第 20 卷 第 4 期 电子科学学刊 Vol.20 No.4
1998 年 7 月 JOURNAL OF ELECTRONICS July 1998

杂凑算法的对偶问题 [1]

朱华飞 杨 波 王新梅 肖国镇
(西安电子科技大学信息保密研究所 西安 710071)

摘 要 杂凑算法往往具有很高的杂凑速率，如 MD_x(x 版本的杂凑算法), SHA(数据杂凑标准) 等. 一个自然的问题是能否利用快速安全的杂凑算法构造相应的分组加密算法呢？本文把这个问题称为杂凑算法的对偶问题. 本文试图用现有的结果证明给定一个安全快速的杂凑算法可以构造一个安全快速的分组加密算法.

关键词 杂凑算法，分组加密算法，对偶问题

中图号 TN918.1

1 引 言

我们知道杂凑算法的构造一般是通过以下两条途径实现的. 一是利用现有的安全分组加密算法如 DES(数据加密标准), IDEA(数据加密算法), HAVAL (可变长度杂凑算法) 等, 通过当前密文组与下一组明文的与或迭代加密得到消息 M 的杂凑值 $H(M)$, 基于分组密码迭代的杂凑算法的安全性是以所选择的分组密码的安全性为尺度. 另一种是不基于任何假设条件，通过简单的布尔运算，小整数取模运算构造杂凑算法，该类杂凑算法往往具有很高的杂凑速率如 MD_x(Message Digest version x), SHA(Standard Hash Algorithm) 等. 一个自然的问题是能否利用快速安全的杂凑算法构造相应的分组加密算法呢？我们把这个问题称为杂凑算法的对偶问题. 本文试图对这个问题的研究作些努力.

2 关于杂凑算法对偶问题的讨论

我们知道杂凑算法是不可逆的，而分组加密算法是可逆的. 利用杂凑算法构造分组加密算法，首先要解决如何处理初始值的设定问题. 这是因为给定一个 $GF(2)^m$ 到 $GF(2)^n$ 上的杂凑函数 h, 若赋以不同的初始值 IV, 导致对同一明文 M 产生不同的杂凑值. 因此，如果不给出一个杂凑初始值的编码规则，我们便无法构造一个可逆的加密算法. 按照对初始值设定的不同原则，把杂凑算法的对偶性问题分成三类：第一型对偶算法，第二型对偶算法及第三型对偶算法. 为了讨论问题方便，不妨认为 $m = 4n$, 现描述一个明文分组 $M = (L, R)$, $|L| = |R| = n$ 的加密 - 解密全过程.

第一型对偶算法设 $K_1, K_2, K_3, K_4 = K_1 \oplus K_2 \oplus K_3$ 是通信双方在每次通信前预先选定的秘密密钥. 记

$$f_1(\bullet) = f(K_1||K_2||\bullet||K_3), \quad f_2(\bullet) = f(K_2||K_3||\bullet||K_1),$$
$$f_3(\bullet) = f(K_3||K_1||\bullet||K_2), \quad f_4(\bullet) = f(K_4||\overline{K_2} \oplus K_1||\bullet||K_3 \oplus K_1),$$
$$\psi(f)(L, R) = (R, L \oplus f(R)).$$

[1] 1997-01-15 收到， 1998-01-04 定稿
国家自然科学基金资助课题

加密过程

$$\begin{aligned} E_1 &\quad \text{计算} \quad \psi(f_1)(L,R)=(R,L\oplus f_1(R)).\\ E_2 &\quad \text{计算} \quad \psi(f_2,f_1)(L,R)=\psi(f_2)\circ\psi(f_1)(L,R).\\ E_3 &\quad \text{计算} \quad \psi(f_3,f_2,f_1)(L,R)=:(S,T).\\ E_4 &\quad \text{计算} \quad \psi(f_4,f_3,f_2,f_1)(L,R)=:(S,T). \end{aligned}$$

解密过程

$$\begin{aligned} D_1 &\quad \text{计算} \quad \sigma(S,T)=(T,S).\\ D_2 &\quad \text{计算} \quad \sigma\circ\psi(f_3,f_2,f_1)\circ\sigma(S,T)=(L,R).\\ D_3 &\quad \text{计算} \quad \sigma\circ\psi(f_4,f_3,f_2,f_1)\circ\sigma(S,T)=(L,R). \end{aligned}$$

利用 Luby & Rackoff[1] 的伪随机理论，我们有

定理 1 若加密算法为 E_1-E_3，解密算法为 D_1-D_2，则第一型对偶算法是 $O(2^{n/2})$ 阶伪随机的．即该分组加密算法能够抵抗 $O(2^{n/2})$ 阶选择明文攻击。

定理 2 若加密算法为 E_1-E_4，解密算法为 D_1-D_3，则第一型对偶算法是 $O(2^{n/2})$ 阶超伪随机的．即该分组加密算法能够抵抗 $O(2^{n/2})$ 阶选择明文－密文或密文－明文攻击。

第二型对偶算法记 $f_1(\bullet)=f(K_1||K_2||\bullet||K_3)$，$f_2(\bullet)=f(K_2||K_3||\bullet||K_1)$，明文 $M=(L,R)$ 的加密过程

$$\begin{aligned} E_1 &\quad \text{计算} \quad \psi(f_1,f_1)(L,R).\\ E_2 &\quad \text{计算} \quad \psi(f_2,f_1,f_1)(L,R)=:(S,T).\\ E_3 &\quad \text{计算} \quad \psi(f_2,f_2,f_1,f_1)(L,R)=:(S,T). \end{aligned}$$

解密过程

$$\begin{aligned} D_1 &\quad \text{计算} \quad \sigma(S,T)=(T,S).\\ D_2 &\quad \text{计算} \quad \sigma\circ\psi(f_2,f_1,f_1)\circ\sigma(S,T)=(L,R).\\ D_3 &\quad \text{计算} \quad \sigma\circ\psi(f_2,f_2,f_1,f_1)\circ\sigma(S,T)=(L,R). \end{aligned}$$

利用 Luby & Rackoff[1] 及 J. Patarin[2,3] 结果，我们有

定理 3 若加密算法为 E_1-E_2，解密算法为 D_1-D_2，则第二型对偶算法是 $O(2^{n/2})$ 阶伪随机的，即该分组加密算法能够抵抗 $O(2^{n/2})$ 阶选择明文攻击。

定理 4 若加密算法为 E_1-E_3，解密算法为 D_1-D_3，则第二型对偶算法是 $O(2^{n/2})$ 阶超伪随机的．即该分组加密算法能够抵抗 $O(2^{n/2})$ 阶选择明文－密文或密文－明文攻击。

第三型对偶算法记 $f=f(K_1||K_2||\bullet||K_3)$，明文 $M=(L,R)$ 的加密过程：

$$\begin{aligned} E_1 &\quad \text{计算} \quad \psi(f,f,f)(L,R).\\ E_2 &\quad \text{计算} \quad \psi(f^2,f,f,f)(L,R)=:(S,T).\\ E_3 &\quad \text{计算} \quad \psi(f,f^2,f,f,f)(L,R)=:(S,T). \end{aligned}$$

解密过程

$$\begin{aligned} D_1 &\quad \text{计算} \quad \sigma(S,T)=(T,S).\\ D_2 &\quad \text{计算} \quad \sigma\circ\psi(f^2,f,f,f)\circ\sigma(S,T)=(L,R).\\ D_3 &\quad \text{计算} \quad \sigma\circ\psi(f,f^2,f,f,f)\circ\sigma(S,T)=(L,R). \end{aligned}$$

利用 J. Patarin[2,3] 及文献 [4] 的结果，我们有下面的结果：

定理 5　若加密算法为 E_1-E_2, 解密算法为 D_1-D_2, 则第三型对偶算法是 $O(2^{n/2})$ 阶伪随机的，即该分组加密算法能够抵抗 $O(2^{n/2})$ 阶选择明文攻击。

定理 6　若加密算法为 E_1-E_3, 解密算法为 D_1-D_3, 则第三型对偶算法是 $O(2^{n/2})$ 阶超伪随机的，即该分组加密算法能够抵抗 $O(2^{n/2})$ 阶选择明文－密文或密文－明文攻击。

上述结果是在 $m=4n$ 的假设下得到的，对 $m=kn$ 的情形，不难按前面的思想同样考虑，我们不再赘述。

3　杂凑算法对偶问题的实例

设通信双方 A, B 在每一次通信前预先选定秘密密钥 $K=K_1||K_2||K_3$, $|K_i|=128$bit, $(i=1,2,3)$; f 为 MD_4(或 MD_5); IV_0 表示 f 的杂凑初始值。要加密的明文 M, 经过 Merkle-Damgaard 强化后，记为 $M=M_1\cdots M_n$, $M_i=(L_i,R_i)$, $|L_i|=|R_i|=128$bit, $(i=1,\cdots,n)$。现描述明文 M 的一个分组 (L,R) 的加密过程：

步骤 1　f 的 IV_0 置为 K_1, 相应于初始值 K_1 的 f 记为 f_1, 并计算：$\psi(f_1)(L,R)=(R,\ L\oplus f_1(K||R))$。

步骤 2　f_1 经过一次加密运算后，IV_0 由 K_1 映射成 $f_1(K||R)$。利用代换置第二轮的杂凑初始值 IV_1 为 K_2, 相应的 f 记为 f_2, 并计算：$\psi(f_2)\circ\psi(f_1)(L,R)=\psi(f_2)(R,L\oplus f_1(K||R))=(L\oplus f_1(K||R),\ R\oplus f_2(K||(L\oplus f_1(K||R))))$。

步骤 3　类似于步骤 2, 把 IV_2 置为 K_3。记 $f=f_3$, 并计算 $\psi(f_3)\circ\psi(f_2)\circ\psi(f_1)(L,R)=(S,T)$, 称 (S,T) 为明文 $M=(L,R)$ 的相应密文。

注意到：若 $\psi(f)(L,R)=(S,T)$, 则 $\psi^{-1}(f)(S,T)=\sigma\circ\psi(f)\circ\sigma(S,T)=(L,R)$, 其中 $\sigma(L,R)=(R,L)$。因此，由 (S,T) 利用与加密运算同样的步骤(仅仅顺序相反)。由 (S,T) 可得到 (L,R), 即 $[\psi(f_3)\circ\psi(f_2)\circ\psi(f_1)]^{-1}(S,T)=\sigma\circ\psi(f_1)\circ\psi(f_2)\circ\psi(f_3)\circ\sigma(S,T)=(L,R)$。我们把上述的加密方法称为杂凑分组加密算法 (Hash Block Enciphering Algorithm-HBEA)。

由 MD_4 的结构及算法的特点知道：对于固定的 $K=K_1||K_2||K_3$, 由指定的杂凑初始值 K_i, $f_i(K||x)$ 可视为 $\sum^{128}$ 到 $\sum^{128}$ 上的关于变量 x 的置换。因此，文中描述的分组密码可以换为这样的说法：通信双方预先选定三个置换 (f_1,f_2,f_3) 作为秘密密钥，$f_i\in F_n(\sum^n$ 到 $\sum^n$ 上所有映射的集合)。注意到：$\psi(f)$ 是 $\sum^{2n}$ 到 $\sum^{2n}$ 上的置换，因此：$\psi(f_1,f_2,f_3)=\psi(f_3)\circ\psi(f_2)\circ\psi(f_1)\in P_{2n}$ (P_n 为 $\sum^n$ 到 $\sum^n$ 所有置换的集合)。应用 Luby, Rackoff 所证明的结果，并令 $n=128$, $m=64$ 知道：$\psi(f_1,f_2,f_3)$ 是 $O(2^{64})$ 伪随机的。因此，我们有：

定理 7　基于 Feistel 型置换的三轮的 HBEA 能够抵抗 $O(2^{64})$ 阶的选择明文－密文攻击。

若我们在步骤 1— 步骤 3 的基础上再附加一轮杂凑值的运算：

步骤 4　令 $K_4=\overline{K_1}\oplus K_2\oplus\overline{K_3}$, 并把 IV_3 置为 K_4, 记 $f=f_4$, 并计算：$\psi(f_1,f_2,f_3,f_4)(L,R)=\psi(f_4)\circ\psi(f_1,f_2,f_3)(L,R)=(S,T)$, 得到明文 (L,R) 的相应密文 (S,T)。解密过程只要作相应的四轮逆序运算即可。注意到 MD_4 具有良好的随机性。因此，当 (K_1,K_2,K_3,K_4) 保密时，可视 (f_1,f_2,f_3,f_4) 是随机选自 F_n 上的。由文献 [3] 知：$\psi(f_1,f_2,f_3,f_4)$ 是 $O(2^{64})$ 阶超伪随机的。因此，我们有

定理 8 基于 Feistel 型置换的四轮的 HBEA 能够抵抗 $O(2^{64})$ 阶的选择明文 - 密文和选择密文 - 明文攻击。

4 结 论

我们在 MD_4 (MD_5) 具有良好密码学特性的基础上，利用三轮 (四轮)Feistel 型置换构造 $O(2^{64})$ 阶的伪随机 (超伪随机) 分组密码。由于 MD_4 (MD_5) 是高速率的杂凑算法而且 MD_4 (MD_5) 并没有 DES 或 IDEA 那样受专利的限制，任何人只要愿意，便能使用该杂凑算法。因此，无论从算法的安全性，快速性还是经济效益的角度考虑，该算法具有明显的优越性。这也是我们对该算法的应用前景持乐观态度的理由。

参 考 文 献

[1] Luby M, Rackoff C. How to construct pseudorandom permutations from pseudorandom functions. SIAM Journal on Computing, 1988, 17(2): 373–386.

[2] Patarin J. New results on pseudorandom permutation generators based on the DES Scheme, Abstracts of Crypto'91, Santa Barbara, CA, USA: 1991, 72–77.

[3] Patarin J. How to construct pseudorandom and super pseudorandom permutations from one single pseudorandom function. In Abstracts of Eurocrypt'92, Balatonfured, Hungary: 1992, 235–245.

[4] 朱华飞. 密码安全杂凑算法的设计与应用: [博士论文]. 西安: 西安电子科技大学, 1996 年 10 月.

DUAL PROBLEM OF HASH ALGORITHM

Zhu Huafei Yang Bo Wang Xinmei Xiao Guozhen

(*Xidian University, Xi'an* 710071)

Abstract Hash algorithm is always with high operation speed such as MD_x, SHA. A natural problem is that could one apply for a fast hash scheme to construct a cipher block algorithm. This paper denotes such a problem as dual problem of hash algorithm. Based on the known result, it is proved that several fast secure block cipher can be constructed if a fast secure hash algorithm is given.

Key words Hash algorithm, Cipher block algorithm, Dual problem

朱华飞: 男, 1966 年生, 博士后, 主要从事密码算法的设计与安全性分析.

杨 波: 男, 1963 年生, 博士生, 主要从事密码算法的设计与安全性分析.

王新梅: 男, 1937 年生, 教授, 博士生导师, 主要从事编码理论和密码算法的设计与安全性分析的教学与研究.

肖国镇: 男, 1935 年生, 教授, 博士生导师, 主要从事密码算法的设计与安全性分析.