基于伪随机序列的 Arnold 加密算法

徐光宪 吴 巍
(辽宁工程技术大学电子与信息工程学院 葫芦岛 125105)

摘 要 Arnold 变换是一种经典的图像置乱算法。由于其具有周期性,导致密钥量不够。提出了一种改进的 Arnold 变换方法,即引入伪随机序列并利用安全哈希算法产生随机参数序列,将数字图像分块,并对每个块图像分别进行 Arnold 变换,最终得到一幅置乱图像。该算法有效地增加了密钥量,可以克服通过穷举分析等手段进行的攻击,增加了图像的安全性。

关键词 图像加密,伪随机序列,SHA1,Arnold Cat 变换,矩阵变换
中图法分类号 TP 309.7 **文献标识码** A

Arnold Encryption Algorithm Based on PN Sequence

XU Guang-xian WU Wei
(School of Electronic Information Engineering, Liaoning Technical University, Huludao125105, China)

Abstract Arnold Cat transformation is a classical algorithm of image scrambling. But its periodicity restrains the times of image scrambling, so the number of the keys is not enough. This paper presented a modified Arnold method, which uses PN sequence and secure hash algorithm to generate a random parameter sequence, then divides an image into 4 pieces and carries out Arnold algorithm to the 4 pieces respectively. The Arnold algorithm effectively increases the number of the keys, and overcomes the attack via the exhaustive analysis, which enhances the security.

Keywords Image encryption, PN sequence, SHA1, Arnold Cat transformation, Matrix transformation

随着数字化技术和 Internet 的飞速发展,其中以数字图像为载体的网络多媒体技术应用逐渐普及,在最大限度地方便人类信息交流的同时,也带来了被窃取的危机。所以如何在网络环境中实施有效的隐私保护和信息安全手段成为一个迫在眉睫的现实问题,而数字图像的加密技术则是解决该问题的重要手段之一。

图像位置置乱是一种常用的数字图像加密方法。经典的图像加密算法有 Arnold Cat 变换、Hilbert 变换、Zigzag 变换、Baker 变换等。Arnold Cat 变换以其计算简单、具有周期性、处理时间短、置乱效果好的优点被广泛应用。但其密钥量小、视觉效果还不够理想,而 Arnold 变换是通过改变图像中各像素点的位置来实现图像加密的目的,这样攻击者就可以通过统计分析等手段进行攻击。本文提出了一种新颖的 Arnold 变换,其将伪随机码作种子序列,用安全哈希算法对该种子序列进行处理产生随机参数序列,将原始图像分块并应用 Arnold 变换。

1 Arnold Cat 变换

Arnold Cat 变换是在遍历理论研究中提出的一种变换,俗称猫脸变换。本意为 cat mapping。设想正方形图像大小为 $N\times N$,其内有一点 (x,y),将点 (x,y) 变换到另一点 (x',y') 用公式表示为:

$$\begin{bmatrix}x'\\y'\end{bmatrix}=\begin{bmatrix}1&1\\1&2\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}(\mathrm{mod}\ N) \tag{1}$$

$x,y\in(0,1,2,\cdots,N-1)$,式(1)称为 Arnold 变换。Arnold 变换实际上是一个点的位置移动过程。

对于一幅大小为 $N\times N$ 的图像,经过若干次迭代后可得到一幅置乱的图像,但是 Arnold 变换具有周期性,迭代到某一步时,将重新恢复出原始图像。如一幅 128×128 的图像的 Arnold 变换周期为 96;256×256 的图像的 Arnold 变换周期为 192;512×512 的图像的 Arnold 变换周期为 384。迭代次数与图像大小呈非线性关系。

2 分块置乱算法

2.1 图像的预处理

数字图像以数据矩阵的形式进行存储。矩阵元素所在的行与列对应图像上各像素的坐标,元素的数值就是像素的灰度值或 RGB 值。因此可以应用矩阵的初等变换对数字图像进行操作,如对矩阵的行(列)进行交换、倍乘、排列组合等变换以及矩阵运算。将明文图像进行预处理,发送方用选取的密码矩阵 $T_1,T_2,\cdots,T_n$,对原始图像矩阵进行变换。

矩阵 T 的选取:设原始图像为 $m\times m$ 的一个矩阵 P,对于

到稿日期:2012-02-10 返修日期:2012-07-24 本文受辽宁高等学校杰出青年学者成长计划项目(LJQ2012029)资助。
徐光宪(1977-),男,博士,副教授,主要研究方向为网络编码与信息处理;E-mail:flybirdxgx@sohu.com;吴 巍(1988-),女,主要研究方向为网络编码。

任意矩阵 T

$$T=\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ \vdots & x_{ij} & \vdots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mm} \end{pmatrix} \tag{2}$$

式中，m 为图像阶数。$x_{ij}\in S(i=1,2,\cdots,m,j=1,2,\cdots,m)$，$S$ 表示任何非空集合。其逆过程表示为

$$T^{-1}=\begin{pmatrix} y_{11} & y_{12} & \cdots & y_{1m} \\ \vdots & x_{ij} & \vdots & \vdots \\ y_{m1} & y_{m2} & \cdots & y_{mm} \end{pmatrix} \tag{3}$$

式中，$(|T|\neq 0)$，$y_{ij}\in S(i=1,2,\cdots,m;j=1,2,\cdots,m)$，$S$ 表示任何非空集合。

预处理过程描述为：对于任意图像矩阵 P，有

$$Q=T_n(T_{n-1}(\cdots(T_1(P)))\cdots) \tag{4}$$

式中，Q 为 n 个连续不同排列 $T_1,T_2,\cdots,T_n$ 组成的密码矩阵。在反序列中用 $T_i(i=1,2,\cdots,n)$ 的反操作来恢复矩阵 P，即

$$P=T_1^{-1}(T_2^{-1}(\cdots(T_n^{-1}(Q)))\cdots) \tag{5}$$

在接收端要恢复出原始图像，接收到的排列顺序应该与发送端的排列顺序一模一样。

2.2 伪随机序列的产生

伪随机序列具有类似于噪声的一些统计特性，同时又便于重复产生和处理。n 阶移位寄存器的反馈函数表示为

$$F(X_1,X_2,\cdots,X_n)=\sum_{i=1}^{n}C_iX^i \tag{6}$$

$C_i\in(0,1)$，当 $C_i=0$ 时，表示反馈端断开；$C_i=1$ 时，表示反馈端连接。可见 C_i 的取值决定了反馈的形式和输出序列结构，而特征多项式决定着 C_i 的取值。

$$f(x)=C_0+C_1x+C_2x^2+\cdots+C_nx^n=\sum_{i=0}^{n}C_iX^i \tag{7}$$

该方程为特征多项式，其中 $C_0=C_n=1$。理论研究表明，若反馈移位寄存器的特征多项式为本原多项式，则输出序列为 m 序列。

本文利用 PN 码中典型的 m 序列（最大长度线性反馈移位寄存器）产生器来产生伪随机序列。一方面，产生的序列具有随机序列的所有特性，而且可以预先确定并实现任意长度。因此可以通过伪随机序列产生器的固定参数来精确地重复产生。另一方面，改变初始状态可以产生一个全新的伪随机序列。

在加密端设定加密的初始值 O 和本原多项式 F，即可产生一个伪随机序列，记为 S。如一个 4 级线性反馈移位寄存器的初始状态为 $O=(1,0,0,0)$，将其作用于本原多项式 $f(x)=x4+x+1$，则可产生周期为 15 的伪随机序列 000111101011001…。代码如下：

```
Function [mseq]=msequence(originalstate)
n=length(originalstate);
L=2^n-1;%所需的移位寄存器的长度
register=[zeros(1,n-1),1];%初始化移位寄存器的初始状态
  mseq=register(1);
  for i=2:L
    newregister(1:n-1)=register(2:n);
    newregister(n)=mod(sum(originalstate. * register),2);
    register=newregister;
    mseq(i)=register(1);
end
```

除了全零状态之外，初值可以随意设定，只要将初值和本原多项式信息传给接收方即可实现解密。伪随机序列的这种类随机噪声特性，使得攻击者在不知道加密初值和本原多项式的时候很难破解加密图像。

2.3 随机参数序列的产生

安全哈希算法（Secure Hash Algorithm）即 SHA 算法的修订版本为 SHA1。利用上一步产生的伪随机序列源 S 作为 SHA1 算法的输入，然后将其转换成一段密文，即为本算法的随机参数序列 U。SHA1 算法的安全性在于其产生散列值的操作过程具有较强的单向性。如果在输入序列中加入密码，那么任何人在不知道密码的情况下都不能产生正确的散列值。上一步产生的伪随机序列源 S 相当于 SHA1 算法中加密的输入报文，从而保证了安全性。

SHA1 算法的基本迭代只能处理 512bit 的数据块。为了能够处理任意长的数据，需要对输入的伪随机序列 S 进行补位和补长度，将输入的序列 S 分块为 512bit 的倍数。然后再对序列 S 进行 SHA1 算法处理，所有序列块处理完后得到一个 160bit（20 字节）的数据，这个数据就是随机参数序列 U。

2.4 基于分块的 Arnold 变换

将上一步得到的 20 字节随机参数序列 U 的前 10 字节序列设为 $U1$，后 10 字节序列设为 $U2$。然后按下述几步进行处理。

（1）利用该序列 $U1$ 和 $U2$ 计算得到 $M1$ 和 $M2$。$M1=U1 \bmod N$，$M2=U2 \bmod N$。其中 N 是待置乱的图像的阶数。

检测 $M1$ 和 $M2$ 是否介于 $[2,N-2]$ 之间。如果在其间，则说明满足分块要求，继续执行。如果不在其间，则重复生成种子序列 S，直到得到符合要求的 $M1$、$M2$。

（2）令 $N1=\max(M1,M2)$、$N2=\max((N-M1),M2)$、$N3=\max(M1,(N-M2))$、$N4=\max((N-M1),(N-M2))$，分别求以 $N1$、$N2$、$N3$、$N4$ 为阶的图像 Arnold 变换周期，记为 $T1$、$T2$、$T3$、$T4$。检测 $T1$、$T2$、$T3$、$T4$ 是否大于 6。如果大于 6，则说明图像有较长的周期，满足变换要求，继续执行下一步；否则回到第一步重新生成 S 和 $M1$、$M2$。直到生成符合条件的 $T1$、$T2$、$T3$、$T4$。

（3）随机生成正整数 $B1$、$B2$、$B3$、$B4$，作为子图像的迭代次数。分别检测 $B1$、$B2$、$B3$、$B4$ 是否在 $[3,T1-3]$、$[3,T2-3]$、$[3,T3-3]$、$[3,T4-3]$ 区间内。如果在，则说明子图像可以获得较好的置乱效果，继续进行下一步；否则重新生成随机正整数 $B1$、$B2$、$B3$、$B4$，直到生成满足条件的 $B1$、$B2$、$B3$、$B4$。

（4）对划分的 4 个子图像依次分别进行置乱。

解密过程只要以相反的步骤（4）、（3）、（2）、（1）进行操作，就可恢复预处理时的图像。

2.5 图像加密流程

根据以上分析，得出图像加密流程如图 1 所示。

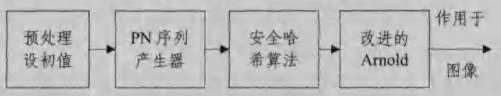

图 1　图像加密流程

3　仿真验证与分析

3.1　算法的仿真验证

为了验证改进的 Arnold 算法的有效性，针对不同图像采用不同参数做了大量的实验。本文以大小为 256×256 的 tulip 图片为例，应用此算法对图像进行加密处理，得到的结果如图 2 所示。

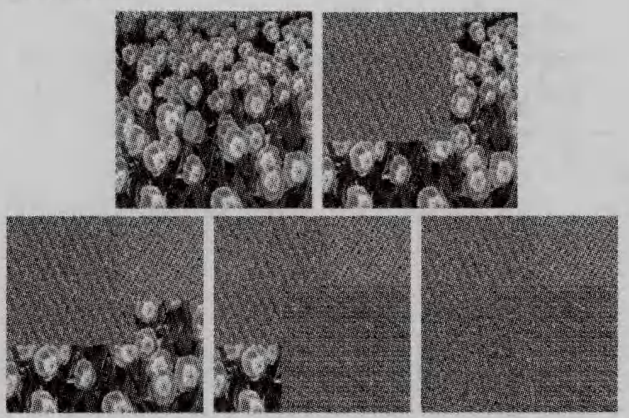

图 2　改进型 Arnold 加密过程中的图像

为了与 Arnold Cat 变换进行对比，对大小为 256×256 的 tulip 图像进行 2、6、191、192 次经典 Arnold 变换，得到的置乱效果如图 3 所示。对于阶数为 256 的图像，其恢复周期为 192。从图 3 可以看出，一般迭代 6 次就有很好的置乱效果。但正是由于 Arnold 变换的周期性，当迭代了 192 次后就完全恢复出原始图像。

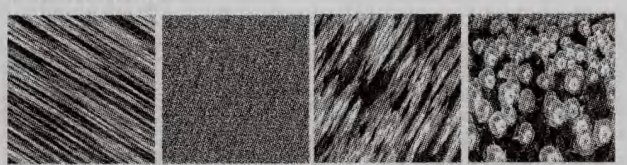

图 3　tulip 图像分别进行 2、6、191、192 次得到的置乱图像

3.2　算法分析

(1)穷举分析

本算法的安全性在于提高了 Arnold Cat 算法的密钥量，其中需要确定的值为加密矩阵 T、种子序列 S、各分块子图的迭代次数 $B1$、$B2$、$B3$、$B4$。虽然 $B1$、$B2$、$B3$、$B4$ 的大小与各子图的阶数有关，但是种子序列 S 的类随机性和未知的加密矩阵，使得破解相当困难。对于一幅 Arnold 变换置乱的图像，最多迭代 $N^2/2$ 次，才能恢复出原始图像。应用本文提出的算法对置乱的图像应用穷举法进行分析，计算得到最大穷举次数为 $T=n!\times(N^2\times2\sum_{0}^{N}(x^2/2+(N-x)^2/2))=n!\times N^4$，由此可以看出，对于大小为 256×256 的 tulip 图像，T 的数量级可达到 $n!\times256^4$，可见较文献[6]很有强的抗穷举攻击。即使在图像的阶数很小的情况下，也可有很强的抗穷举性。

(2)时间复杂分析

只要知道置乱的方法和所采用的参数，基于置乱变换的时间复杂度一般不超过 $O(N^2)$ 量级。本算法是将图像分为 4 块，每块子图像的时间复杂度分别为 $O(N1^2)$、$O(N2^2)$、$O(N3^2)$、$O(N4^2)$，其中 $N1$、$N2$、$N3$、$N4$ 为各子图像的阶数。所以总的时间复杂度上限为 $O(N1^2)+O(N2^2)+O(N3^2)+O(N4^2)\leqslant O(N1+N2+N3+N4)^2<O((4N)^2)=O(N^2)$。该算法的时间复杂度小于 $O(N^2)$。

(3)相邻像素相关性分析

对大小为 256×256 的 tulip 原始图像、对该图像进行整体 Arnold 置乱后的图像和应用本文改进的 Arnold 算法置乱的 tulip 图像分别对应抽取水平和垂直方向相邻像素对各 1000 对，应用式(8)－式(11)计算其水平和垂直方向的相邻像素的相关系数，结果如表 1 所列。可以看出，应用本文改进的加密算法得到的垂直和水平相邻素相关系数都相对较小，表明置乱效果程度较高。

$$E(x)=\frac{1}{N}\sum_{1}^{N}x_i \tag{8}$$

$$D(x)=\frac{1}{N}\sum_{1}^{N}[x_i-E(x)]^2 \tag{9}$$

$$\mathrm{Cov}(x,y)=\frac{1}{N}\sum_{1}^{N}[x_i-E(x)][y_i-E(y)] \tag{10}$$

$$r_{xy}=\mathrm{Cov}(x,y)/\sqrt{D(x)}\sqrt{D(y)} \tag{11}$$

式中，x 和 y 分别表示相邻两个像素的像素值，r_{xy} 是相邻两像素的相关系数。像素越接近 1，表明相邻两像素的相关性越高。加密后的图像像素相关性接近零，所以降低了相邻两像素的相关性。

表 1　相邻像素相关性比较

像素关系	垂直方向	水平方向
256×256tulip 原始图像	0.956	0.930
整体 Arnold 置乱后的图像	0.060	0.055
基于划分的 Arnold 置乱图像	0.051	0.046
本文改进的 Arnold 算法置乱图像	0.032	0.020

(4)密钥敏感性分析

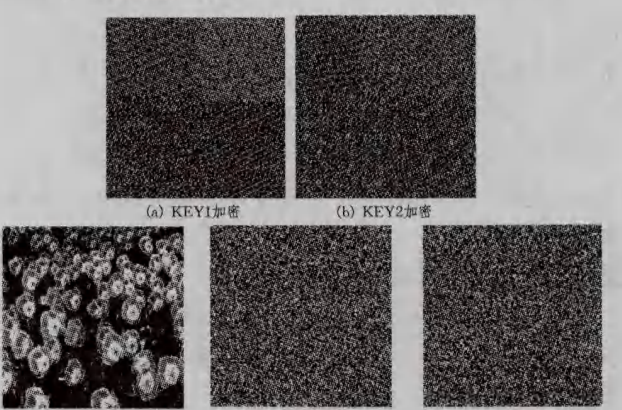

图 4　敏感性分析结果

算法的密钥参数为 Q、S、$B1$、$B2$、$B3$、$B4$，对 256×256 大小 tulip 图像加密。设 3 组密钥 KEY1：S=341769205802、$B1$=35、$B2$=17、$B3$=24、$B4$=40；KEY2：S=341769205803、$B1$

=35、$B2$=17、$B3$=24、$B4$=40；KEY3：S=341769205802、$B1$=35、$B2$=17、$B3$=24、$B4$=39，密码矩阵 $T_1, T_2, \cdots, T_n$ 由算法随机选取。图 4 中，图 4(a)为应用 KEY1 加密的结果；图 4(b)为应用 KEY2 加密的结果；图 4(c)为应用 KEY1 对图 4(a)进行解密的结果；图 4(d)为应用 KEY3 对图 4(a)进行解密的结果(密钥 KEY3 与 KEY1 的差，$SUB_{KEY3-KEY1}$=1)；图 4(e)为应用 KEY1 对图 4(b)进行解密的结果(密钥 KEY2 与 KEY1 的差，$SUB_{KEY2-KEY1}$=1)。

以上结果表明，密钥的细微差别导致完全不能解密，该算法具有高度敏感性。

结束语 本文提出了一种改进 Arnold 变换算法，其充分利用矩阵的特性，结合伪随机序列和安全哈希函数，并进行了实验验证和性能分析。该算法在抗穷举攻击和时间复杂度方面比经典 Arnold Cat 变换有明显的优势，而且密钥量大、置乱效果好，具有可行性和有效性。

参 考 文 献

[1] Kumar B K S, Patil C R. JPEG image encryption using fuzzy PN sequences[J]. SIViP, 2009, 0131(6): 2-3

[2] 陈铭. 基于 Arnold 变换的图像信息伪装算法[J]. 计算机应用研究, 2006(1): 235-237

[3] 樊昌信, 曹丽娜. 通信原理(第六版)[M]. 北京：国防工业出版社, 2006: 326-338

[4] 黄谆, 白国强, 陈弘毅. 快速实现 SHA-1 算法的硬件结构[J]. 清华大学学报, 2005, 45(1): 123-129

[5] 任洪娥, 尚振伟, 张健. 一种基于 Arnold 变换的数字图像加密算法[J]. 光学技术, 2009, 35(3): 384-390

[6] 马进, 卢雷, 朱宁. 基于划分思想的 Arnold 变换算法[A]// 中国电子学会第十六届信息论学术年会论文集[C]. 北京：电子工业出版社, 2009: 90-94

[7] 陈亦欢, 严伟超. 应用 SIMD 并行技术的 SHA-1 加密算法的批量实现[J]. 重庆理工大学学报：自然科学版, 2012, 26(7): 74-80

(上接第 46 页)

果见图 15，其中，使用网卡选择，选择具体的网卡类型进行数据包捕获；可以通过混杂模式查看所有的消息捕获结果，包括消息的接收者帐号、发送者帐号、消息类型、获取时间、源 IP、源端口、目的 IP、目的端口、消息内容；也可以选择特定的即时通信软件进行查阅截获结果。通过记录查询可以查询到特定时间段截获的结果，并且能够另存为文本文件。同时通过设置过滤规则，可以设置需要匹配的有害信息关键字，以实时拦截有害信息。利用发包工具进行漏检测试，结果如表 2 所列。

表 2 漏检测试结果

每秒发包个数	共发送包数	共检测包数	漏检率
100	1456	1456	0.00%
300	4823	4823	0.00%
1000	11729	11704	0.21%
1500	20304	19396	4.47%

系统漏检率指在网络正常情况下，系统应该检测出的数据占全部通过数据的百分比。随着发包率的增加，在一定时间内底层提交给文本识别引擎的数据量相应增加，当超出文本识别引擎的最大吞吐量时，部分数据会被文本识别引擎丢弃，造成漏检率的出现，系统应工作在可接收的漏检率下。通过表 2 发现，当每秒发包率小于 300 时，系统的漏检率为 0；当每秒发包率大于 300 小于 1000 时，出现微量漏检；当每秒发包率大于 1500 时，丢包率急剧增加。所以，本系统可以完全满足中小型企业网对于监控的需求。

结束语 本文针对目前对即时通信监控的需要，通过分析 MSN 和其他几个主流的通信机制，得出了 IM 软件的文本消息传输协议。同时，针对 MSN 文本消息传输协议的特殊性，提出了一种会话关联识别来对 MSN 文本信息进行有效完整的提取。并且分析了 MSN 和雅虎通这两个即时通信软件之间相互传输文本消息的格式。在提取消息时采用正则表达式，使得匹配提取过程更为精确，并且只需通过修改相应正则表达式即可实现对不同版本即时通信软件的监控。通过实验数据证明，IMSMMS 可对中小企业进行有效监控。下一步将研究对即时通信中文件传输、视频、音频等的统一监控进行分析实现。

参 考 文 献

[1] Lu Rui, Mi Jia, Huang Bo. Design and Implementation of Instant Messenger Security Monitoring System Based on Protocol Analysis[C]// 2010 Chinese Control and Decision Conference. 2010: 4290-4293

[2] Osullivan S. Instant messaging vs. instant compromise [J]. Network Security, 2006(7): 4-6

[3] Hindocha N. Threats to Instant Messaging[EB/OL]. Symantec Security- Response WHITE PAPER, http://securityresponse.symante-c.com/avcenter/reference/threats.to.instant.messaging.pdf

[4] 付安民, 张玉清. 即时通实时监控系统的设计与实现[J]. 通信学报, 2008, 29(10): 165-172

[5] Xu Guo-tian. Design and realization of the MSN monitoring system[C]// Information Theory and Information Security (ICITIS). 2010: 178-181

[6] 胡振宇, 刘在强, 苏璞睿, 等. 基于协议分析的 IM 阻断策略及算法分析[J]. 电子学报, 2005, 33(10): 1830-1834

[7] 刘彬, 赵彩荣, 丛建刚. 即时通信协议分析与监控技术研究[J]. 计算机应用研究, 2007, 24(9): 260-265

[8] 付安民, 张玉清. 飞信即时通监控系统的设计与实现[J]. 计算机工程, 2008, 34(13): 229-231

[9] 章秩. 基于以太网的即时通信监控系统研究与实现[D]. 上海：复旦大学研究生院, 2010

[10] MSN Messenger Service 1.0 Protocol[EB/OL]. http://www.hypothetic.org/docs/msn/sitev1/index.php, 2003.9

[11] 郑有才, 郑光华, 张玉清. 即时通信息监听技术的研究与实现[J]. 计算机应用研究, 2005, 22(8): 113-116

[12] 严华, 蔡瑞英. 即时通信系统的设计与实现[J]. 计算机技术与发展, 2009, 7(19): 242-248