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)02－00127－ 文献标识码：A 中图分类号：TN918.1

不可能差分攻击中的明文对筛选方法

张庆贵
(解放军信息工程大学电子技术学院，郑州 450004)

摘 要：基于快速排序原理，提出用于筛选明文对的基本算法和改进算法，改进算法的计算复杂性可以将由直接检测方法的 $O(n^2)$降为 $O(n\log n)$。基于上述结果以改进算法分析对 ARIA 等分组密码算法的几个不可能攻击的计算复杂性，证明 ICISA2008 上发表的某个针对对 ARIA 的不可能攻击的数据筛选过程的计算复杂性远高于密钥求解过程的计算复杂性。

关键词：密码学；密码分析；不可能差分攻击；明文对筛选；计算复杂性；ARIA 算法

Plaintext Pair Sieve Methods in Impossible Differential Attack

ZHANG Qing-gui
(Institute of Electronic Technology, the PLA University of Information Engineering, Zhengzhou 450004)

【Abstract】 Based on quicksort theory, this paper presents a basic plaintext pair sieve algorithm and an improved algorithmone, and the computational complexity of improved algorithm is $O(n\log n)$, which is less than $O(n^2)$ of the method by checking each pairs. It analyzes the computational complexity of plaintext pair sieve in impossible differential attacks on ARIA etc with the new algorithm, and proves that the computational complexity is higher remarkably than that in the key solving process for the one impossible differential attack on ARIA presented in ICISA 2008.

【Key words】 cryptology; cryptanalysis; impossible differential attack; plaintext pair sieve; computational complexity; ARIA algorithm

1 概述

不可能差分攻击[1]自提出以来逐渐成为分析分组密码算法强度的一种重要方法，其主要利用了分组密码算法中某些概率为 0 的差分对应实施攻击，目前用于对 Skipjack, AES, DES, CLEFIA, ARIA 等许多分组密码的进行分析。在很多不可能差分攻击[1-5]中，需要先对明文对进行筛选，并从中挑选出密文差满足特定要求的明文对，再利用挑选出的明文对求解密钥。从选择明文中选出合适明文对的过程称为数据筛选。数据筛选的方法及其计算复杂性分析往往被不可能攻击方案的设计者忽视。虽然数据筛选只是进行简单的比较运算，但在特殊情形下，数据筛选过程的计算复杂性可能远高于密钥求解过程的计算复杂性，造成对不可能差分攻击计算复杂性的误判。大多数针对不可能攻击的文献都分析了通过数据筛选过程的明文对的数量，但未指出如何对明文对进行筛选以及筛选过程的计算复杂性。

文献[1]根据密文在指定字节的值对明密对进行分类从而进行数据筛选。由于该文针对的不可能攻击只需检测密文差在 2 个字节上是否都为 0，因此只需将明文分成 2^{16} 类即可完成数据筛选。该方法在具体实现时需要增加一定的存储空间。文献[5]利用生日悖论原则，每次随机地选取若干明文，并利用碰撞攻击方法以一定的概率找出一个通过检测的明文对，从而实现数据筛选。该方法在具体实现时也需要增加一定的存储空间，而且未必能找出所有能通过检测的明文对。

本文系统地研究了不可能差分攻击中数据筛选方案的设计问题，并基于快速排序思想，提出了一个有效的数据筛选方法，据此证明了文献[23]提出的不可能攻击的计算复杂性是 $2^{132.8}$ 次比较而不是其中声称的 2^{96} 次 6 轮 ARIA 加密运算。

2 不可能差分攻击中数据筛选的原理和方法

不可能差分攻击选择的明文具有十分特殊的结构特点，其中，“结构”是一个重要的概念。

定义 1 在指定比特位置上取固定值、在其它其他位置上任意取值的所有明文构成的集合称为一个结构。

当一个结构中共有 t 个可以任意取值的比特位置时，该结构中共有 2^t 个明文。在一个结构中，任意 2 个不同明文的模 2 和在指定位置上都是 0，因此，一个结构共可产生 $2^{t-1}(2^t-1)\approx 2^{2t-1}$ 个不同的明文对。在不可能攻击中，通常需要选择很多结构。如果不可能差分攻击共选择 2^m 个结构，则这 2^m 个结构平均共可产生 2^{m+2t-1} 个不同的明文对。不可能差分攻击首先需要对这些明文对进行筛选，然后利用挑出的明文对，借助分组密码算法的不可能差分对应求解密钥。

在不可能差分攻击中，明文对的筛选通常是根据密文差进行的。筛选的主要依据是仅保留对应密文的异或差在指定位置上为 0 的明文对。有些不可能差分攻击还要求密文的异或差在特定位置上非 0。如果不可能差分攻击只要求密文差在 n 个指定的比特位上为 0，则此时明文对的保留率为 2^{-n}。因此，经过数据筛选过程后，平均将保留 $2^{m+2t-n-1}$ 个明文对。当一个结构中共有 2^t 个明文时，如果采取通过检测每个密文对的指定比特位的差是否为 0 的方法进行数据筛选，则完成 2^{m+2t-1} 个明文对的检测共需要执行 2^{m+2t-1} 次比较运算。当明文对的数量很大时，其计算复杂性很高，甚至超过密钥求解过程的计算复杂性。因此，需要针对不可能差分攻击中所选明文的特殊结构，研究如何更有效地筛选明文对。

例 1 文献[2]提出，对 6 轮 ARIA 的不可能差分攻击需要选择 2^{64} 个结构，每个结构中的明文在第 3, 4, 6, 8, 9, 13, 14 共 7 个字节上可自由取值，因此，一个结构中共有 2^{56} 个明

基金项目：河南省杰出青年科学基金资助项目(0312001800)
作者简介：张庆贵(1963－)，男，博士博士生，主研方向：密码学
收稿日期：2009-05-12 **E-mail：**zhangqingui@126.com

文。在数据筛选时，需要检测密文的模 2 差在第 0, 2, 3, 4, 5, 6, 7, 11, 12, 13, 15 共 11 个字节上是否都为 0。因此，该不可能攻击共可产生约 2^{175} 个不同的明文对，经过数据筛选后，每个结构平均保留 $2^{56+55}/2^{8\times11}=2^{23}$ 个明文对，2^{64} 个结构平均共保留 2^{87} 个明文对。如果对 2^{175} 个明文对直接按定义检测，则需要执行 2^{175} 次比较，其计算复杂性远远大于求解密钥过程的计算复杂性，即 2^{112} 次 6 轮 ARIA 加密。

例 2 文献[3]提出，对 6 轮 ARIA 的改进的不可能差分攻击需要选择 2^{71} 个结构，每个结构中的明文在第 1, 5, 6, 8, 11, 15 共 6 个字节位置上可自由取值，因此，一个结构中共有 2^{48} 个明文。在数据筛选时，需要检测密文的模 2 差在第 0, 1, 2, 4, 5, 6, 7, 8, 9, 10, 14, 15 共 12 个字节位置上是否都为 0。因此，该不可能攻击共可产生约 2^{166} 个不同的明文对，经过数据筛选后，每个结构平均保留 $2^{48+47}/2^{8\times12}=0.5$ 个明文对，2^{71} 个结构平均共保留 2^{70} 个明文对。如果对 2^{166} 个明文对直接按定义检测，则需要执行 2^{166} 次比较，其计算复杂性远远大于求解密钥过程的计算复杂性，即 2^{96} 次 6 轮 ARIA 加密。

不可能差分攻击的数据筛选问题有一个显著的特点，即在选择的每个结构中，任何 2 个明文的差都符合对明文差的要求。因此，只要对应的密文差满足在指定位置上为 0 或非 0 的要求，该明文对就是符合不可能攻击要求的明文对。由于要求密文差在指定位置上为 0 这个条件的排除率远高于要求密文差在指定位置上非 0 这个条件的排除率，因此可以采取先利用前者进行筛选，再利用后者进行筛选。由于第 1 步的筛选可以排除绝大多数的明文对，因此本文只研究如何利用密文差在指定位置上为 0 进行筛选，对于第 2 步的筛选，按定义进行检测即可。显然，有如下定理：

定理 设 m 和 c 分别是明文及对应的密文，且明文对的筛选条件是 2 个密文在第 $i_1,i_2,\cdots,i_u$ 比特的差都为 0。再设密文 c 在 $i_1,i_2,\cdots,i_u$ 比特位上的值依次为 $c_{i_1},c_{i_2},\cdots,c_{i_u}$，令

$$\psi(m,c)=\sum_{i=1}^{u}c_{i_j}2^{j-1},\ \psi(m_1,c_1)\le\psi(m_2,c_2)\le\cdots\le\psi(m_N,c_N)$$

则当 $i<j$ 时，(m_i,c_i) 和 (m_j,c_j) 是符合要求的一个明文对的充要条件是：$\psi(m_i,c_i)=\psi(m_{i+1},c_{i+1})=\cdots=\psi(m_j,c_j)$。

以下称 $\psi(m,c)$ 为明密对 (m,c) 的检测值。本文根据上述定理，基于快速排序思想，回避对密文对的检测，提出对一个结构产生的所有明文对进行筛选的快速方法，包括一个基本算法和一个改进算法。

在基本算法中，首先对一个结构中的明密对 (m,c) 按照 $\psi(m,c)$ 的大小进行排序，记录并输出使 $\psi(m,c)$ 连续相等的明密对片段的首序号和尾序号。根据定理，一个明文对满足筛选条件等价于这 2 个明密对位于某个 $\psi(m,c)$ 连续相等的片段之中。如果还需要将这些明密对显示显式地存储，只需根据这些连续片段的首序号和尾序号，将这些明文对全部挑选选中并存储即可。

当一个结构由 2^t 个明文对组成时，对该结构产生的所有明文对进行筛选的计算复杂性为 $O(2^t t)$，远远小于按定义直接检测的计算复杂性 2^{2t-1}。

设 $a_1,a_2,\cdots,a_n$ 是一个结构中的 n 个明密对，排序算法 Quicksort(a,k,m) Quicksort(a,k,m)*Quicksort*(a,k,m)根据 $\psi(a_i)$ 的大小对 $a_k,a_{k+1},\cdots,a_m$ 进行排序，则数据筛选基本算法步骤如下：

Step1 进行执行算法 Quicksort$(a,1,n)$ Quicksort$(a,1,n)$。

Step2 记录并输出排序后得到的 $a_1,a_2,\cdots,a_n$ 中长度大于等于 2 且 $\psi(a_i)$ 连续相等片段的首序号和尾序号。此时，每个首序号和尾序号之间任何 2 个序号对应的明密对都是满足要求的选择明文对，反之亦然。

快速排序算法 Quicksort(a,k,m) Quicksort(a,k,m) [5]步骤如下：

If $k<m$ k＜m then do
{
Split(a,k,m,i) Split(a,k,m,i);
Quicksort$(a,k,i-1)$ Quicksort(a,k,i-1)；;
Quicksort$(a,i+1,m)$ Quicksort(a,i+1,m)；;
}
End do

其中，k 和 m 的初始值分别为 1 和 n；分割算法 Split(a,k,m,i)*Split*(a,k,m,i) Split(a,k,m,i) 先选择一个整数 i，再对 $a_k,a_{k+1},\cdots,a_m$ 调换顺序，从而使下式成立：

$$\begin{cases}\psi(a_j)<\psi(a_i) & k\le j<i\\ \psi(a_i)\le\psi(a_j) & i<j\le m\end{cases}$$

Split(a,k,m,i) Split(a,k,m,i) 算法[5]具体步骤如下：

Step 1 $v\leftarrow0$ v←0；;
Step 2 $i\leftarrow k+v$ i←k+v; $x\leftarrow a_k$ x←aₖ; $j\leftarrow m$ j←m;
Step 3 Repeat follwings until $j=i$ j＝=i:
{
Repeat $j\leftarrow j-1$ j←j-1 until ψ(aⱼ)＜ψ(x) $\psi(a_j)<\psi(x)$;
Do $a_i\leftarrow a_j$ aᵢ←aⱼ;
Repeat i ← i+1 $i\leftarrow i+1$ until ψ (aᵢ) ≥ ψ (x) $\psi(a_i)\ge\psi(x)$;
Do aⱼ←aᵢ $a_j\leftarrow a_i$;
}
Step 4 if i＝=k $i=k$ do v←v+1 $v\leftarrow v+1$ and return to Step 2.
Else if i＞k $i>k$ Do aᵢ←x $a_i\leftarrow x$；;
Step 5 output i.

由于对 $n=2^t$ 个元素的快速排序算法的平均计算复杂性[6]为 $O(n\log n)=O(t2^t)$，记录连续片段的首序号和尾序号的计算复杂性为 $O(n)=O(2^t)$，因此基本算法的计算复杂性为 $O(t2^t)$，它远远小于按定义直接检测的计算复杂性 2^{2t-1}。此外，由于排序算法只需使用几个中间变量，并不增加额外的存储空间，因而因此存储复杂性仍为存储 n 个明密对的存储复杂性：$O(n)=O(2^t)$。

3 基于快速排序思想的碰撞搜索算法

在基本算法中，由于需要记录使 $\psi(a_i)$ 连续相等的片段的首序号和尾序号，因此需要在 Step2 中对 N 个已排好序的明密对重新检测一遍，即增加了 N 个比较。下面对该算法进行改进，将碰撞的挑选与快速排序算法有机地结合，同时进行排序和挑选，从而进一步降低数据筛选过程的计算量。

设 $a_1,a_2,\cdots,a_n$ 是一个结构中的 n 个明密对，碰撞搜索改进算法 Colision_finder(a,k,m) Collision_finder(a,k,m) 根据 $\psi(a_i)$ 的大小采用快速排序算法的原理，从 $a_k,a_{k+1},\cdots,a_m$ 中挑选出所有连续片段，其具体步骤如下：

If k＜m $k<m$ then Do
{
Split _and _choice(a,k,m,i) Split_and_choice(a,k,m,i) and obtain u;

Collision_finder$(a,k+u,i-1)$ Collision_finder(a,k+u,i-1)；;

Collision_finder(a,i+1,m) *Collision_finder*$(a,i+1,m)$ ；;

}

End do

其中，k 和 m 的初始值分别为 1 和 n；分割与挑选算法 Split_and_choice(a,k,m,i) Split_and_choice(*a*,*k*,*m*,*i*)*Split_and_choice*(a,k,m,i)选择一个整数 i，并对 $a_k,a_{k+1},\cdots,a_m$ 调换顺序，同时输出调整过程产生的使 $\psi(a_j)$ 与 $\psi(a_i)$ 相等的元素 a_j (共 u 个)，且使调整顺序后的 $a_{k+u},a_{k+u+1},\cdots,a_m$ 对任意 $k+u\le j_1<i$ 和 $i<j_2\le m$，都有 $\psi(a_{j_1})<\psi(a_{j_2})$。

Split_and_choice(*a*,*k*,*m*,*i*) Split_and_choice(a,k,m,i) 算法具体过程如下：

Step 1 $v\leftarrow 0$ v←0；; u←0 $u\leftarrow 0$ ；;

Step 2 i←k+v $i\leftarrow k+v$; j←m $j\leftarrow m$; $x\leftarrow a_i$ x←a_i;

Step 3 Repeat follwings until j＝i $j=i$:

{

Repeat j←j-1 $j\leftarrow j-1$ until ψ(a_j)＜ψ(x) $\psi(a_j)<\psi(x)$;

Do $a_i\leftarrow a_j$ a_i←a_j;

L: L: Repeat i ← i+1 $i\leftarrow i+1$ until ψ (a_i) ≥ ψ (x) $\psi(a_i)\ge\psi(x)$;

If ψ(a_i)≥ψ(x) $\psi(a_i)=\psi(x)$ do

{

output a_i a_i ;

u←u+1 $u\leftarrow u+1$;

If i＞k+v+u $i>k+v+u$, Do a_i←a_{k+u} $a_i\leftarrow a_{k+u}$;

Return to L;

}

Else if ψ (a_i) ≥ ψ (x) $\psi(a_i)>\psi(x)$ Do a_j ← a_i $a_j\leftarrow a_i$;

}

Step 4 If i＝k+v $i=k+v$ Do

{

$v\leftarrow v+1$ v←v+1；;

If ψ (a_{k+v}) ≥ ψ (a_{k+v-1}) $\psi(a_{k+v})=\psi(a_{k+v-1})$ Output a_{k+v-1} a_{k+v-1} and Do u←1 $u\leftarrow 1$;

Else Do u←0 $u\leftarrow 0$;

Return to Step 2 .

}

Else if i＞k+v $i>k+v$ Do

{

If u=0, Do a_i←x $a_i\leftarrow x$ ；;

output i and u

.

}

由于省略了基本算法中的 Step2，因此改进算法使碰撞搜索的计算复杂性减少了 $O(\log n)$。

下面对例 1 和例 2 中的有关数据进行分析，结果如表 1 所示。其中，基于改进算法的数据筛选的计算复杂性按“结构总数×$n$lb$n$”的方法计算，筛选过程计算复杂性的单位是比较次数。

表 1 例 1 和例 2 中数据筛选过程的计算复杂性

	结构中明文数	检测值总数	结构总数	计算复杂性		
				直接筛选	用改进算法筛选	密钥求解
例 1	2^{48}	2^{96}	2^{64}	2^{155}	$2^{117.6}$	2^{112} 次加密
例 2	2^{56}	2^{88}	2^{71}	2^{166}	$2^{132.8}$	2^{96} 次加密

由此可见，即使利用改进算法进行数据筛选，例 2 文献[3](即文献[3]例 2)提出的对 ARIA 的不可能差分攻击)中数据筛选的计算复杂性也是 $2^{132.8}$ 次比较，远大于密钥求解的计算复杂性，也大于文献[2]提出的攻击方法的计算复杂性。然而，例 1 文献[2](即文献[2]例 1) 提出的对 ARIA 的不可能差分攻击)中数据筛选的计算复杂性都小于密钥求解的计算复杂性。

文献[2]还提出了对 12 轮 Camellia 的不可能差分攻击，文献[4]提出了对低轮 CLEFIA 的不可能差分攻击。本文通过对其中几个不可能攻击数据筛选过程的计算复杂性进行分析，结果如表 2 所示，可以看出发现，其计算复杂性都不高于密钥求解的计算复杂性。

表 2 对 Camellia 和 CLEFIA 不可能差分攻击中数据筛选的复杂性

攻击类型	密码算法	结构中明文数	结构总数	计算复杂性		
				直接筛选	用改进算法筛选	密钥求解
文献[2]	12 轮 Camellia	2^{112}	2^{8}	2^{231}	$2^{126.8}$	2^{181} 次加密
文献[4]	12 轮 CLEFIA	2^{32}	$2^{78.93}$	$2^{205.93}$	$2^{115.93}$	2^{111} 次加密
文献[4]	13 轮 CLEFIA	2^{64}	$2^{47.73}$	$2^{174.73}$	$2^{117.73}$	2^{158} 次加密
文献[4]	14 轮 CLEFIA	2^{64}	$2^{48.23}$	$2^{175.23}$	$2^{118.23}$	2^{199} 次加密
文献[4]	15 轮 CLEFIA	2^{113}	2	2^{226}	$2^{120.8}$	2^{248} 次加密

4 结束语

本文研究了不可能差分攻击中数据筛选的方法问题，并基于快速排序思想，提出了对明密对筛选的有效算法，并据此指出文献[3]提出的不可能差分攻击中数据筛选的计算复杂性远大于密钥求解的计算复杂性。由于本文数据筛选算法的本质是根据具体的检测值找出所有的碰撞，因此算法还可直接应用于碰撞的搜索问题。

参考文献

[1] Biham E, Biryukov A, Shamir A. Cryptanalysis of Skipjack Reduced to 31 Rounds Using Impossible Differentials[C]//Proc. of EUROCRYPT'99. [S. l.]: Springer-Verlag, 1999.

[2] Wu Wenling, Zhang Wentao, Feng Dengguo. Impossible Differential Cryptanalysis of Reduced Round ARIA and Camellia[J]. Journal of Computer Science Technology, 2007, 22(3): 453-460.

[3] Li Shenhua, Song Chunyan. Improved Impossible Differential Cryptanalysis of ARIA[C]//Proc. of International Conference on Information Security and Assurance. Busan, Korean: [s. n.], 2008.

[4] Sun Bing. Impossible Differential Cryptanalysis of CLEFIA[Z]. [2008-12-11]. http://www.eprint/2008/151/pdf.

[5] Wang Wei, Wang Xiaoyun. Improved Impossible Differential Cryptanalysis of CLEFIA[Z]. [2009-01-11]. http://www.eprint/2007/466/pdf.

[6] Baase S. 计算机算法: 设计和分析引论[M]. 朱 洪, 译. 上海: 复旦大学出版社, 1985.

编辑 张 帆

(上接第 126 页)

5 结束语

随着异构网络间切换对安全性的要求越来越高，切换中认证机制的运行需要花费更多的时间。本文通过基于网络层 PANA 认证架构，采用预认证与 AAT 证书联合的方法，有效地克服了这一矛盾和不同管理域间切换与单一域内切换时延

差距大的问题。

参考文献

[1] Tuladhar S R, Caicedo C E, Joshi J B D. Inter-domain Authentication for Seamless Roaming in Heterogeneous Wireless Networks[C]//Proc. of EEE International Conference on SUT Capos. Sensor Networks, Ubiquitous and Trustworthy Computing. [S. 1.]: IEEE Press, 2008: 249-255.

[2] Gaabab B, Binet D, Bonnin J M. Authentication Optimization for Seamless Handovers[C]//Proc. of the 10th IFIP/IEEE International Symposium on Integrated Network Management. Munich, Germany: [s. n.], 2007: 829-832.

[3] Forth R. AAA Context Transfer for Fast Authenticated Inter-domain Handover[Z]. (2003-04-07). http://rangiroa.essi.fr/DEA-RSD/rapports/ruckforth.pdf.

[4] Aura T, Roe M. Reducing Reauthentication Delay in Wireless Networks[C]//Proc. of the 1st International Conference on Security and Privacy for Emerqing Areas in Communications Networks. NY, USA: [s. n.], 2005: 139-148.

[5] 伍华凤, 戴新发, 陈　鹏. 一种层次化移动 IP 接入认证机制[J]. 计算机工程, 2008, 34(24): 131-133.

[6] Lopez R M, Dutta A, Ohba Y, et al. Network-layer Assisted Mechanism to Optimize Authentication Delay During Handoff in 802.11 Networks[C]//Proc. of the 4th Annual International Conference on Mobile and Ubiquitous Systems: Networking & Services. Philadelphia, PA, USA: [s. n.], 2007: 1-8.

编辑　索书志