文章编号:2095-476X(2013)04-0091-04

以细菌DNA为载体的信息隐藏方案

崔光照, 赵晓航

(郑州轻工业学院 电气信息工程学院, 河南 郑州 450002)

摘要:利用重组DNA技术,结合细菌的抗药性,提出了以细菌DNA为载体的信息隐藏方案:重组DNA实验条件和细菌抗药性是密钥,载体DNA和冗余DNA混合在一起,受体细菌和冗余细菌混合在一起,明文得到了两度隐藏.随机尝试可能既破坏密文又得不到明文.安全性分析表明,本方案具有很强的安全性.

关键词:DNA密码;重组DNA技术;信息隐藏

中图分类号:TP309;TP18 **文献标志码**:A DOI:10.3969/j.issn.2095-476X.2013.04.022

A bacterial DNA based information hiding scheme

CUI Guang-zhao, ZHAO Xiao-hang

(*College of Electric and Information Engineering, Zhengzhou University of Light Industry, Zhengzhou* 450002, *China*)

Abstract: A bacterial DNA based information hiding scheme was proposed by utilizing the recombinant DNA technology as well as the resistance of bacteria. In this scheme recombinant DNA experimental conditions and bacterial resistance were the key. Vector DNA molecules and redundant DNA molecules were mixed together and then receptor bacteria and redundant bacteria were mixed together. Thus the plaintext was hidden twice. Random attempt may not only destroy the ciphertext but also miss the plaintext. Security analysis showed that the proposed scheme had strong security.

Key words: DNA cryptography; recombinant DNA technique; information hiding

0 引言

信息是国家安全、社会稳定和经济发展的战略资源,密码学是信息安全领域的核心.随着DNA计算的发展进步,基于数学困难问题的密码受到了威胁.2005年,W. L. Chang等[1]提出了试除法分解大数的DNA计算模型,分析了破译RSA的计算复杂度.2008年,X. C. Zhang等[2]提出了用DNA自组装破译公钥密码系统NTRU的非确定性算法.2010年杨学庆等[3]针对RSA陷门库特点,提出一种新的DNA计算模型并破译了RSA.DNA计算也给密码学带来了机遇.1999年,C. Clelland等[4]将信息隐藏在DNA微点中,实现了以DNA为载体的信息隐藏.2006年,肖国镇等[5]预测DNA密码将会成为未来密码学的三大主要领域之一.2007年,卢明欣等[6]提出了基于DNA技术的对称加密方法.2010年,来学嘉等[7]提出了基于DNA技术的非对称加密方法.2010年,Z. H. Chen等[8]用DNA自组装实现了一次一密密码.同年,M. Hirabayashi等[9]改进了一次一密密码本的生成过程.2012年,蒋君等[10]分析了

收稿日期:2013-03-10

基金项目:2012年度河南省科技创新人才计划项目(124200510017)

作者简介:崔光照(1957—),男,河南省洛宁县人,郑州轻工业学院教授,博士,硕士研究生导师,主要研究方向为生物信息计算、信息安全.

DNA 密码与传统密码和量子密码的优势与不足.

信息隐藏是 DNA 密码的一个重要研究方向. 重组 DNA 技术是一种先在细胞外重新组合 DNA 分子,再将重组 DNA 分子导入受体细胞中增殖的分子生物学技术. 细菌对抗生素有抗药性,不同的细菌具有不同的抗药性. 重组 DNA 技术和细菌的抗药性在信息隐藏方面有着特殊的用途. 本文将利用重组 DNA 技术,结合细菌的抗药性,提出以细菌 DNA 为载体的信息隐藏方案.

1 密码设计

重组 DNA 实验一般包括 4 步:

1)获得目的基因;

2)目的基因与克隆载体连接,形成重组 DNA 分子;

3)将重组 DNA 分子导入受体细胞;

4)筛选出含有重组 DNA 分子的细胞,进而提取出目的基因.

前 3 步可以看作隐藏目的基因,第 4 步可以看作提取隐藏的目的基因.

抗生素对细菌有杀灭作用,细菌对抗生素有抗药性. 根据细菌的抗药性,选择恰当的抗生素可以从大量细菌中筛选出需要的细菌.

重组 DNA 技术和细菌的抗药性在信息隐藏方面有特殊的用途.

1.1 加密过程

将重组 DNA 技术和细菌的抗药性应用于信息隐藏,可得到具有较高安全性的密码. 加密过程如图 1 所示.

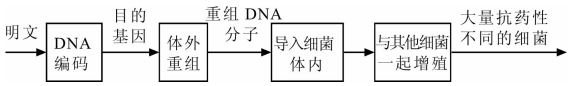

图 1 加密过程

1)DNA 编码. 对信息隐藏来说,编码方式对信息安全的影响不大,可以采用简单的替换编码. 三联体编码方式见表 1, 3 个核苷酸表示 1 个字符[4].

2)制作含有目的基因的 DNA 片段. 根据表 1 所示编码方式,明文映射成 DNA 序列,该 DNA 序列即为目的基因. 在目的基因前后各加上一段 DNA 序列,得到含有目的基因的 DNA 片段. 其中目的基因存储明文,目的基因前后的片段是后续剪接处理的对象[11].

表 1 三联体编码方式

A = CGA	I = ATG	Q = AAC	Y = AAA	6 = TTA
B = CCA	J = AGT	R = TCA	Z = CTT	7 = ACA
C = GTT	K = AAG	S = ACG	0 = ACT	8 = AGG
D = TTG	L = TGC	T = TTC	1 = ACC	9 = GCG
E = GGC	M = TCC	U = CTG	2 = TAG	= ATA
F = GGT	N = TCT	V = CCT	3 = GCA	, = TCG
G = TTT	O = GGA	W = CCG	4 = GAG	. = GAT
H = CGC	P = GTG	X = CTA	5 = AGA	: = GCT

3)制作重组 DNA 分子(载体 DNA 分子). 选择恰当的克隆载体,在适当的条件下,剪接含有目的基因的 DNA 片段和克隆载体,得到重组 DNA 分子[11].

4)隐藏重组 DNA 分子. 选择适当的细菌做受体细菌,将重组 DNA 分子导入受体细菌内,再用大量冗余细菌隐藏受体细菌.

受体细菌体内还有大量其他 DNA 分子(冗余 DNA 分子),冗余 DNA 分子对重组 DNA 分子起到了很好的隐藏作用. 冗余细菌对含有重组 DNA 分子的细菌(受体细菌)起到了很好的隐藏作用.

1.2 解密过程

根据细菌的抗药性和重组 DNA 实验条件解密,解密过程如图 2 所示.

1)杀灭不含重组 DNA 分子的细菌. 冗余细菌和受体细菌具有不同的抗药性,据此选择恰当的抗生素杀灭冗余细菌,筛选出受体细菌.

2)分离出重组 DNA 分子. 根据克隆载体的性质设计恰当的实验条件,使重组 DNA 分子扩增,冗余 DNA 分子不扩增. 再经过分离、纯化筛选出重组 DNA 分子.

3)译码. 解开目的基因和克隆载体的连接,得到目的基因,根据表 1 所示编码方式译码得到明文.

图 2 解密过程

2 安全性分析

细菌抗药性和重组 DNA 实验条件组成联合密钥,加密密钥和解密密钥相同. 知道了细菌的抗药性,可以选择恰当的抗生素杀灭冗余细菌,留下受体细菌,从而分离出受体细菌;知道了重组 DNA 实验条件,可以在恰当条件下使重组 DNA 分子扩增,

冗余 DNA 分子不扩增,从而分离出重组 DNA 分子. 随机尝试可能既破坏密文又得不到明文.

抗生素的种类已经达到数千种,临床上常用的亦有几百种. 假定攻击者得知解密所用的抗生素为某 N 种抗生素中的 M 种抗生素(这 N 种抗生素已知,具体使用哪 M 种抗生素未知),随机尝试,一次成功的概率用 P 表示.

$$P = \frac{1}{A_N^M}$$

假定细菌可以反复利用,需要尝试的次数

$$C = A_N^M$$

假定 M 的值为 3,使用 Matlab 在 2.40 GHz,Intel(R)Core(TM)i3 CPU,2 G 内存和 1 G 独显的电脑上绘制 NP 曲线与 NC 曲线,如图 3 所示.

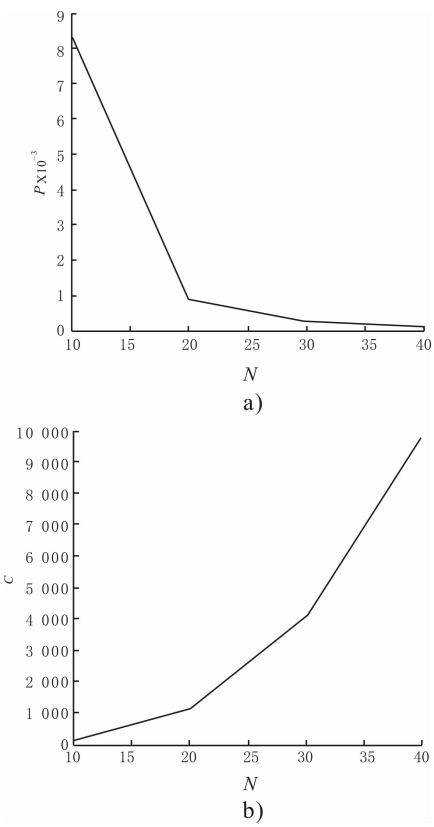

图 3　$M=3$ 时 NP 曲线与 NC 曲线

由图 3 可知,P 的值随 N 的值的增加而迅速减小,C 的值随 N 的值的增加而迅速增大. 尝试一次就获得成功的概率是很低的,即使 N 的值很小,也要尝试多次才能取得成功. 然而,细菌是不可反复利用的,尝试一次而没有成功,含有重组 DNA 分子的细菌极有可能被杀灭,从而无法得到明文.

假定 N 的值为 10,使用 Matlab 在 2.40 GHz,Intel(R)Core(TM)i3 CPU,2 G 内存和 1 G 独显的电脑上绘制 MP 曲线与 MC 曲线,如图 4 所示.

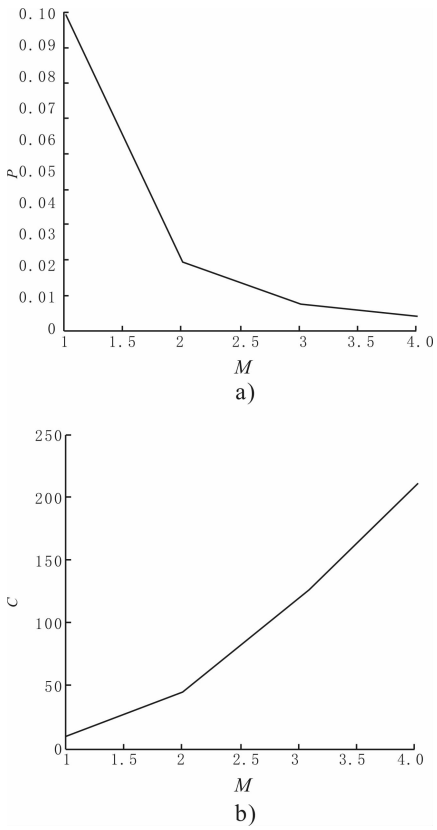

图 4　$N=10$ 时 MP 曲线与 MC 曲线

由图 4 可知,即使攻击者可以将 N 的值从几千缩小到 10,只要稍微增加 M 的值就可以大大提高破译难度.

如果攻击者得知了细菌的抗药性,选择恰当的抗生素分离出了受体细菌,接下来要做的是从众多 DNA 分子中分离出重组 DNA 分子. 随机设定实验条件,可能破坏重组 DNA 分子,从而无法得到明文.

如果攻击者得到了重组 DNA 分子,剩余的安全性就很弱了. 攻击者可以比较每次截获的重组 DNA 分子,不变的部分是载体,变的部分是含有目的基因的 DNA 片段. 再根据语言统计特性就可以破译编码方式,获得明文. 比如,目的基因中出现频度最高的是 GGC,英文中出现频度最高的是 E,GGC 可能就是英文字母 E.

如果公开编码方式,攻击者可以根据编码方式猜测哪些 DNA 序列是目的基因. 比如某一 DNA 序列为 TTGTCTCGAGGGTTTCCCGTGCGCAAA,这一 DNA 序列不是目的基因,因为表 1 所示编码方式中没有 GGG,CCC 这些编码方式. 又如另一 DNA 序列

为 TTGTCTCGAGTTCGAGTGCGCAAA,这一 DNA 序列很有可能是目的基因,因为三联体的排列完全符合表 1 所示编码方式. 有了编码方式,攻击者可以验证分离出的 DNA 分子是否为重组 DNA 分子,因此,公开编码方式会降低安全性.

3 特点与缺陷

特点是重组 DNA 与冗余 DNA 混合在一起,受体细菌与冗余细菌混合在一起,明文得到了两度隐藏;随机尝试,可能既破坏明文又得不到密文.

缺陷是细菌抗药性可能发生变异,冗余细菌的抗药性可能变得和受体细菌相同,从而使正确的接收方也无法解密;大量使用抗生素可能导致细菌产生更强的抗药性,对人类产生危害.

4 结论

利用重组 DNA 技术、结合抗生素杀菌作用以及细菌抗药性提出了以细菌 DNA 为载体的信息隐藏方案. 信息得到了两度隐藏,不知道密钥,可能会既毁灭了密文也得不到明文. 今后的研究还需进一步对该方案进行评估与改进.

参考文献:

[1] Chang W L, Guo M, Ho M S H. Fast parallel molecular algorithms for DNA-based computation: factoring integers [J]//IEEE Transactions on Nanobioscience, 2005, 4(2):149.

[2] Zhang X C, Niu Y, Cui G, et al. Breaking the NTRU public-key cryptosystem using self-assembly of DNA tilings [J]. Chinese Journal of Computers, 2008, 12(31):2129.

[3] 杨学庆,柳重堪. 基于 DNA 计算的 RSA 密码系统攻击方法[J]. 计算机工程,2010,36(2):1.

[4] Celland C, Risca V, Bancroft C. Hiding messages in DNA microdots[J]. Nature, 1999, 399(6736):533.

[5] 肖国镇,卢明欣,秦磊,等. 密码学的新领域——DNA 密码[J]. 科学通报,2006,51(10):1139.

[6] 卢明欣,来学嘉,肖国镇,等. 基于 DNA 技术的对称加密方法[J]. 中国科学,2007,37(2):175.

[7] 来学嘉,卢明欣,秦磊,等. 基于 DNA 技术的非对称加密与签名方法[J]. 中国科学,2010,40(2):240.

[8] Chen Z H, Xu J. One-Time-Pads encryption in the tile assembly model[J]. Journal of Computational and Theoretical Nanoscience, 2010, 7(5):848.

[9] Hirabayashi M, Kojima H, Oiwa K. Design of true random one-time pads in DNA XOR cryptosystem[J]. Proceedings in Information and Communications Technology, 2010(2):174.

[10] 蒋君,殷志祥. DNA 密码对比传统密码学与量子密码学的优势与不足[J]. 科技视界,2012,24:24.

[11] 饶妮妮. 一种基于重组 DNA 技术的密码方案[J]. 电子学报,2004,32(7):1216.

(上接第 85 页)

[4] Benedetto S, Divsalar D, Montorsi G, et al. Pollara. Serial concatenation of interleaved codes: Performance analysis, design, and iterative decoding[J]. IEEE Transactions on Information Theory, 1998, 44(3):909.

[5] Ming Xiao, Tor Aulin. Serially concatenated continuous phase modulation with convolutional codes over rings[J]. IEEE Transaction on Communication, 2006, 54(8):1387.

[6] Maw R L, Taylor D P. Space-time coded systems using continuous phase modulation[J]. IEEE Transaction on Communication, 2007, 55(11):2047.

[7] Zhang Zhijun, Zhang Aili, Fu Huan. Performance study on convolutional coded MSK system[C]//The 7th International Conference on Wireless Communication, Networking and Mobile Computing, Piscataway: Institute of Electrical and Electronics Engineers Inc, 2011:1074 – 1076.

[8] Zhang Zhijun, Yang Yujie, Wang Xianfang, et al. Serially concatenated coded M-ary continuous phase modulation with bit-interleaver[J]. Journal of Theoretical and Applied Information Technology, 2012, 45(1):144.

[9] Zhang Zhijun, Wang Weifeng, Xintao Duan. Performance analysis of Q-ary low-density parity-check coded continuous phase modulations[J]. International Review on Computers and Software, 2012, 7(7):3665.

[10] Narayanan K R, Altunbas I, Narayanswami R S. Design of serial concatenated MSK schemes based on density evolution[J]. IEEE Transaction on Communication, 2003, 51(8):1283.

[11] Bahl L, Cocke J, Jelinek F, et al. Optimal decoding of linear codes for minimizing symbol error rate[J]. IEEE Transactions on Information Theory, 1974, 20(2):284.