第 36 卷 第 2 期 计 算 机 工 程 2010 年 1 月
Vol.36 No.2 Computer Engineering January 2010
·安全技术· 文章编号：1000—3428(2010)02—0135—02 文献标识码：A 中图分类号：TP309.2

基于 DNA 计算的 IDEA 密码攻击方法

杨学庆[1,2]，柳重堪[1,2]
(1. 数学、信息与行为教育部重点实验室，北京 100083；2. 北京航空航天大学电子信息工程学院，北京 100083)

摘 要：针对国际数据加密算法(IDEA)密码的特点，提出一种基于 DNA 计算的粘附子模型的 IDEA 密码系统攻击方法。该方法使用已知明文进行攻击，采用 DNA 储存链编码各种可能的密钥与已知明文，通过组合、分离、设置、清除 4 种操作筛选出密钥，由凝胶电泳确定密钥的具体值。该攻击方法所需的数据量仅为一组明文密文对，时间复杂度为 $O(n^2)$。

关键词：DNA 计算；国际数据加密算法；粘附子模型

Attacking Method on International Data Encryption Algorithm Code Based on DNA Computing

YANG Xue-qing[1,2], LIU Zhong-kan[1,2]
(1. Key Laboratory of Mathematics, Informatics and Behavioral Semantics of Ministry of Education of China, Beijing 100083;
2. School of Electronic and Information Engineering, Beihang University, Beijing 100083)

【Abstract】 In terms of the features of International Data Encryption Algorithm(IDEA), this paper presents an attacking method on IDEA code system, which is based on the sticker model of DNA computing. It uses known-plaintext to realize attack. All possible key and the known plaintext are encoded by DNA strands, and the desired key is sorted out by applying combination, separation, set and clear. The corresponding key is read out by gel electrophoresis. The method requires only one pair of plaintext-ciphertext and its time complexity is $O(n^2)$.

【Key words】 DNA computing; International Data Encryption Algorithm(IDEA); sticker model

1 概述

DNA[1]因为具有极强的并行计算能力与海量的储存密度，所以一经创建就被尝试用来处理密码攻击问题[2]。在诸多的 DNA 计算模型中，粘附子模型[3]的突出优点为：实验实现比较容易。国际数据加密算法(International Data Encryption Algorithm, IDEA)[4]是 20 世纪 90 年代出现的分组加密算法，现已成为目前公开的最好、最安全的分组算法之一，并已被纳入欧洲 Nessie 的密码大计划。由于 IDEA 算法比较容易实现，因此被广泛应用于维护网络信息安全。

本文针对 IDEA 密码系统的特点，重点研究了基于 DNA 计算的用粘附子模型实现 IDEA 加密规则的操作方法。

2 粘附子模型简介

粘附子模型[3]使用 2 种类型的 DNA 单链，长的 DNA 单链称为储存链。将储存链划分成数个长度相等的区域，与这些区域互补的短的 DNA 单链称为粘附子。粘附子与其互补的区域一起编码一个比特。当储存链上的某个区域与其相对应的粘附子结合时，该区域编码的比特的值为 0；反之，比特的值为 1，如图 1 所示。

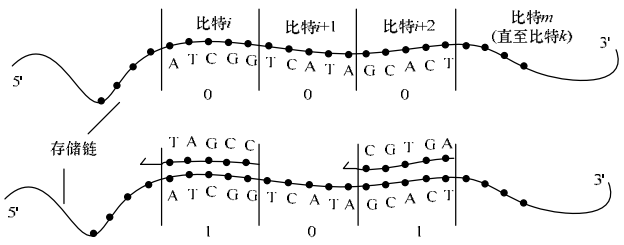

图 1 粘附子模型

粘附子模型包括以下 4 种基本操作：

(1)组合：将 2 个由字符串组成的集合组合成一个新的多重集合，其元素由原来 2 个集合的所有字符串组成，其中，元素有可能重复出现。

(2)分离：将一个由字符串组成的集合分成 2 个新的集合，把指定比特为 0 的字符串归为一个集合，指定比特为 1 的字符串归为另一个集合。

(3)设置：将一个集合中每个字符串中指定比特的值赋为 1。

(4)清除：将一个集合中每个字符串中指定比特的值赋为 0。

3 IDEA 简介

IDEA 是一个分组加密算法，密钥长度为 128 位。它以 64 位为分组对数据加密。明文从算法的一端输入，经过 8 轮运算变成密文从另一端输出。在每一轮中，64 位的数据分组被分成 4 个 16 位子分组：X_1, X_2, X_3 和 X_4，执行图 2 所示的 14 步运算。经过 8 轮运算之后，通过一个最终输出变换产生密文。图中，X_i 表示 16 位明文子分组；Y_i 表示 16 位密文子分组；$Z_i(r)$表示 16 位子密钥；⊕表示 16 位子分组的相异或；⊞表示 16 位整数的模 216 加；⊙表示 16 位整数与 216 对应 0 子分组的模 216+1 乘。

基金项目：国家自然科学基金资助重点项目(11037705)
作者简介：杨学庆(1972－)，女，博士研究生，主研方向：DNA 计算，密码破译；柳重堪，教授、博士生导师
收稿日期：2009-05-05 **E-mail：**yxqjessica@sohu.com

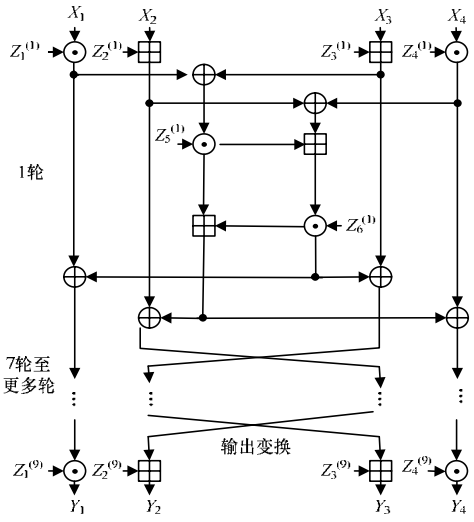

图 2　IDEA 算法

4　攻击 IDEA 密码系统的 DNA 算法

用 DNA 计算攻击 IDEA 密码系统的方法为已知明文攻击法，这时知道密钥空间、一些明文以及与这些明文相对应的密文。设 IDEA 的密钥库为 k_d，已知明文中的任一组明文为 p，与 p 相对应的密文为 c。将 p 根据 IDEA 的加密规则与 k_d 中的每一个密钥进行加密运算，将运算结果与 c 进行比较，找出其中满足 $E_k(p)=c$ 的 k，便可确定密钥。使用 DNA 计算攻击 IDEA 有以下 4 步：

(1)用 DNA 分子表示 p 与 k_d。具体实现过程如下：

1)合成 1 635 碱基长的 DNA 分子，将该 DNA 分子均分为 325 个区域，每个区域对应于一位，将 DNA 链的这些区域记为 $B_n(0\leq n\leq 324)$，其中，B_1~B_{128} 区储存密钥，记区域 B_n 所对应的粘附子为 S_n。

2)为了生成各种可能的密钥，首先合成若干 1 635 碱基长的 DNA 单链与充足的粘附子 S_1~S_{128}；然后将 DNA 单链平均分成 2 份，分别放入试管 T_1 与 T_2 中，在 T_1 中放入过量的粘附子 S_1~S_{128}，充分反应后移去多余的粘附子；最后组合 T_1 与 T_2，产生试管 T，加热后再冷却。

3)用 DNA 存储链的 B_{129}~B_{192} 区域来编码已知明文。根据已知明文的值向 T 中放入过量的粘附子，充分反应后分离出多余的粘附子。

(2)将 p 与 k_d 中的各种密钥进行 IDEA 加密运算，产生密文库。这一步的加密过程本质上是模乘、模加与异或运算的组合。具体实现如下：

1)异或运算

①设需对第 n 位进行异或运算，第 k 位存储运算结果。

②对储存链的第 n 位进行分离操作，产生 T_1 与 T_0。

③根据储存链的第 n 位对 T_1 与 T_0 并行实施分离操作，产生 T_{11}, T_{10}, T_{01} 与 T_{00}。清除 T_{11} 与 T_{00} 的第 k 位，设置 T_{10} 与 T_{01} 的第 k 位。

2)加法运算

for(j=a+1;j<= a+p;j++)

{根据储存链的第 j 位对试管 T 进行分离操作，产生 2 个试管 T_1 与 T_0;

根据储存链的第 j+p+1 位对 2 个试管 T1, T0 同时并行进行分离操作，产生 4 个试管 T_{11}, T_{10}, T_{01} 与 T_{00};

根据储存链的第 a＋3p＋2 对 4 个试管同时进行分离操作，产生 8 个试管 T_{111}, T_{101}, T_{011}, T_{001}, T_{110}, T_{100}, T_{010} 与 T_{000};

设置试管 T_{100}, T_{010}, T_{001} 与 T_{111} 第 j+2p+1 位，设置试管 T_{011} 的第 a+3p+2 位，清除试管 T_{100} 的第 a+3p+2 位，清除试管 T_{101}, T_{011}, T_{000} 与 T_{110} 第 j+2p+1 位;

将以上 8 个试管组合，生成试管 T;}

3)模 2^{16}+1 乘运算

用 DNA 计算实现模 $2^{16}+1$ 乘运算时，先运用数学原理将它转化成加法，然后执行基于 DNA 计算的加法运算。将模乘运算转化成加法运算的方法如下：记 $\alpha=a_n a_{n-1}\cdots a_1$，$\beta=\beta_n\beta_{n-1}\cdots\beta_1$，$\gamma=\gamma_{2n}\gamma_{2n-1}\cdots\gamma_1=\alpha\times\beta$，其中，$\alpha_i,\beta_i\in\{0,1\}$ 为 n 位二进制数，记 $X[\alpha]=\{i,\alpha_i=1\}$，即 $X[\alpha]$ 为 α =1 的位的指标的集合，则 $\alpha\times\beta=\sum_{i\in X[\alpha]}\beta\times 2^{i-1}$。运用以下方法，可以将模 2^{16}+1 乘运算转化成加法运算：

γ=0

for i=1

if α_i=1 then

for j=1 to n-i+1

$\gamma_{i+j-1}=\gamma_{i+j-1}+\beta_j$

4)储存空间的分配方案

IDEA 由 8 轮相同的运算组成，每一轮包括 14 步，计算这 14 步时储存空间的分配方案如下：第 1 步~第 4 步的计算过程使用 B_{257}~B_{323} 区，这 4 步的计算结果存储于 B_{129}~B_{192} 区；第 5 步、第 6 步的计算结果存储于 B_{257}~B_{288} 区；第 7 步的计算过程使用 B_{291}~B_{323} 区，其计算结果存储于 B_{257}~B_{272} 区；第 8 步~第 10 步的计算结果存储于 B_{274}~B_{321} 区；最后 4 步的计算结果存储于 B_{193}~B_{256} 区。

(3)从产生的密文库中筛选与 c 相同的密文。其 DNA 计算操作过程如下：

for(i=1;i<=64;i++)

{根据 B_{i+128} 区对试管 T 进行分离操作，产生试管 T_1 与 T_0;

当已知密文的第 i 位为 1，记试管 T_1 为 T；否则，记试管 T_0 为 T;}

通过 64 次分离，筛选出符合条件的密钥。

(4)确定加密 c_i 的密钥 k_i 的具体值。其 DNA 计算操作过程为：将 DNA 复合体的粘附子洗提出来，通过走杂交胶判断存在何种粘附子，从而确定密钥的具体值。

5　算法复杂度与可行性分析

5.1　算法复杂度分析

算法的第(1)步与第(4)步的时间复杂度为 $O(1)$。第(3)步的时间复杂度为 $O(n)$，n 是分组的大小。第(2)步最耗时，也最复杂。它一共包括 24 次模乘运算、40 次模加运算与 48 次异或运算，由于模乘运算的时间复杂度为 $O(n^2)$，模加运算与异或运算的时间复杂度均为 $O(n)$，因此第(2)步的时间复杂度为 $O(n^2)$。即整个算法的时间复杂度为 $O(n^2)$。

5.2　可行性分析

本算法是基于粘附子模型的，粘附子模型的 4 种操作中，清除操作是最容易发生错误的。虽然本算法使用了清除操作，但其并不是必需的，可以通过增加存储空间取消清除操作。这样算法只包括组合、设置、分离这 3 种操作，而这 3 种操作的可操作性与可行性已经通过实验[5]得到了证明，所以，本算法具有较高的实验可行性。

6　仿真

由于 DNA 计算的特殊性，其仿真必须设计专门的软件来进行，并且不同的 DNA 计算模型其仿真软件也不一样，

(下转第 140 页)