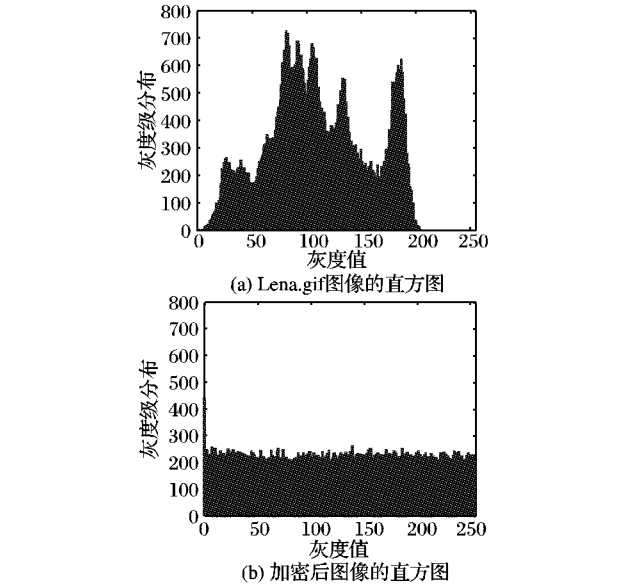


图 3 Lena 图像及其加密后图像的直方图

在相关性测试中,测试了水平、垂直和对角三个方向上两个相邻像素之间的相关性。对实验图像分别随机取出 1 000 对像素点进行测试,结果见图 4 和表 2。从图 4(a), 4(c), 4(e) 和表 2 的第 2 列可以看出,明文图像中相邻像素值呈带状分布,其相关系数都在 0.9 以上,具有很高的相关性;而图 4(b), 4(d), 4(f) 和表 2 的第 3 列表明,密文图像中相邻像素值的分布比较均匀,其相关性在 0.02 左右,从而表明密文图像中相邻像素的相关性已经非常小,满足安全性要求。

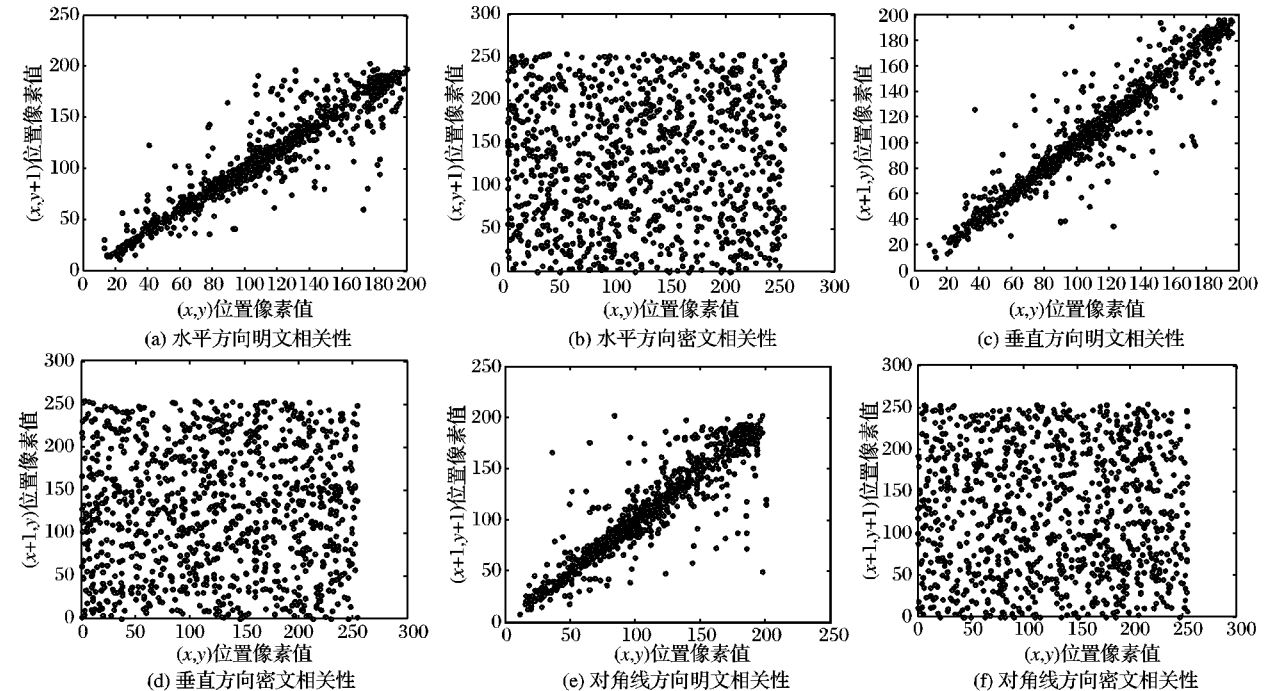


图 4 Lena 明文和密文相关性

表2 Lena 图像及其加密后图像相邻两点的相关系数

方向	明文图像	密文图像
水平方向	0.915 4	0.023 8
垂直方向	0.959 9	0.020 8
对角线方向	0.921 7	0.012 4

### 3.3.3 敏感性分析

攻击者可能利用差分分析手段来获取明文图像和密文图像之间的相关信息,这就要求加密算法具有良好的密钥敏感性和明文敏感性。本实验测试了SLZW算法对密钥和明文的敏感性。

为了测试密钥敏感性,本文分别改变 $a, x_1^1$ 和 $x_1^2$ 值中的一位得到新的密文图像,然后求其与原密文图像之间的像素改变比率(Number of Pixels Change Rate, NPCR)和像素值的平均标准改变强度<sup>[13]</sup>(Unified Average Change Intensity, UACI)。测试明文敏感性时,通过改变明文中的一个像素值来得到一幅新的密文图像,然后求其与原密文图像之间的NPCR和UACI。其中, $a, x_1^1$ 和 $x_1^2$ 改变一位后的新值分别为0.663 36, 0.167 924 022和0.870 856 579,明文敏感性测试中改变的像素值坐标为(100, 100)。

表3列出了密钥和明文敏感性实验的结果。观察可以看出:改变密钥或者明文的某一位,得到的NPCR均大于0.995, UACI在0.33左右,这表明算法具有良好的密钥敏感性和明文敏感性,能够抵抗差分攻击。

表3 Lena 图像各加密相关因素的NPCR和UACI

改变值	NPCR	UACI
$a$	0.996 1	0.337 7
$x_1^1$	0.995 8	0.339 4
$x_1^2$	0.995 1	0.339 2
明文像素点	0.995 3	0.339 2

## 4 结语

本文提出了一种安全的LZW编码算法SLZW,并将其应用于GIF图像加密中。该算法通过动态Huffman树对标准LZW算法中的字典进行编码,通过二维的CML产生密钥流控制Huffman树的初始化和动态更新,并且将编码输出和密钥流进行异或后产生密文输出。该算法将LZW编码和动态

Huffman编码无缝地融合进一个编码过程,不仅具有较好的安全性,而且显著提高了标准LZW算法的压缩率。SLZW算法具有良好的压缩效果和安全性,能够广泛应用于文本和图像的压缩及加密中,具有较强的实用性。

### 参考文献:

- [1] CHENG H, LI X. Partial encryption of compressed images and videos[J]. IEEE Transactions on Signal Processing, 2000, 48(8): 2439-2451.
- [2] WU C-P, KUO C-C. Design of integrated multimedia compression and encryption systems[J]. IEEE Transactions on Multimedia, 2005, 7(5): 828-839.
- [3] WELCH T. A technique of high-performance data compression[J]. IEEE Computer, 1984, 17(6): 8-19.
- [4] XIE D, KUO C-C. Secure Lempel-Ziv compression with embedded encryption[C]// Proceedings of SPIE. Boston: SPIE, 2005, 5681: 318-327.
- [5] ZHOU JIANTAO, AU O C, FAN XIAOPENG, et al. Secure Lempel-Ziv-Welch (LZW) algorithm with random dictionary insertion and permutation[C]// IEEE International Conference on Multimedia and Expo. Piscataway: IEEE, 2008: 25-248.
- [6] LI SHUJUN, LI CHENGQING, KUO J. On the security of a secure Lempel-Ziv-Welch (LZW) algorithm[C]// IEEE International Conference on Multimedia and Expo. Piscataway: IEEE, 2011: 1-5.
- [7] HUFFMAN D. A method for the construction of minimum redundancy codes[J]. Proceedings of the IRE, 1952, 40(9): 1098-1101.
- [8] KNUTH D. Dynamic Huffman coding[J]. Journal of Algorithms, 1985, 6(2): 163-180.
- [9] PERL Y, MEHTA A. Cascading LZW algorithm with Huffman coding: a variable to variable length compression algorithm[C]// Proceedings of the First Great Lakes Computer Science Conference on Computing. London: Springer-Verlag, 1991: 170-178.
- [10] 廖晓峰,肖迪,陈勇,等.混沌密码学原理及其应用[M].北京:科学出版社,2009.
- [11] LI SHUJUN, MOU XUANQIN, CAI YUANLONG. Pseudo-random bit generator based on couple chaotic systems and its applications in stream-cipher cryptography[C]// Proceedings of the Second International Conference on Cryptology in India: Progress in Cryptology. London: Springer-Verlag, 2001, 2247: 316-329.
- [12] The USC-SIPI Image Database[EB/OL]. [2012-05-01]. <http://sipi.usc.edu/database/>.
- [13] CHEN GUANRONG, MAO YAOBIN, CHUI C. A symmetric image encryption scheme based on 3D chaotic cat maps[J]. Chaos, Solitons and Fractals, 2004, 21(3): 749-761.

(上接第3455页)

- [3] MARKUS J, SUSANNE W. Security weaknesses in Bluetooth[C]// Proceedings of the 2001 Conference on Topics in Cryptology. London: Springer-Verlag, 2001: 179-191.
- [4] SINGEL D, PRENEEL B. Security overview of Bluetooth[EB/OL]. [2012-05-20]. <http://www.cosic.esat.kuleuven.be/publications/article-565.pdf>.
- [5] YANIV S, AVISHAI W. Cracking the Bluetooth PIN[C]// Proceedings of the 3rd International Conference on Mobile Systems, Applications, and Services. New York: ACM, 2005: 39-50.
- [6] 郭锋,庄奕琪.蓝牙E0加密算法安全分析[J].电子科技大学学报,2006,35(2): 160-163.
- [7] 王利涛,郁滨.蓝牙用户认证方案的设计与仿真实现[J].计算机工程与设计,2008,29(14): 3607-3609.
- [8] YU BIN, LI HAIYAN. Design and implementation of one key agreement scheme in Bluetooth[C]// Proceedings of the 2008 International Conference on Computer Science and Software Engineering. Washington, DC: IEEE Computer Society, 2008: 665-668.
- [9] WONG F L, STAJANO F, CLULOW J. Repairing the Bluetooth pairing protocol[C]// Proceedings of the 13th International Conference on Security Protocols. Berlin: Springer-Verlag, 2005: 1-17.
- [10] LEE G, PARK S. Bluetooth security implementation based on software oriented hardware-software partition[C]// Proceedings of 2005 IEEE International Conference on Communications. New York: IEEE, 2005: 2070-2074.
- [11] ZHANG SUN, LIU LIANDONG, YU BIN. Integrality authentication scheme of Bluetooth baseband packet header based on key stream[C]// 2010 The 3rd International Conference on Computational Intelligence and Industrial Application. Berlin: Springer, 2010: 113-121.
- [12] 郁滨,黄一才.基于蓝牙单芯片的密码算法实现方案研究[C]// 全国第20届计算机技术与应用学术会议论文集.合肥:中国科学技术大学出版社,2009: 112-116.
- [13] Radio Cambridge Silicon. BlueCore5-multimedia external[EB/OL]. [2012-05-20]. <http://www.csr.com>.