文章编号:1671-4229(2021)02-0023-05

基于聚合酶链置换反应的 2D-LASM 混沌文本加密算法

王夕远[1]，殷志祥[1,2*]，唐　震[1]，杨　静[1]，崔建中[3]，徐如解[1]
(1. 安徽理工大学 数学与大数据学院，安徽 淮南　232001；
2. 上海工程技术大学 数学与统计学院，上海　201620；3. 淮南联合大学 计算机系，安徽 淮南　232001)

摘　要：针对现有的基于 DNA 序列进行信息加密算法中没有涉及到 DNA 计算中化学反应这一问题,提出了基于聚合酶链置换反应的 2D-LASM 混沌映射文本加密算法. 该算法将混沌映射产生的伪随机序列和明文信息进行异或操作,然后转换成 DNA 序列;随后再将 DNA 序列进行聚合酶链置换反应得到新的 DNA 序列. 对新的 DNA 序列进行解码,得到密文. 最后对该算法的密钥空间进行分析,可知该算法具有较好的加密效果.

关键词：DNA 计算；聚合酶链置换反应；2D-LASM 混沌映射；DNA 编码
中图分类号：TP 301　　**文献标志码**：A

2D-LASM chaotic text encryption algorithm based on polymerase strand displacement reaction

WANG Xi-yuan[1], *YIN Zhi-xiang*[1,2*], *TANG Zhen*[1], *YANG Jing*[1], *CUI Jian-zhong*[3], *XU Ru-jie*[1]
(1. School of Mathematics and Big Data, Anhui University of Science and Technology, Huainan 232001, China;
2. School of Mathematics, Physics and Statistics, Shanghai University of Engineering Science, Shanghai 201620, China;
3. Department of Computer Science, Huainan Union University, Huainan 232001, China)

Abstract: Aiming at the problem that the existing information encryption algorithm based on DNA sequence does not involve chemical reactions in DNA computing, a 2D-LASM chaotic map text encryption algorithm based on polymerase chain reaction permutations was proposed. This algorithm performs an XOR operation on the pseudo-random sequences generated by chaotic mapping and the plaintext information, and then converts them into DNA sequences. Subsequently, DNA sequences were subjected to polymerase strand displacement reactions to obtain new DNA sequences. Decoding the new DNA sequences then generates the ciphertext. Finally, the key space of the algorithm was analyzed, which shows that the algorithm has a good encryption effect.

Key words: DNA computing; polymerase strand displacement reaction; 2D-LASM chaotic mapping; DNA encoding

基金项目：国家自然科学基金资助项目(61672001, 61702008, 62072296)；安徽省自然科学基金资助项目(1808085MF193)；军委科技委前沿创新计划重点资助项目子课题(18163ZT00500901)

作者简介：王夕远(1998—)，男，硕士研究生. E-mail：wxy683106@163.com
* 通信作者. E-mail：zxyin66@163.com

引文格式：王夕远，殷志祥，唐震，等. 基于聚合酶链置换反应的 2D-LASM 混沌文本加密算法[J]. 广州大学学报(自然科学版)，2021,20(2):23-27,34.

近年来由于科技的迅猛发展,传统电子计算机已经难以满足人们的需求. DNA 计算机是近些年最有独创性和出乎意料的发现之一,其具有高度的并行性、运算速度快、存储容量大、耗能低和 DNA 分子资源丰富等优点. 为了制造出 DNA 计算机,人们开始对 DNA 计算进行研究. DNA 计算的第一个例子解决了一个 7 城市哈密顿路径问题[1]. 2000 年,Head 等[2]提出了使用 DNA 质粒的一种新的计算方法,列出了潜在的优势. 通过报告计算图顶点集最大独立子集基数的 NP 完备算法题的一个实例计算,说明了新方法的有效性. 2003 年,殷志祥等[3]提出了在基于表面的 DNA 计算采用荧光标记策略,解决了简单的 0-1 规划问题. 2011 年,Qian 等[4]提出了 DNA 链置换级联的神经网络计算. 此外,DNA 计算还可以用来构建半加器、半减器、全加器、全减器[5-9]. 2019 年,Chao 等[10]在 DNA 折纸上利用 DNA 单分子导航仪求解迷宫问题. 同年,唐震等[11-12]利用了 DNA 折纸术解决了一类特殊的整数规划问题,还提出了在 DNA 折纸基底上的一种动态的与非门计算系统.

随着计算机技术的发展和运用,人们通过网络进行信息传输的频率越来越频繁. 为了防止传输的信息被截获破解,人们越来越重视对信息的传输加密. 由于混沌系统具有有界性、遍历性、伪随机性、对初值条件和控制参数敏感性等特点,人们把混沌系统用于信息加密领域. 2010 年,王林林[13]提出了基于混沌的密码算法设计与研究. 2015 年,贾嫣等[14]提出了基于改进混沌映射的图像加密算法,该算法采用的是雅克比椭圆映射对初始密钥进行迭代产生新的密钥. 2020 年,严利民等[15]提出了混沌映射和流密码结合的图像加密算法. 2021 年,蔡敏等[16]设计并实现了基于混沌时间序列的图像加密算法. 同年,曾祥秋等[17]提出了基于改进的 Logistic 映射的混沌图像加密算法. 考虑到 DNA 可以存储大量信息的特点,人们开始将 DNA 和混沌系统进行结合,并用于信息加密领域. 2016 年,Wang 等[18]提出使用 CML 和 DNA 序列操作的彩色图像加密方案. 该方案将图像的每一个像素进行 DNA 编码,并给出了 8 种相对应的 DNA 编码规则. 2017 年,孙倩等[19]提出了基于 DNA 编码与统计信息优化的图像加密算法. 2020 年,朱凯歌等[20]设计了基于 DNA 动态编码和混沌系统的彩色图像无损加密算法.

之前,人们所提出的将 DNA 用于信息加密的算法中,并没有涉及到 DNA 计算中的化学反应,只是将数据转换成了 DNA 编码的形式. 本文提出了一种基于聚合酶链置换反应的 2D-LASM 混沌映射文本加密算法,成功将 DNA 计算中的化学反应结合到加密过程中,并对该算法进行了密钥空间分析,表明该算法具有较好的加密效果.

1 准备工作

1.1 二维 Logistic-Adjusticed-sine (2D-LASM)映射

本文的加密算法采用结构简单、性能优良的二维 Logistic-Adjusticed-sine (2D-LASM) 映射,其数学表达式如下:

$$\begin{cases} X_{n+1} = \sin(\pi\mu(Y_n + 3)X_n(1 - X_n)) \\ Y_{n+1} = \sin(\pi\mu(X_{n+1} + 3)Y_n(1 - Y_n)) \end{cases}$$

其中,$\mu \in [0,1]$,$X_n, Y_n \in (0,1)$. 这里给定初值,设 $X_1 = 0.5$, $Y_1 = 0.5$,为了消除瞬态效应迭代 n 次,产生两个长度为 n 的伪随机数组 X 和 Y.

1.2 DNA 序列编码

DNA(脱氧核糖核酸)是染色体的主要组成成分,同时也是主要遗传物质. DNA 是双螺旋结构,有两条脱氧核苷酸链,一个脱氧核苷酸分子由三个分子组成:一分子含氮碱基、一分子脱氧核糖及一分子磷酸. 脱氧核苷酸共有 4 种含氮碱基分别为:A(腺嘌呤)、T(胸腺嘧啶)、G(鸟嘌呤)和 C(胞嘧啶). 按照 Watson-Crick 碱基互补配对原则,A 和 T 互补,C 和 G 互补. 在计算机中,信息的存储是用二进制 0 和 1 进行表示. 二进制中 0 和 1 是互补的,因此 00 和 11,10 和 01 也是互补的. 若将 A、T、C、G 分别用 00,01,10,11 进行表示共有 24 种,而符合互补原则的只有 8 种,见表 1.

表 1 DNA 编码规则

Table 1 The cading rules of DNA

Rule	1	2	3	4	5	6	7	8
00	A	A	C	G	C	G	T	T
01	C	G	A	A	T	T	C	G
10	G	C	T	T	A	A	G	C
11	T	T	G	C	G	C	A	A

1.3　聚合酶链置换反应原理

聚合酶链置换反应(Polymerase Strand Displacement, PSD)是一种基于聚合酶的链置换反应,其和一般的链置换反应不同的是该反应需要有酶的参与. 反应原理见图 1,这里的 A 和 A * 互补,B 和 B * 互补,C 和 C * 互补,D 和 D * 互补. 这种反应类似于聚合酶链式反应(PCR). 首先,引物和部分互补的双链 DNA 粘性末端按碱基互补配对原则结合,即 A 和 A * 相结合;其次,将温度调至 DNA 聚合酶最适反应温度,在 DNA 聚合酶的作用下,从引物的 3′端开始以 5′→3′端的方向延伸,合成与模板 5′-D-C-B-A-3′互补的 DNA 链,将部分互补的双链 DNA 中的单链 5′-C * -D * -3′置换出来.

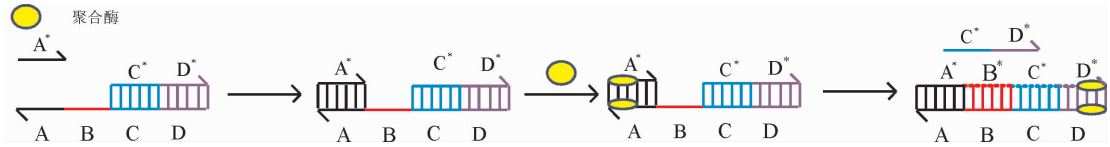

图 1　聚合酶链置换反应原理

Fig. 1　Principle of polymerase strand displacement reaction

2　本文的加密算法

2.1　算法框架

本文算法的加密过程主要分为两个部分:

(1)将待加密的文本信息进行置乱处理,随后和 2D-LASM 混沌映射产生的伪随机序列进行异或操作,之后再根据产生的随机种子选择相应的 DNA 编码规则得到相对应的 DNA 序列;

(2)将所得到的 DNA 序列进行 PSD 反应,得到一组新的 DNA 序列. 根据随机种子选择相应的 DNA 解码规则,得到加密的文本.

本文的加密算法具体框架如图 2 所示.

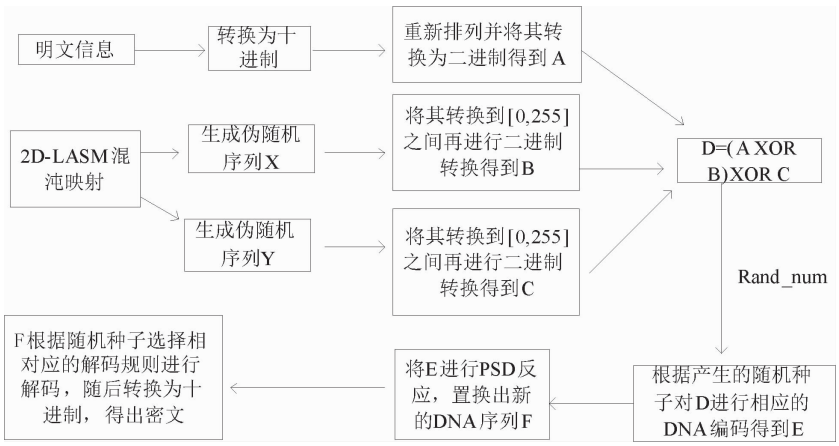

图 2　算法框架

Fig. 2　Algorithm framework

2.2　加密算法的工作步骤

Step1:输入一段明文,将其转换为十进制的数,长度为 L;

Step2:采用 *randperm* 函数将十进制数进行重新排列,并将其转换为二进制数,记为 A;

Step3:通过二维 Logistic-Adjusted-Sine 混沌映射迭代 1 000 + L 次,得到两个伪随机序列 X,Y;

Step4:分别取 X,Y 的后 L 位,经过 $\mathrm{mod}(floor(k*10^8),256)$ 将其转换到[0,255]之间并进行二进制处理得到 B,C,这里 mod 表示取余,*floor* 表示向下取整;

Step5:对 Step2 中的 A 和 Step4 中所得到的 B,C 进行异或操作得到 D;

Step6:根据产生的随机种子选择相应的 DNA 编码规则对 D 进行编码,得到 DNA 序列 E;

Step7:将所得到的 DNA 序列 E 按照长度为 *m*

进行分割,得到 u 条 DNA 序列. 再将这 u 条序列分别进行 PSD 反应,可以置换出 u 条新的 DNA 序列;

Step8:将 Step7 所得到的 u 条 DNA 序列在 DNA 连接酶的作用下组成一条新的 DNA 序列 F;

Step9:根据随机种子选择相应的 DNA 解码规则对 F 进行解码,并将其转换为十进制数 G;

Step10:对 G 使用 *char* 函数得到相应的密文.

解密过程是加密过程的逆过程,这里由于篇幅问题就不再阐述.

2.3 实例分析

该实例的仿真是在 Matlab 2016a 仿真软件上进行的,这里所采用的明文为 you are a better man;将其转换为十进制数得到[121,111,117,32,97,114,101,32,97,32,98,101,116,116,101,114,32,109,97,110],可知 L 为 20. 并将其进行重新排列得到[97,101,98,114,110,121,97,109,116,32,32,32,117,32,101,116,111,97,101]. 随后转换为二进制数得到 A;

通过二维 Logistic-Adjusted-Sine 混沌映射迭代 1 000 + 20 次,得到两个伪随机序列 X,Y;

分别取 X,Y 的后 20 位,经过 mod(*floor*($k * 10^8$),256) 将其转换到[0,255]之间并进行二进制处理得到 B,C,这里 mod 表示取余,*floor* 表示向下取整;

对 A,B,C 进行异或操作得到 D;这里产生的随机种子为 2,即选择第二种 DNA 编码规则进行编码得到 DNA 序列 E: CGACCGTCGTTTCTACAAAGGGAGTAATTCGGCTATACTTTGTTGACGGGCGAAGCCTAGCCCTGTGACTCCTTCCCCAT 长度为 80 nt;对 DNA 序列 E 按照长度为 16 nt 进行分割,可得到 5 条 DNA 序列,见图 3.

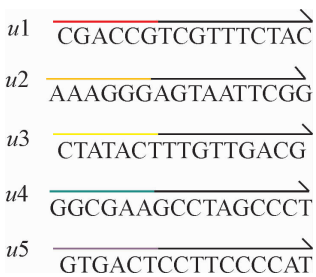

图 3 DNA 序列 E 的分割

Fig. 3 Segmentation of DNA sequence E

将所得到的 5 条 DNA 序列进行 PSD 反应,具体过程见图 5. 为了保证和 DNA 序列 E 相对应,生成一组随机数,根据上述的方法进行 DNA 编码得到 DNA 序列 F:CCGCAGTTCATGGACCGCTCGGTTAGACTCGAGATCTTTCATTAAAAGATTCCGACGCCACCTCGGTGTTTGGCGTACCT 长度为 80 nt;对 F 按照 16 nt 进行分割得到 5 条 DNA 序列,见图 4.

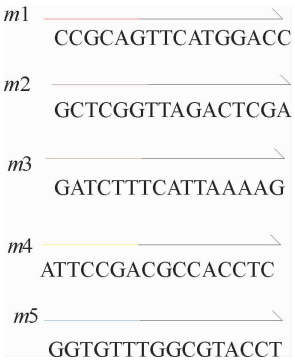

图 4 DNA 序列 F 的分割

Fig. 4 Segmentation of DNA sequence F

经过 PSD 反应之后得到 5 条长度为 16 nt 的 DNA 单链,将这 5 条单链按照 *m*1-*m*2-*m*3-*m*4-*m*5 的顺序进行合并得到 F. 这样通过 PSD 反应将文本信息转换的 DNA 序列 E 成功变成了 DNA 序列 F;接下来根据随机种子选取相对应的 DNA 解码规则将其转换为二进制数 G,将 G 转换成十进制数 H,之后得到密文"|Jn_āNþ < >h®]ýg +".

通过以上步骤成功将明文 you are a better man 转换为|□Jn_āNþ < >□h®]ýg +. 需要注意的是这里的 DNA 序列 F 是根据产生的随机数进行 DNA 编码的,产生的随机数不同这里的 F 就不同,最终得到的密文也就有所不同.

3 算法安全性分析

若想对密文进行解密需要知道明文所对应的 DNA 序列 E 和所产生的随机种子,但由于 DNA 序列 F 是由产生的随机数所得到的,再加上 PSD 反应是不可逆的,因此,得到 DNA 序列 E 就显得很困难,此外还需要知道明文信息转换为十进制之后是如何进行重新排列的. 本文将经过 PSD 反应(图 5)所置换出来的 DNA 单链设置为 16 nt,由于 DNA 具有 4 种含氮碱基,因此,每条 DNA 单链共有 4^{16} 种可能性. 由于满足互补条件的 DNA 编码规则共有 8 种,因此,相对应的 DNA 解码规则也有 8

种. 对明文信息转换成十进制的数进行重新排列共有 20! 种,因此,该算法的密钥空间为(20!) * 8 * 8 * 4^{16} * 4^{16} * 4^{16} * 4^{16} ≈ 2^{184} 远大于 2^{100},这说明该算法可以抵抗穷举攻击,安全性能高.

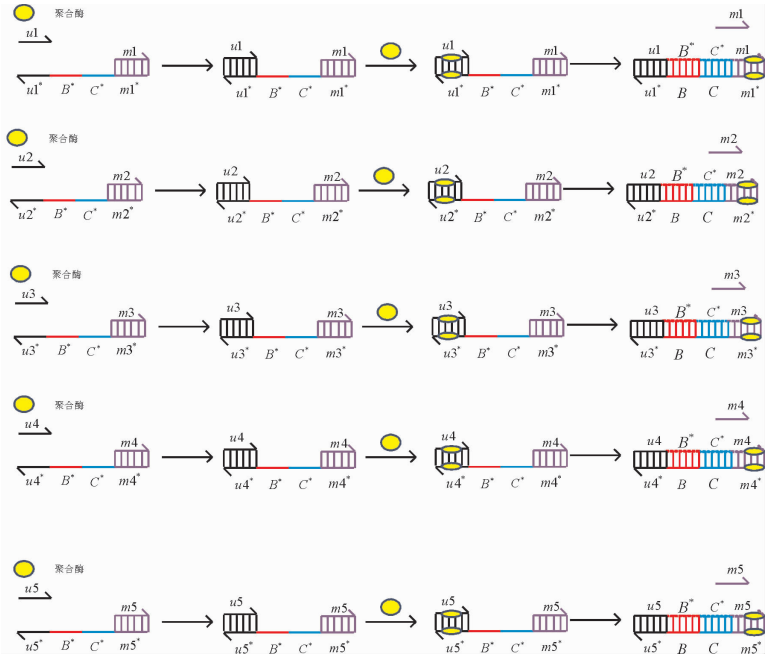

图 5　PSD 反应

Fig. 5　Polymerase stand displacement reation

4　结　论

本文提出了一种基于聚合酶链置换反应的 2D-LASM 混沌映射文本加密算法,克服了之前 DNA 加密算法中没有涉及到 DNA 计算加入化学反应这一缺点. 加入化学反应之后,算法的密钥空间有了很大的改善. 同时,这一算法还为图像加密提供了思路. 如何将 DNA 计算中的化学反应用于图像加密,这是以后需要解决的问题.

参考文献:

[1] Adleman L. Molecular computation of solutions to combinatorial problems[J]. Science (American Association for the Advancement of Science), 1994, 266(5187): 1021-1024.

[2] Head T, Rozenberg G, Bladergroen R S, et al. Computing with DNA by operating on plasmids[J]. Biosystems, 2000, 57(2): 87-93.

[3] 殷志祥,张凤月,许进. 0-1 规划问题的 DNA 计算[J]. 电子与信息学报, 2003(1): 62-66.

[4] Qian L, Winfree E, Bruck J. Neural network computation with DNA strand displacement cascades[J]. Nature, 2011, 475(7356): 368-372.

[5] Orbach R, Wang F, Lioubashevski O, et al. A full-adder based on reconfigurable DNA-hairpin inputs and DNAzyme computing modules[J]. Chemical Science (Cambridge), 2014, 5(9): 3381-3387.

[6] Lin H, Chen J, Li H, et al. A simple three-input DNA-based system works as a full-subtractor[J]. Scientific Reports, 2015, 5(1):10686-10691.

[7] Yang C, Chen Y, Lin H, et al. An optical deoxyribonucleic acid-based half-subtractor[J]. Chemical Communications, 2013, 49(78): 8860-8862.

[8] Li W, Zhang F, Yan H, et al. DNA based arithmetic function: A half adder based on DNA strand displacement[J]. Nanoscale, 2016, 8(6): 3775-3784.

(下转第 34 页)