第 59 卷 第 10 期 2010 年 10 月
1000-3290/2010/59(10)/6744-08

物 理 学 报
ACTA PHYSICA SINICA

Vol. 59, No. 10, October 2010

基于 Markov 性质的一阶安全算术编码及应用*

段黎力† 廖晓峰 向 涛

(重庆大学计算机学院,重庆 400044)

(2009 年 11 月 16 日收到; 2009 年 12 月 24 日收到修改稿)

基于压缩编码的加密方式能够同时完成加密和压缩的功能,通过压缩减少了信息的冗余,同时引入加密使对试图推测出明文信息和找到密钥的攻击具有非常好的鲁棒性. 本文提出了一种基于一阶 Markov 模型的安全算术编码,在编码过程中通过随机密钥保证图像压缩编码的安全性,且显著提高编码的压缩效率,使其方便在网络中安全传输. 实验结果和安全性分析表明,算法能够抵抗现有的各种基于算术编码的安全性攻击和其他密码学分析.

关键词: 算术编码, 图像加密, Markov 模型, 动态模型

PACC: 0250, 2841C

1. 引 言

早期简单的编码压缩算法仅仅能将文件压缩到原始大小的 75% 左右,而更多当代的系统压缩算法能轻松的将其压缩到原始大小的 50% 左右,更好的算法甚至能达到原文件的 20% 左右[1—3]. 而大家都公认 Huffman 编码为最优编码,且有其不能被改进提高的误解. 事实上,Huffman 编码只有在所有概率都为 2 的幂次方的情况下才能达到最优,而概率趋于一致分布的情况下,编码效果并不理想. 在 20 世纪 60 年代 Elias 就提出了算术编码的思想,1987 年 Witten 等[1]提出了算术编码在数据压缩方面的应用,指出其比 Huffman 编码具有更好的压缩效率,它能够在不要求概率分布形式的情况下表现出良好的性质.

包括 Huffman 编码在内的早期编码模式都是采用固定的码字来表示每一个需要编码的符号,而从加密的角度来看,这些算法都是使用简单的字母替换,即用一个符号或字符串替换另一个符号或字符串,所以都很容易被破解,不能提供真正意义上的数据安全. 算术编码并不是采用固定码字来表示每个符号,它的压缩模式是将一段消息用一个[0,1)的真子集(子区间)来表示,而这个区间被初始化为[0,1),并且每编码一个符号区间就缩小一次. 使每一个新区间都能唯一地表示一段消息. 它可以根据所使用的模型,保证用一段无限逼近信息熵的比特数来传输消息. 算术编码的模型指出每个符号的概率值,分为静态模型和动态模型两种. 静态模型是根据大量的实验数据预先给出每个符号的先验概率,而动态模型里符号的概率值则是根据当前已编码的字符动态改变的,以更加真实地逼近消息中符号的概率分布. 如果我们采用静态模型对每个符号进行编码,最好状态可以达到每个字符编码后的码字长度为 $-\mathrm{lb}P_i$(P_i 为第 i 个字符的概率),当且仅当输入符号的概率分布与预测模型的概率分布一致,而通常情况下不同的消息符号的概率分布并不相同,所以静态模型并不能真实的反映每段消息的概率分布而不能达到最佳的压缩效果,相反,动态模型就能更真实地反映出当前编码的消息中符合的概率分布,这将更有利于数据的压缩.

由于算术编码的良好压缩性能,近年来,算术编码被广泛应用到一些最新的多媒体压缩标准中,例如 JPEG2000, H. 264 等. 正是它在多媒体压缩中的广泛应用使其安全性引起了人们的重视,这就使基于压缩编码的加密算法得到了广泛研究[4—17]. 在

* 国家自然科学基金(批准号: 60973114)、输配电装备及系统安全与新技术国家重点实验室自主研究项目(批准号: 2007DA10512709207)、重庆市自然科学基金(批准号: 2008BB2193, 2009BA2024)、中央高校基本科研业务费(批准号: CDJZR10180020)、中国博士后科学基金(批准号:20100470817)和重庆市重点自然科学基金(批准号:2009BA2024)资助的课题.

† E-mail: wing1687@ yahoo. cn

20 世纪 80 年代末，业界还一致认为[3,4]不管是静态模型和动态模型的算术编码都能提供很好的安全性，但是很快便发现单纯的算术编码并不能够提供可靠的安全性能，文献[5]中提出了对动态模型的算术编码的选择明文攻击，文献[6]中提出了另一种对静态模型的已知明文攻击和选择明文攻击. 文献[18]中给出了一个用二值算术码实现数据加密的算法 ZNJ，但很快文献[19,20]就提出了对 ZNJ 算法的改进和破解方法，其中文献[18]还针对原文献中指出该方法可以抵抗已知明文攻击，提出了对它的已知明文攻击方法. 另外，Kim 等[10]提出了一种基于对连续区间进行划分的安全算术编码也在最近的研究[11]中被攻破. 目前还比较安全的是 Grangetto 等[8]提出的 RAC 算法，该算法实际采用的是流密码的形式，它的安全性完全依赖于 PRBG 的安全性.

但是因为现在对安全的算术编码的研究大多只涉及了算法的安全性，而未考虑到如何能更好地利用算术编码最大的优点——良好的压缩性能，所以本文提出一种压缩性能良好的安全算术编码，即基于 Markov 性质的一阶安全算术编码，并将其应用于图像压缩加密中，而这又有别于像文献[15]中提出的普通的图像加密.

2. 算术编码简介

算术编码是将一段消息表示成[0,1)之间的一个实数区间，这个区间随着每个信源符号的加入而逐步减小，每次减少的程度取决于当前加入的信源符号的概率，下面给出一个算术编码的例子. 设信源符号集由 5 个信源符号{a,e,i,o,u}组成，其出现概率和顺序如表 1 所示.

表 1　信源符号及其范围

信源符号	概率	范围
a	0.2	[0,0.2)
e	0.3	[0.2,0.5)
i	0.2	[0.5,0.7)
o	0.1	[0.7,0.8)
u	0.2	[0.8,1)

假设需要进行编码的符号流为 eaii. 在编码开始前，对应的数值范围是整个半开区间[0,1). 编码完第一个信源符号后，范围缩小到[0.2,0.5). 然后当编码符号 a 时，范围压缩到新区间的前 1/5，即区间为[0.2,0.26). 如此继续下去，编码器的输出数值范围不断缩小，最终的区间为[0.236,0.2384)，则最后编码的值可为区间[0.236,0.2384)上的任一值，设为 0.2365. 图 1 显示了整个编码的区间变化过程. 而解码过程与其正好相反，首先查看最后编码的值落入哪一个信源符号的数值范围(表 1 中的范围列)，本例所示编码值 0.2365 在区间[0.2,0.5)之间，所以第一个符号为 e；然后从编码数值中消去第一个符号 e 的影响，即减去 e 的下界值，然后除以 e 对应的宽度(概率值)，即(0.2365 - 0.2)/0.3 = 0.1216，查找这一结果落入哪个符号对应的数值范围，得到第二个符号 a. 重复上述两步直到解出整个符号流为止.

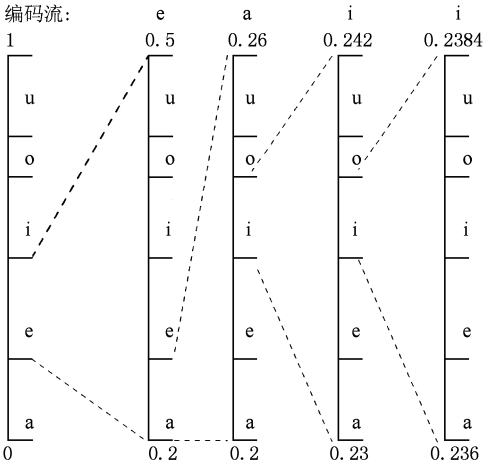

图 1　算术编码每一步区间变化过程

2.1. 算术编码的模型

在给定的上下文中，算术编码的模型是为下一个要编码的符号提供概率预测，分为静态模型和动态模型两种. 表 1 是一个 5 个符号的静态模型，它可以用于编码这个符号集中任意符号组合的消息，如果模型中的概率与消息中符号出现的概率一致便能达到最好的压缩效果. 但更多的时候，编码过程中采用的是动态模型，因为一般情况下，并不能预先知道信源符号的先验概率. 而动态模型仅表示当前已经被编码的符号概率，且是随编码符号增多动态更新的. 模型最初的时候，所有符号概率相同，可以全部设置为 1(频数). 每编码一个字符，其频数增加 1 以逼近观测到的概率. 另一方面，编码和解码的模型必须使用同样的初始值(相同的频数)和更新算法，才能使编码器和解码器达到同步；即编码器

在接收一个符号后,先编码再更新模型;而解码器也根据当前的模型进行解码,然后依据解码出来的符号更新模型.然而单一字符的概率模型使得输出依然存在着大量的统计冗余信息,所以需要更好的模型来提高压缩效率,减少冗余.这是因为最简单的模型对上下文并不敏感,每个符号都是固定的概率分布.为了使压缩效率更好,我们将在下一节讨论更加成熟的模型.

在模型中,任何真实出现的符号概率都不可能为零,否则该符号将无法进行编码.除此之外,算术编码对概率分布就没有其他的限制了.只要符号集的概率分布能真实的反映信息中符号的概率,压缩效率就能达到最好状态.

2.2. 高阶模型

提高模型的预测能力的方法是使符号的概率分布与在它之前的一些符号相联系,使概率的值对上下文敏感性增加,建立高阶的 Markov 模型,这对提高编码的压缩效率有很大的帮助.但如此一来,模型就会变得非常大:如果使用 k 个相邻的字符,模型中概率的个数将会达到 a^k,其中 a 为整个符号集中符号的个数.所以我们可以根据需要自行选择 k 的大小.但是在这里存在一个"零概率"问题,即在任何事件未发生之前都不存在理论方法估计它的先验概率[13].同理,也不存在预先的方法表明哪几个相邻字符能够给出最佳的概率预测,因此必须通过实验来解决.Cleary 和 Witten[14] 提出一种具有良好性能的模式—只用最长的邻接字符串来作概率预测,而忽略相对较短的相邻字符串的影响,这被称为"部分匹配策略",在本文提出的算法中我们采用了这种模式.

考虑到本文提出的算法主要针对二元符号进行研究,所以采用的模型为一阶模型,即二元一阶 Markov 模型.图 2 为字符串 0101100 的一阶 Markov 模型.图中输入字符串下面对应的是每个字符在模型中支持概率预测(条件概率)的最大阶数,接下来是模型中该阶符号的概率值.零阶模型就是单一的符号概率模型(二元模型即有两个概率值,符号 0 和 1 各一个值),没有条件的限制.而一阶模型中条件概率有 4 种情况,即 $P(0|1)$,$P(1|1)$,$P(1|0)$,$P(0|0)$.

利用高阶模型进行算术编码的过程与低阶模型略有不同,主要是模型是在编码过程中根据每次

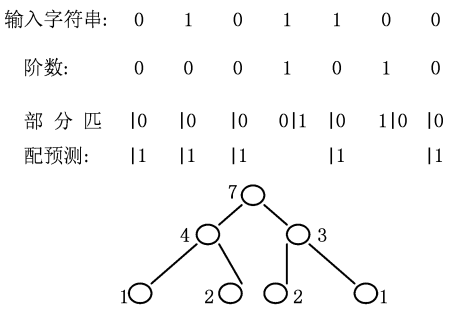

图 2 字符串 0101100 的一阶 Markov 模型

编码的符号逐步由初始模型建立的.初始模型如图 3 所示.编码过程如下,如果当前需要被编码的符号为 0,则查找模型中是否存在一阶条件概率(非初始状态),若存在则查看前一个编码的符号的具体值,假设为 1,则用对应的一阶条件概率($P(0|1)$)进行编码,并向解码器方传输特殊标志(escape symbol),指明用的是一阶概率模型使解码器方知道应该如何解码;若前一个符号为 0,则使用一阶条件概率 $P(0|0)$ 进行编码,发送特殊标志给解码器方.而如果模型中一阶模型尚未建立好,则使用符号 0 的当前的零阶概率进行编码,解码时用对应的零阶概率解码即可;另一方面,被编码的符号为 1 时过程亦如此.

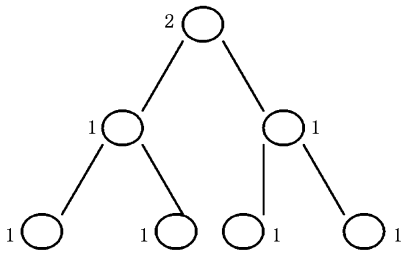

图 3 初始模型

3. Markov 信源编码性质简介

定义 3.1 若信源输出的符号序列和状态序列满足下列条件,则称此信源为 Markov 信源[16].

(1)某一时刻信源符号的输出只与当时的信源状态有关,而与以前的状态无关,即

$p(x_l = a_k | u_l = s_j, x_{l-1} = a_k, u_{l-1} = s_i, \cdots) = p(x_l = a_k | u_l = s_j)$ 其中,$a_k \in A(a_1, a_2, \cdots, a_q)$,$s_i, s_j \in s = (s_1, s_2, \cdots, s_j)$.

(2)信源状态只由当前输出符号和前一刻信源

状态唯一确定，即

$$p(u_l = s_i \mid x_l = a_k, u_{l-1} = s_j) = \begin{cases} 1, \\ 0. \end{cases}$$

类似地，可以定义 m 阶 Markov 信源.

实际中，很多信源发出的消息各符号之间都是相互关联的，而这种信源被成为有记忆信源，通常用联合概率或者条件概率来描述这种关联性，也就是有记忆信源模型，Markov 信源是一种非常重要的有记忆信源. 信源的输出符号之间依赖关系越强，其信源熵就越小，更趋于极限熵 H_∞，则熵的相对率越大，信源的剩余度越小，而这正是使压缩效果好的关键所在. 所以本文算法使用一阶 Markov 模型以提高压缩效率.

4. 采用一阶 Markov 模型的安全算术编码及在图像中的应用

RAC 算法的提出是基于算术编码区别于其他编码的特殊性，即其对压缩数据进行解码的过程对差错非常敏感，差错会在极短的时间内扩散到整个解码块中，事实上，只要在解码过程中有一个错误，就会导致随后解码的信息没有任何的意义. 这种特性正是设计一个具有良好鲁棒性的图像加密算法所需要的特性. 因为 RAC 是到目前为止最安全的算术编码，但压缩性能还不算好，本文提出一种对 RAC 的改进算法——基于一阶 Markov 模型的安全算术编码，使编码的压缩效率进一步提高并且能够保证其具有良好的安全性.

本算法是通过密钥控制编码模型是否置换，即图 2 中模型树的内部节点是否交换，以进行加密，以图 2 中的例子来说，其模型置换过程如图 4 所示.

下面给出一种模型置换（加密）的规则. 由于树状模型是从初始状态根据每次编码的符号逐步建立起来的，所以我们的置换规则也可应用于模型逐步建立的过程中，即加密过程为模型的建立过程. 加密过程分为以下两步：1）用随机数生成器产生 256 位的随机数 r_i；其中 r_i 可能的值为 0 或 1，用于控制是否需要进行模型的置换. 对不同的输入图像采用不同的种子，生成对应的随机数串. 而随机数发生器的选取可以随意，像细胞自动机做随机数发生器以及文献[16,22—23]等中提出的随机数产生器等，但并不是这里讨论的重点. 2）如果 $r_i = 1$ 且当前使用的模型阶数为 0，则交换第零阶概率模型即符号 0 和 1 在模型中的概率值. 若当前使用的模型阶数为 1，则交换对应的第一阶模型概率值，如果前一个被编码的字符为 0，则交换的是 0|0 和 1|0 的概率值；相反，前一个被编码的字符为 1，则交换的是 0|1 和 1|1 的概率值，整个过程如图 4 所示. 另一方面，如果 $r_i = 0$，则用正常的高阶模型进行编码.

我们仍以图 2 中的例子来说明加密及模型建立的过程，假设密钥为 0101100. 模型建立及加密过程如图 5 所示.

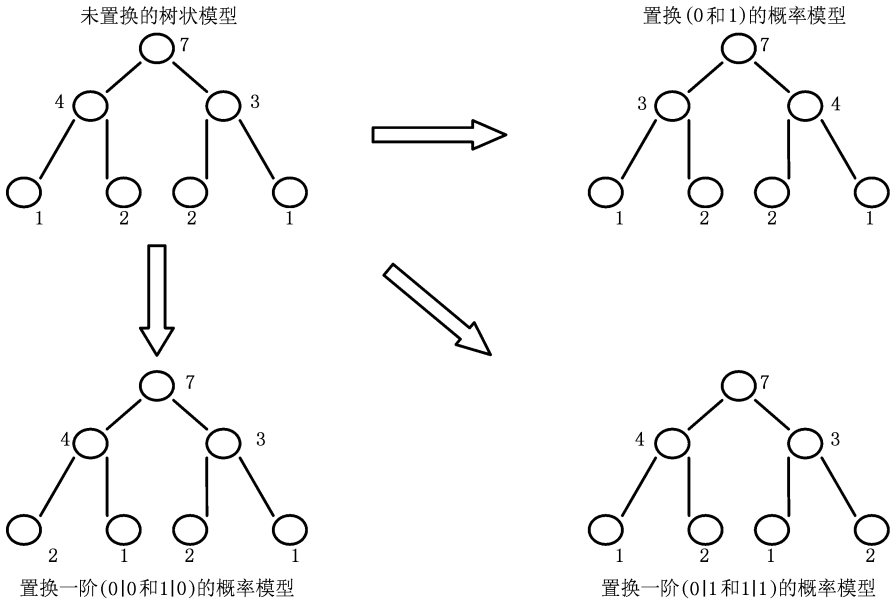

图 4　模型置换示意图

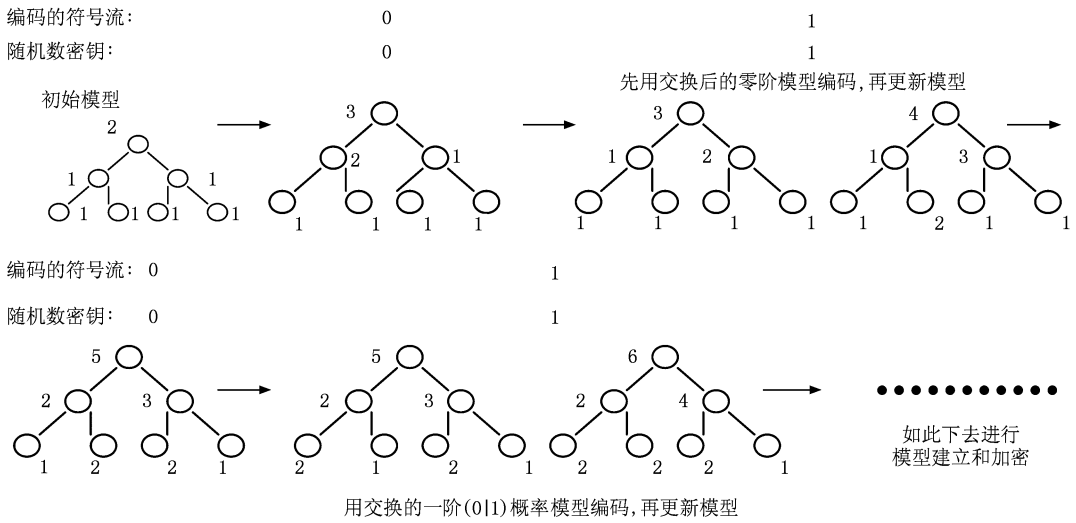

图 5 模型建立及加密过程

一阶 Markov 模型的使用使编码加密的压缩效果得到很大提高,由于我们并未违反编码过程的语法,而只是对模型进行了改变,所以解码过程与传统的算术编码解码过程相同,只是需要密钥的控制才能解密出源消息.解密过程如下:

1)用与加密端(编码器端)相同的种子生成随机数序列,即生成解密密钥.

2)如果 $r_i = 1$,检查特殊标志符若存在则说明用一阶概率模型解码,然后查看前一个解码的符号,用相应的交换后的一阶条件概率模型进行解码;否则用交换后的零阶概率模型进行解码,然后更新模型;

如果 $r_i = 0$,同样检查特殊标志符,若存在则说明用一阶概率模型解码,然后查看前一个解码的符号,用相应的一阶条件概率模型进行解码;否则用零阶概率模型进行解码,然后更新模型.

3)然后读取下一个 r_{i+1}.

可以从中看出本算法是完全可逆的.使用本算法进行图像加密时,首先对图像数据进行预处理,将十进制的像素值转换为八位二进制,即只有符号 0 和 1,再对其进行编码压缩及加密.具体实现过程如 3 小节中算法所示.解密时只需将正确解码的序列转换为十进制恢复图像数据.

5. 实验结果及安全性分析

5.1. 压缩性能分析

为了分析本算法的压缩性能,我们分别对图像,文本文件用 RAC 算法和本算法进行压缩加密,测试实例采用通用的压缩算法文件库(the canterbury corpus)中的文件,其压缩效果如表 2 所示.从中我们可以很明显的看出利用一阶 Markov 模型的算术编码比 RAC 具有更好的压缩性能,这是因为一阶 Markov 模型中符号的概率值更逼近于实际传输的消息中各符号的概率值.

5.2. 时间效率分析

为了更好的分析本算法的性能,我们针对不同的文件类型(包括图像和文本文件,其中的文本文件同上采用(the canterbury corpus)中的通用测试文件)将其运行时间与基本算术编码的运行时间进行了比较,表 3 为两者的比较结果,从中可以看出牺牲一小部分的运算时间来达到良好的压缩加密效果是非常值得的,而压缩加密性能的实验结果将随后给出.

表 2 RAC 算法和本文算法对不同文件的压缩结果

文件	类型	名称	字节数	比特数	RAC 比特数	本文算法比特数
1		Lenna	524288	4194304	786432	532480
2	图像	Cameraman	131072	1048576	200648	131136
3		Rice	294912	2359296	440912	294984
4		Lect10	2987278	23898224	4441160	2987440
5	文本	Paper2	575393	4603144	850120	575472
6		Paper3	325682	2605456	479200	325752

表 3　基本算术编码与本算法运行时间结果比较

文件	类型	名称	字节数	运行时间/s	
				原始 AC	本算法
1		Lenna	524288	0.156	0.203
2	图像	Cameraman	131072	0.031	0.046
3		Rice	294912	0.078	0.109
4		Lect10	2987278	0.671	1.063
5	本文	Paper2	575393	0.219	0.125
6		Paper3	325682	0.078	0.110

5.3. 抵抗从编码角度的攻击

因为 Bergen 等[4]提出的选择明文攻击是通过确定固定模型里符号的顺序及概率来进行的，但本算法由密钥控制模型中符号概率的顺序使其顺序并不固定，所以整个算法的安全性及抵抗此种攻击的关键在于密钥的安全性. 而密钥的安全性完全由 PRBG 来控制. 因此，本文提出的算法能够抵抗 Bergen 提出的选择明文攻击.

而 John 等[6]提出的攻击只针对二元固定模型的编码，通过选择适当的明文推测出模型中 A，B 的概率，并且在假设 A，B 顺序已知的条件下进行的. 而本文算法采用动态高阶模型进行编码，就使得文献[6]中提出的攻击方法由于无法确定每次编码时模型中符号的阶数及顺序而失效.

5.4. 密文均衡性测试分析

为了测试密文中符号的均衡性(0，1 分布的均衡性)，我们将产生的随机数值的种子进行改变，测试了密文中 0，1 的个数比. 测试结果如表 1 所示. 从表 4 可以看出密文的均衡性在 0.5 左右，即算法有很好的加密效果.

表 4　0，1 均衡性测试结果

不同种子	1	2	3	4	5
密文中 0/1 个数比	0.4927	0.5052	0.4813	0.5000	0.5043

5.5. 密钥敏感性分析

设计加密算法的一个衡量标准就是雪崩效应准则. 若改变密钥中的任意 1 位，将导致密文分组中几乎所有数据位的变化[12]，即密钥的雪崩效应. 实验对仅改变密钥任意 1 位的密文进行分析，整个数据的改变率如表 5 所示.

表 5　密钥改变 1 位加密数据变化率

密钥变化 1 位	加密数据变化率/%
$K1$	99.6140
$K2$	99.6185
$K3$	99.6277
$K1$，$K2$ 和 $K3$	99.6201

对算法密钥的敏感性测试结果如图 6 所示，我们可以从图 6(a)看到未进行压缩加密的 Lena 图像，图 6(b)显示了 Lena 图像经加密压缩后，由没有密钥的标准解码器解码后的效果，仅改变其中 1 位密钥的解码器解码之后的效果以及用正确密钥解码之后的图像分别在图 6(c)和图 6(d)中可以看到，在以上两种情况下解密后的图像都是没有任何意义的.

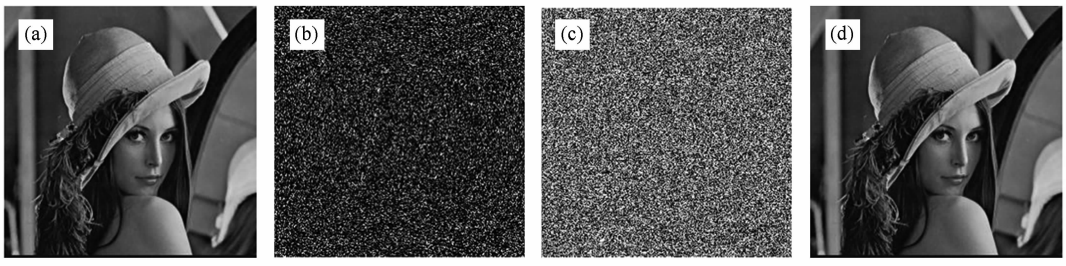

图 6　算法的密钥敏感性测试　(a) 未压缩编码的 Lena 图像，(b) 没有密钥解码之后的图像，(c) 改变 1 位密钥解码之后的图像，(d) 正确密钥解密后图像

5.6. 统计分析

PSNR 是广泛使用的评鉴画质的客观量测法，PSNR 值越大，就代表失真越少，而 PSNR 值越小，失真就越大，其值小于 9 的时候图像是毫无意义的[8]. 图 7 为不同的随机数种子产生的密钥(不同于编码方的加密密钥)解码后图像的 PSNR 值. 从图中可以看出图像经过加密后，如果被非法用户截

获,而使用与加密方不同的密钥进行解码时,得到的图像几乎是没有任何意义的.

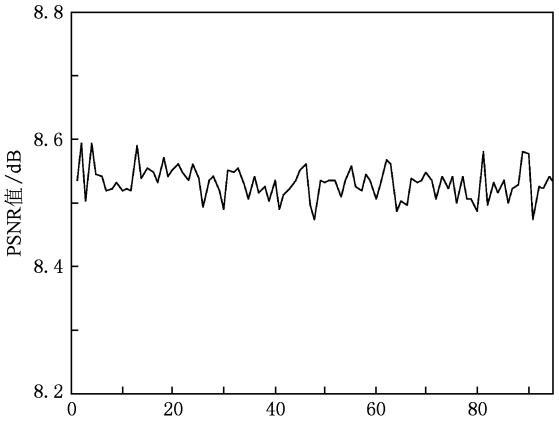

图 7 根据不同种子产生密钥得到解密图像 PSNR 值

5.7. 明文敏感性分析

对于明文图像来说,如果图像像素中有一个细微的变换可以导致在密文图像中有重大的差别,则对于图像来说,差分攻击就是无效的. 于是为了测试改变明文图像中一个像素值的改变对密文带来的影响,我们利用文献[12]中提出的 NPCR 作为测试的标准. 由于测试对象的不同,我们将 NPCR 进行重新定义,首先给出一个像素值不同的明文图像所对应的两个经压缩加密之后的密文 C_1, C_2, 位置 i 处对应的密文记为 $C_1(i)$, $C_2(i)$, 定义数组 $D(i)$, 其值由 $C_1(i)$, $C_2(i)$ 决定,如果 $C_1(i) = C_2(i)$, $D(i) = 0$, 否则 $D(i) = 1$. NPCR 值定义为

$$\sigma_{\mathrm{NPCR}} = \frac{\sum_i D(i)}{size(D)} \times 100\%$$

对作用于大小为 256 × 256 的灰度图像的算法进行了测试,结果如图 8,NPCR 值在 0.996 附近波动,说明算法对明文的敏感性非常好.

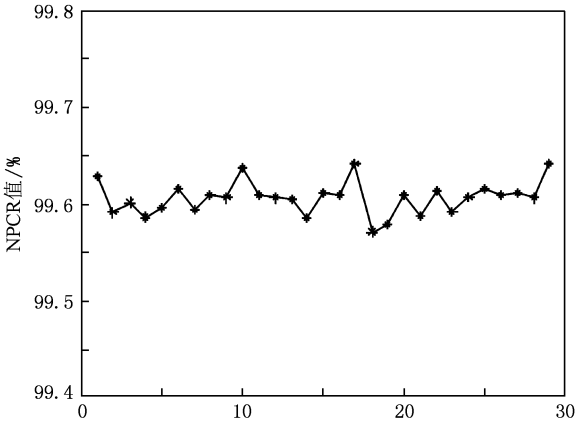

图 8 图像改变一个像素值的 NPCR 值

6. 结 论

本文提出了一种基于一阶 Markov 模型的安全的算术编码,并将其应用到对图像的加密中. 该算法利用高阶模型能够极大地提高编码压缩效率的优点,结合算术编码对模型的敏感性,通过密钥控制编码过程中模型的阶数,及符号概率的顺序对输入符号进行压缩加密. 实验和安全性分析结果表明:该加密算法具有比 RAC 更好的压缩性能,且安全性非常好,有很大的密钥空间,对密钥十分敏感,从密码学和编码学分析的角度,对多种攻击手段都具有良好的免疫性,适用于软件加密系统,而且很容易移植到硬件平台上,具有良好的应用前景.

[1] Witten I H, Neal R M, Cleary J G 1987 *Comput. Pract.* **30** 520
[2] Langdon G G Jr 1984 *IBM J. RES. Develop.* **28** 135
[3] Witten I H, Cleary J G 1988 *Comput. Secur.* **7** 397
[4] Bergen H A, Hogan J M 1992 *Comput. Secur.* **11** 445
[5] Bergen H A, Hogan J M 1993 *Comput. Secur.* **12** 157
[6] Cleary J G, Irvine S A, Rinsma-Melchert I 1995 *Comput. Secur.* **14** 167
[7] Liu X, Farrell P G, Boyd C A 1997 *Cryptography. Cod.* **1355** 199
[8] Grangetto M, Magli E, Olmo G 2006 *IEEE Trans. Multimed.* **8** 905
[9] Wen J T, Kim H J, Villasenor J D 2006 *IEEE Signal Process. Lett.* **13** 69
[10] Kim H J, Wen J T, Villasenor J D 2007 *IEEE Trans. Signal Process.* **55** 2263
[11] Jakimoski G, Subbalakshmi K P 2008 *IEEE Trans. Multimed.* **10** 330
[12] Pareek N K, Patidar V, Sud K K 2006 *Imag. Vision Comput.* **24** 926
[13] Good L J 1965 *The Estimation of Probabilities* (Cambridge: Massachusetts Institute of Technology Press)
[14] Cleary J G, Witten I H 1984 *IEEE Trans. Commun.* **32** 396

[15] Wang J Z, Xu S J, Yang S X 2008 *Chin. Phys.* B **17** 4027
[16] Liu J S, Liu S B, Sun J, Xu Z Q 2009 *Chin. Phys.* B **18** 5219
[17] Li Z C 2006 *Information Theory and Coding* (Xuzhou: China University of Mining Press) p95 (in Chinese) [李子臣 2006 信息论与编码（徐州：中国矿业大学出版社）第 95 页]
[18] Zhao F G, Ni X F, Jiang F 1999 *J. Commun.* **20** 92 (in Chinese) [赵风光、倪兴芳、姜　峰 1999 通信学报 **20** 92]
[19] Xie D Q, Xie Z J, Li C, Leng J 2001 *J. Commun.* **22** 41 (in Chinese) [谢冬青、谢志坚、李　超、冷　健 2001 通信学报 **22** 41]
[20] Zheng H R, Jin C H 2003 *J. Commun.* **24** 73 (in Chinese) [郑浩然、金晨辉 2003 通信学报 **24** 73]
[21] Gao L J, Yang X P, Li Z L, Wang X L, Zhai H C, Wang M W 2009 *Acta Phys. Sin.* **58** 1053 (in Chinese) [高丽娟、杨晓苹、李智磊、王晓雷、翟宏琛、王明伟 2009 物理学报 **58** 1053]
[22] Zhou Q, Hu Y, Liao X F 2008 *Acta Phys. Sin.* **57** 5413 (in Chinese) [周　庆、胡　月、廖晓峰 2008 物理学报 **57** 5413]
[23] Wang F P, Wang L, Wang Z J 2006 *Acta Phys. Sin.* **55** 3964 (in Chinese) [汪芙平、王　蕾、王赞基 2006 物理学报 **55** 3964]

Image encryption based on arithmetic coding with order-1 Markov model*

Duan Li-Li†　Liao Xiao-Feng　Xiang Tao

(*Department of Computer Science, Chongqing University, Chongqing* 400044, *China*)

(Received 16 November 2009; revised manuscript received 24 December 2009)

Abstract

Encryption based on compression can provide compression and encryption in a single step, in which redundancy is removed by compression and security is guaranteed by encryption. A novel secure arithmetic coding scheme based on order-1 Markov model is proposed in this paper, and it is applied to image encryption where images can be transmitted securely on the Internet and the proposed algorithm significantly improve the compression efficiency of coding. Experimental results and security analyses indicate that, the algorithm can not only resist existing attacks based on arithmetic coding, but also be immune to other cryptanalysis.

Keywords: arithmetic coding, image encryption, markov model, adaptive-model
PACC: 0250, 2841C

* Project supported by the National Natural Science Foundation of China (Grant No. 60973114), the Independent Research Projects of Transmission and Distribution Equipment and System Security with the New State Key Laboratory (Grant No. 2007DA10512709207), the Natural Science Foundation of Chongqing, China (Grant Nos. 2008BB2193, 2009BA2024), the Fundamental Research Funds for the Central Universities (Grant No. CDJZR10180020), the China Post-Doctoral Science Foundation (Grant No. 20100470817) and the Key Natural Science Foundation of Chongqing, China (Grant No. 2009BA2024).

† E-mail: wing1687@yahoo.cn