ISSN 1000-9825, CODEN RUXUEW
Journal of Software, Vol.18, No.7, July 2007, pp.1756−1764
DOI: 10.1360/jos181756

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel/Fax: +86-10-62562563

一种可并行的消息认证码*

王大印[1,2+], 林东岱[1], 吴文玲[1]

[1](中国科学院 软件研究所 信息安全国家重点实验室,北京 100080)

[2](中国科学院 研究生院,北京 100049)

A Parallelizable Message Authentication Code

WANG Da-Yin[1,2+], LIN Dong-Dai[1], WU Wen-Ling[1]

[1](State Key Laboratory of Information Security, Institute of Software, The Chinese Academy of Sciences, Beijing 100080, China)

[2](Graduate University, The Chinese Academy of Sciences, Beijing 100049, China)

+ Corresponding author: Phn: +86-13381031335, E-mail: wdy@is.iscas.ac.cn

Wang DY, Lin DD, Wu WL. A parallelizable message authentication code. *Journal of Software*, 2007,18(7): 1756−1764. http://www.jos.org.cn/1000-9825/18/1756.htm

Abstract: This paper defines and analyzes a fully deterministic parallelizable block-cipher mode of operation for message authentication — DPMAC (deterministic parallelizable message authentication code). DPMAC is constructed based on a 128-bit block cipher, works for strings of any bit length, and employs a single block-cipher key. Its security is proved, using the Game-Playing technique to quantify an adversary's forgery probability in terms of the quality of the block cipher as a pseudo-random permutation.

Key words: message authentication code; pseudo-random permutation; provable security; Game-Playing

摘 要: 提出并分析了一种确定的、可并行的消息认证码——DPMAC(deterministic parallelizable message authentication code).它基于分组长度为 128-bit 的分组密码来构造.使用一个密钥,可以处理任意长度的消息.在底层分组密码是伪随机置换的假设下,使用 Game-Playing 技术量化了攻击者成功伪造的概率,从而证明了其安全性.

关键词: 消息认证码;伪随机置换;可证明安全;Game-Playing

中图法分类号: TP309 文献标识码: A

消息完整性是信息安全中最基本的需求,它的目的是为了防止消息在传输或存储过程中被未经授权地篡改.而使用对称密码学来保证消息完整性的基本工具就是消息认证码(message authentication codes,简称 MACs).

消息认证码通过以下机制来保护消息的完整性:首先,在参与通信的两方之间共享一个密钥,通信时(这里使用 *A* 和 *B* 代表参与通信的两方),*A* 传送一个消息给 *B*,并将这一消息使用 MAC 算法和共享密钥计算出一个值,这个值称为认证标记;然后,将这一值附加在该消息之后传送给 *B*.*B* 在接收后,使用同样的机制计算接收的消息的认证标记,并与其接收到的标记进行比较:如果这两个标记相同,*B* 就认为消息在由 *A* 传送到 *B* 的过程中没

* Supported by the National Natural Science Foundation of China under Grant Nos.60373048, 90604036 (国家自然科学基金); the National Basic Research Program of China under Grant No.2004CB318004 (国家重点基础研究发展计划(973))

Received 2006-01-20; Accepted 2006-04-26

有被篡改;如果不相同,B 就认为消息在传送过程中被篡改了.在 A 和 B 采用以上算法通信的过程中,攻击者可以自由地窃听 A 和 B 之间通信的信道,并且可以自由地修改或创建信道中传送的消息,如果攻击者能够成功地使 B 相信一个经过他篡改的消息(该消息不曾被 A 传送过)来自于 A,就称攻击者进行了成功的伪造.好的消息认证码算法应该使这种被伪造的概率尽可能地小.消息认证码的安全性定义也就来自于此.

许多常见的消息认证码,像 HMAC(hash based message authentication code)[1]和 CBC MAC(cipher block chaining message authentication code)[2]本质上都是串行的,也就是说,必须在处理完前 i–1 个消息块后才能处理第 i 个消息块.随着科技的发展,人们对加密硬件的处理速度上的要求越来越高,而且,新型的处理器对并行计算的支持度也越来越大.所以,人们对能支持并行处理的消息认证码的兴趣越来越浓,陆续地出现了一些可并行的消息认证码[3].下面简要介绍一下本文所提出的 DPMAC(deterministic parallelizable message authentication code)以及其他几种有代表性的可并行的消息认证码的特点,并对它们进行比较.

XOR MAC(exclusive or message authentication code)[4].Bellare 等人提出的这个可并行的消息认证码需要为每一个分组增加一个索引信息,这样就引起了分组密码调用次数的成倍增加,而且,该算法需要维持一个随机值或者一个状态值,于是增加了系统的计算量和通信量.

PMAC(parallelizable message authentication code)[5].这种算法是由 Black 等人提出来的.它是一种确定性的可并行的算法,而且调用分组密码的次数做到了最少.它是一个不错的消息认证码,但是这种算法受专利保护,不能免费使用.

XECB MAC(extended electronic codebook message authentication code)[6].这种算法是由 Gligor 等人提出来的,它借鉴了 XOR MAC 的构造方法,避免了为消息引入索引信息,减少了分组密码调用次数.与 XOR MAC 一样,该算法需要维持一个随机值或一个状态值.

本文所提出的 DPMAC 是一种确定性的可并行的算法.它不需要维持一个随机值或一个状态值,性能明显优于 XOR MAC,略高于 XECB MAC,但略低于 PMAC.不过与 PMAC 相比,该算法不受专利保护.以下将给出该算法的具体构造方法及其安全性证明.

1 预备知识

在给出 DPMAC 的定义和证明之前,首先介绍可证明安全理论中的一些符号和定义.

1.1 一些符号和定义

分组密码是这样一个函数:E:k×{0,1}^{n}→{0,1}^{n},其中:k 是有限集,每个 $E_K(\cdot)=E(K,\cdot)$是一个在集合{0,1}^{n} 上的置换.

$Perm(n)$表示{0,1}^{n}上所有置换的集合.这个集合也可以看成是所有的由集合 k 中元素唯一确定的置换的集合. $\pi \xleftarrow{\$} Perm(n)$ 表示随机地从集合 $Perm(n)$中选取一个置换.

在可证明安全理论中,一般都使用优势函数来度量一种算法与一种理想的算法之间的差别.如果这个差别可以忽略,就认为该算法是安全的.这里给出几个优势函数的定义.

定义 **1**. 假定 A 是一个拥有 Oracle 的攻击者,A^o 表示 A 可以查询 Oracle o.不失一般性,假定攻击者从来不查询该 Oracle 定义域之外的值,并且从来不重复查询已经查询过的值.那么在进行了一定数量的查询之后,攻击者 A 输出一个值,这个值只能是 0 或 1,定义:

$$Adv_E^{prp}(A)=\Pr[K\xleftarrow{\$}k:A^{E_K(\cdot)}=1]-\Pr[\pi\xleftarrow{\$}Perm(n):A^{\pi(\cdot)}=1].$$

该式表示在经过了一定数量的查询之后,当给定的 Oracle 为 $E_K(\cdot)$时,攻击者 A 输出 1 的概率与当给定的 Oracle 为$\pi(\cdot)$时攻击者 A 输出 1 的概率之差.其中:K 随机地从集合 k 中选取;π随机地从置换族 $Perm(n)$中选取.该式度量了分组密码 E:k×{0,1}^{n}→{0,1}^{n} 和随机置换之间区分的概率.在安全性证明中,一般都假设算法中所使用的分组密码的 $Adv_E^{prp}(A)$ 是一个可以忽略的值.

类似地,定义一个从{0,1}^{n} 到{0,1}^{n} 的函数族 F:k×{0,1}^{n}→{0,1}^{n},其中,k 是有限集.这里写作 $F_K(\cdot)=F(K,\cdot)$.

$Rand(n)$表示所有$\{0,1\}^n\to\{0,1\}^n$的函数的集合.定义:

$$Adv_F^{prf}(A)=\Pr[K\xleftarrow{\$}k:A^{F_K(\cdot)}=1]-\Pr[\rho\xleftarrow{\$}Rand(n):A^{\rho(\cdot)}=1].$$

该式度量了函数 $F:k\times\{0,1\}^n\to\{0,1\}^n$和随机函数之间区分的概率.

类似地,如果 $f:k\times\{0,1\}^{n^*}\to\{0,1\}^l$,其中:$\{0,1\}^{n^*}$表示长度是 n 的整数倍的所有二进制串的集合;k 是一个具有一定分布的集合.那么,f 就是一个从$\{0,1\}^{n^*}$到$\{0,1\}^l$的函数族.这里使用 $f_K(\cdot)$表示 $f(K,\cdot)$.

$Rand(n^*,l)$表示所有从$\{0,1\}^{n^*}$到$\{0,1\}^l$ 的函数的集合.也就是说,对于任意的随机函数$\lambda\in Rand(n^*,l)$,每一个$x\in\{0,1\}^{n^*}$,$\lambda(x)$的值都是均匀地从$\{0,1\}^l$ 中选取的长度为 l 位的串.具体来说,就是对于任何的$\lambda\in Rand(n^*,l)$,$x\in\{0,1\}^{n^*}$,$y\in\{0,1\}^l$,$\Pr[\lambda(x)=y]=2^{-l}$.定义:

$$Adv_f^{prf}(A)=\Pr[K\xleftarrow{\$}k:A^{f_K(\cdot)}=1]-\Pr[\lambda\xleftarrow{\$}Rand(n^*,l):A^{\lambda(\cdot)}=1].$$

该式度量了攻击者 A 区分随机函数 f 和 $Rand(n^*,l)$的概率.以上这些定义来自文献[5,7].

1.2 消息认证码的安全性定义

消息认证码的安全性是通过消息认证码不可伪造的概率定义的.文献[7]中消息认证码一章将消息认证码的不可伪造性使用优势函数 Adv_{MAC}^{uf-cma} 来度量,并且给出了以下定义:

定义 **2**. MAC 是一个消息认证码,A 是对 MAC 进行伪造的攻击者,B_A是区分 MAC 和 $Rand(n^*,l)$的攻击者,那么有

$$Adv_{MAC}^{prf}(B_A)\ge Adv_{MAC}^{uf-cma}(A)-\frac{1}{2^l}.$$

从上式中可以知道:如果要证明消息认证码的不可伪造性,除了直接求 $Adv_{MAC}^{uf-cma}(A)$ 之外,也可以求消息认证码区分于随机函数的优势 $Adv_{MAC}^{prf}(B_A)$.只要该优势可以忽略,那么 $Adv_{MAC}^{uf-cma}(A)$ 也就是可忽略的.文献[2,8]中也指出了这一点.我们下面的证明就是从这个角度着手的.

2 DPMAC 的定义

DPMAC 是一种确定性的消息认证码.它不需要随机值或维持一个状态,而且可支持并行处理.

消息格式与符号定义:DPMAC 所使用的底层的分组密码的分组长度是 128-bits.为描述方便,用 $E:k\times\{0,1\}^n\to\{0,1\}^n$来表示,这里 n=128,若无特别说明,下文中 n 的取值均为 128.

M 表示消息,$|M|$表示消息的长度.$Pad(M)$表示通过填充将消息填充为分组密码分组长度 n 的整数倍(比如,可以在消息最后添加 1,然后补上若干个 0),这样,消息 M 就可以看成 $M=M[1]M[2]\ldots M[m]$,其中$|M[i]|=n$,$i=1,\ldots,m$.$||M||$表示消息被分成的块数.

设 L,M,i 均是长度为 128-bits 的任意的无符号整数串,其中,$i>0$.定义函数:

$$H(L,M,i)=i\times L \bmod (2^{128}+51)+M \bmod 2^{128}.$$

那么,DPMAC 算法如图 1 所示(其中,M 是要处理的消息,K 是分组密码的密钥).

Algorithm $\mathrm{DPMAC}_K(M)$
1. $L\leftarrow E_K(0^n)$
2. $pad(M)$
3. Partition M into $M[1]\ldots M[m]$
4. for $i\leftarrow 1$ to m do
5. $\quad X[i]\leftarrow H(L,M[i],i)$
6. $\quad Y[i]\leftarrow E_K(X[i])$
7. $\Sigma\leftarrow Y[1]\oplus Y[2]\oplus\ldots\oplus Y[m]$
8. $Tag=E_K(\Sigma)$[first l bits]
9. return Tag

Fig.1 Definition of DPMAC

图 1 DPMAC 的定义

3 定 理

这里,我们给出 DPMAC 在信息论中安全性的界,也就是把分组密码理想化为一个随机置换.

定理 **1**(**DPMAC** 的安全性). A 是一个拥有 Oracle 的攻击者,假设 A 进行了 q 次查询,总的块数为σ,那么有

$$Adv^{prf}_{\mathrm{DPMAC}[Perm(n),n]}(A)\le\frac{2(\sigma+2q)^2}{2^{126}},$$

其中,攻击者 A 要查询的消息 $M_1,\ldots,M_q$ 的总长度 $\sigma=\sum_{r=1}^{q}\|M_r\|$,$\|M\|$的定义如前所述.从上面的定理我们可以很容易地将其过渡到复杂性理论中.分组密码 $E:k\times\{0,1\}^n\to\{0,1\}^n$.$A$ 是拥有 Oracle 的攻击者,假设 A 进行了 q 次查询,总的块数为σ.那么,存在攻击 E 的攻击者 B,获得的优势 $Adv^{prp}_E(B)\ge Adv^{prf}_{DPMAC[perm(128),128]}(A)-\frac{2(\sigma+2q)^2}{2^{126}}$,攻击者 B 最多进行σ+q+1 次查询,运行的时间等于 A 运行的时间加上在σ+q+1 个点上计算 E 的时间.

4 定理证明

这里,我们使用文献[9,10]中的 Game-Playing 技术并借鉴文献[5,11]中的证明方法来证明定理 1.

由于我们采用的填充方式是一种一一映射,所以,在下面的证明中仅考虑扩展后的消息.

如果 A 是一个攻击 DPMAC[$Perm$(n),n]的攻击者,由于 A 具有无限的计算能力,不失一般性,假定 A 是确定性的.如图 2 所示的 Game 1 准确地模拟了 DPMAC[$Perm$(n),n].而 Game 0 准确地模拟了 $Rand$(n*,n).那么,攻击者 A 攻击的过程就可以看成是 A 和 Game1,Game0 交互的过程(将 Game 1 中方框内的语句删除后形成的 Game 称为 Game 0).

Game1
simulation DPMAC
Initialization
10 $L\xleftarrow{\$}\{0,1\}^n;\pi(0^n)\leftarrow L$
When A makes its $r-th\ query, M_r=M_r[1]\ldots M_r[m_r]$ where $r\in[1,\ldots,q]$
20 for $i\leftarrow 1$ to m_r do
21 $\{X_r[i]\leftarrow H(L,M_r[i],i)$
22 if $X_r[i]\in Domain(\pi)$ then $Y_r[i]\leftarrow\pi(X_r[i])$
23 else $\{Y_r[i]\xleftarrow{\$}\{0,1\}^n$
24 if $Y_r[i]\in Range(\pi)$ then {bad $\leftarrow$ true; $\boxed{Y_r[i]\xleftarrow{\$}\overline{Range(\pi)}}$}
25 $\pi(X_r[i])\leftarrow Y_r[i]\}$
26 $\Sigma_r\leftarrow Y_r[1]\oplus Y_r[i]\oplus\ldots Y_r[m_r]$
27 $Tag_r\xleftarrow{\$}\{0,1\}^n$
28 if $\Sigma_r\in Domain(\pi)$ then {bad $\leftarrow$ true; $\boxed{Tag_r\leftarrow\pi(\Sigma_r)}$}
29 if $Tag_r\in Range(\pi)$ then {bad $\leftarrow$ true; $\boxed{Tag_r\xleftarrow{\$}\overline{Range(\pi)}}$}
30 $\pi(\Sigma_r)\leftarrow Tag_r$
31 return Tag_r

Fig.2 Game1

图 2 Game1

从图 2 中可以看出:攻击者 A 在与 Game 1 进行交互时,A 的 Oracle 是 DPMAC[$Perm$(n),n];而在 A 与 Game 0 进行交互时,A 的 Oracle 是 $Rand$(n*,n).由于这两个 Game 仅在 bad 设置为 true 时不同,所以,由文献[9]中引理 5 和文献[10]中的引理 1 可知,

$$Adv^{prf}_{\mathrm{DPMAC}[Perm(n),n]}(A)\le\Pr_0[\mathrm{bad}=\mathrm{true}]=\Pr_1[\mathrm{bad}=\mathrm{true}],$$

其中,下标 0 和 1 标明所对应的 Game.由于求 Game 0 中 bad 设置为 true 的概率相对简单一些,因此,这里界定

$\Pr_0[\text{bad} = \text{true}]$这个概率.

我们首先考虑在 24 行和 29 行 bad 设置为 true 的概率.在这两种情况下,相当于每次随机地选择一个长度为 n 位的串,然后测试其是否在一个集合里.这个集合起始时有一个元素 0^n(因为第 10 行),每次增加一个随机元素,直到元素增加到σ+q,那么就有

$$\Pr_0[\text{bad} = \text{true in lines 24 or 29}] \le \frac{1+2+...(\sigma+q)}{2^n} \le \frac{(\sigma+q+1)^2}{2^{n+1}} \quad (1)$$

这里的下标 0 标明是在 Game 0 中的概率.

现在改变 Game 0,将 24 行和 29 行从 Game 0 中删去.将 Game 0 改写为 Game 2.如图 3 所示.

Game 2
simulation DPMAC
Initialization
10 $L \xleftarrow{\$} \{0,1\}^n; \pi(0^n) \leftarrow L$
When A makes its $r-th\ query, M_r = M_r[1]...M_r[m_r]$ where $r \in [1,...,q]$
20 for $i \leftarrow 1$ to m_r do
21 $\{X_r[i] \leftarrow H(L, M_r[i], i)$
22 if $X_r[i] \in Domain(\pi)$ then $Y_r[i] \leftarrow \pi(X_r[i])$
23 else $\{Y_r[i] \xleftarrow{\$} \{0,1\}^n;\ \pi(X_r[i]) \leftarrow Y_r[i]\}\}$
24 $\Sigma_r \leftarrow Y_r[1] \oplus Y_r[i] \oplus ... Y_r[m_r]$
25 $Tag_r \xleftarrow{\$} \{0,1\}^n$
26 if $\Sigma_r \in Domain(\pi)$ then bad $\leftarrow$ true;
27 $\pi(\Sigma_r) \leftarrow Tag_r$
28 return Tag_r

Fig.3 Game 2

图 3 Game 2

由于我们删除的仅是 24 行和 29 行,由文献[7]可知,这是一个 lossy transition.因此有,

$$Adv^{prf}_{\text{DPMAC}[Perm(n),n]}(A) \le \Pr_2[\text{bad} = \text{true}] + \frac{(\sigma+q+1)^2}{2^{n+1}} \quad (2)$$

注意到,在 Game 2 中,由于 $Tag_1, Tag_2, \ldots, Tag_q$ 是完全随机选取的,它不受任何因素的影响,所以,该 Game 2 是 oblivious 的,也就满足文献[9]中第 4.2 节给出的 coin-fixing 定理.该定理指出,在保证所查询的消息不重复的情况下,即使攻击者固定查询序列为 $M_1, \ldots, M_q$(以正体表示)以及算法返回的标记序列,也就是如图 4 所示将 Game 2 改写为 Game 3,也有

$$\Pr_2[\text{bad} = \text{true}] \le \Pr_3[\text{bad} = \text{true}] \quad (3)$$

Game 3
simulation DPMAC
Initialization
10 $L \xleftarrow{\$} \{0,1\}^n; \pi(0^n) \leftarrow L$
When A makes its $r-th\ query, M_r = M_r[1]...M_r[m_r]$ where $r \in [1,...,q]$
20 for $i \leftarrow 1$ to m_r do
21 $\{X_r[i] \leftarrow H(L, M_r[i], i)$
22 if $X_r[i] \in Domain(\pi)$ then $Y_r[i] \leftarrow \pi(X_r[i])$
23 else $\{Y_r[i] \xleftarrow{\$} \{0,1\}^n;\ \pi(X_r[i]) \leftarrow Y_r[i]\}\}$
24 $\Sigma_r \leftarrow Y_r[1] \oplus Y_r[i] \oplus ... Y_r[m_r]$
25 if $\Sigma_r \in Domain(\pi)$ then bad $\leftarrow$ true;

Fig.4 Game 3

图 4 Game 3

为了更便于求概率,下一步我们修改 Game 3 成为 Game 4,使得 Game 3 中 bad 设置为 true 的概率小于等于 Game 4 中 bad 设置为 true 的概率.在 Game 3 的第 22 行,如果 $M_r[i]=M_s[i]$(其中,$s<r,i<m_s$),则必然有 $M_r[i]$已经在定义域中了,这种情况属于平凡的情况.此外还有一些非平凡的情况:$M_r[i]=0^n$ 或者 $M_r[i]=M_s[j]$(其中,$s<r$);或者 $M_r[i]=M_r[j]$(其中,$j<i$).在非平凡的情况下,如果 $M_r[i]$落在π的定义域中,我们就将 bad 设置为 true.这样就得到了 Game 4,如图 5 所示.

Game 4
simulation DPMAC
Initialization
10 $L \xleftarrow{\$} \{0,1\}^n; \pi(0^n) \leftarrow L$
When A makes its $r-th\ query, M_r = M_r[1]...M_r[m_r]$ where $r \in [1,...,q]$
20 for $i \leftarrow 1$ to m_r do
21 $\{X_r[i] \leftarrow H(L, M_r[i], i); Y_r[i] \xleftarrow{\$} \{0,1\}^n$
22 if $M_r[i] = M_s[i]$ for some $s < r$ and $i < m_s$ then $Y_r[i] \leftarrow \pi(X_r[i])$
23 else if $X_r[i] \in Domain(\pi)$ then bad $\leftarrow$ true;
24 $\pi(X_r[i]) \leftarrow Y_r[i]\}$
25 $\Sigma_r \leftarrow Y_r[1] \oplus Y_r[i] \oplus ... Y_r[m_r]$
26 if $\Sigma_r \in Domain(\pi)$ then bad $\leftarrow$ true

Fig.5 Game 4

图 5 Game 4

显然,在 Game 4 中,bad 被设置为 true 的概率大于等于 Game 3 中 bad 被设置为 true 的概率.假设在 Game 4 的执行过程中 bad 被设置为 true,也就是说,随机置换π的输入发生了非平凡的碰撞.这时考虑以下两种情况:

情况 1. 碰撞发生在某个消息的内部.这可能是由于 $M_s[i]=0^n$ 或者 $M_s[i]=M_s[j]$(其中,$j\in[1,...,i-1]$)或者 $M_s[i]=\Sigma_s$ 或者 $\Sigma_s=0^n$(其中,$s\in[1,...,q]$).

情况 2. 碰撞发生在两个消息之间.这可能是由于 $\Sigma_r=\Sigma_s, s\in[1,\ldots,r-1]$ 或者 $\Sigma_r=X_s[i], s\in[1,\ldots,r-1], i\le m_s$ 或者 $M_r[i]=\Sigma_s, s\in[1,\ldots,r-1], i\le m_r$,或者 $M_r[i]=M_s[j]$(其中,$s<r$).

这里,使用 $Mcoll_n(m)$度量攻击者查询一个具有 m 块的消息 M 时发生碰撞的概率;使用 $MMcoll_n(m,\overline{m})$ 度量攻击者查询两个长度分别为 $m,\overline{m}$ 的消息 $M, \overline{M}$ 时发生碰撞的概率.那么,由不等式(2)、不等式(3)及以上的推理可知,

$$\begin{aligned} Adv^{prf}_{\mathrm{DPMAC}[Perm(n),n]}(A) &\le \Pr_2[\mathrm{bad}=\mathrm{true}] + \frac{(\sigma+q+1)^2}{2^{n+1}} \\ &\le \Pr_4[\mathrm{bad}=\mathrm{true}] + \frac{(\sigma+q+1)^2}{2^{n+1}} \\ &\le \max_{\substack{m_1,\ldots,m_q \\ \sigma=\sum_{m_i\ge 1} m_i}} \left\{ \sum_{1\le s\le q} Mcoll_n(m_s) + \sum_{1\le s<r\le q} MMcoll_n(m_s, m_r) \right\} + \frac{(\sigma+q+1)^2}{2^{n+1}} \end{aligned} \tag{4}$$

为求出以上不等式,先证明以下引理.

引理 **1**. 假设 a,b,c,d 均是长度为 128-bits 的任意的无符号整数,且 $a\ne b\ne 0$.那么,最多存在 4 个整数 $L\in\{0,1\}^{128}$,使得等式 $H(L,c,a)=H(L,d,b)$成立.

证明:原式可写为 $a\times L \bmod(2^{128}+51)-b\times L \bmod(2^{128}+51)\equiv(d-c)\bmod 2^{128}$,定义 U 是属于区间$[-2^{128}-50, 2^{128}+50]$且模 2^{128} 同余 $d-c$ 的数的集合.注意到,集合 U 的元素个数$\#U$ 最多为 4.

$a\times L \bmod(2^{128}+51)-b\times L \bmod(2^{128}+51)\equiv(d-c)\bmod 2^{128}$ 等于 $a\times L \bmod(2^{128}+51)-b\times L \bmod(2^{128}+51)=u$,其中,$u\in U$.也就是 $(a-b)\times L\equiv u \bmod(2^{128}+51)$.由于 $2^{128}+51$ 是素数,且 $a-b$ 不为 0,因此,$(a-b)$在此域中存在逆,那么,$L\equiv u(a-b)^{-1} \bmod 2^{128}+51$.因为 u 最多有 4 种可能,所以,L 可能的取值最多有 4 个.引理得证. □

由以上证明可知,对于任意的长度为 128-bits 的无符号整数 d,因为 $a\ne 0$,所以,使 $H(L,c,a)=d$ 成立的 L 的可能

的取值最多有 4 个.这样,由引理 1 可知,如果 L 从 $\{0,1\}^{128}$ 中随机地选择,那么,使 $H(L,c,a)=H(L,d,b)$ 成立的概率是 $\frac{4}{2^{128}}=\frac{1}{2^{126}}$.

引理 **2**. $Mcoll_n(\cdot)$和 $MMcoll_n(\cdot)$分别代表一个消息内的碰撞的概率和两个消息间碰撞的概率,那么有不等式 $Mcoll_n(m)\le C_2^{m+2}\frac{1}{2^{126}}$ 和 $MMcoll_n(m,\overline{m})\le\frac{(m+1)(\overline{m}+1)}{2^{126}}$ 成立.

证明:定义 $D_1=\{0^n\}$,$D_2=\{X[1],\ldots,X[m]\}$,$D_3=\{\Sigma\}$, $D_4=\{\overline{X}[j]: j\in\{j:\overline{M}[j]\neq M[j]\}\}$, $D_5=\overline{\Sigma}$.

考虑碰撞发生在一个消息内的情况:

Case(D_1,D_2):Pr[$0^n=X[i]$]=Pr[$H(L,M[i],i)=0^n$].由于 L 是一个随机选取的整数,且 $i\neq 0$.由引理 1 可以很容易地证得这个概率要小于等于 2^{-126}.

Case(D_1,D_3):Pr[$0^n=\Sigma$].由于$\Sigma=Y[1]\oplus Y[2]\oplus\ldots\oplus Y[m]$是一个随机且独立于 L 的 n 比特串,因此,这个概率是 2^{-n},也就是 2^{-128}.显然小于 2^{-126}.

Case(D_2,D_2):对于 $i,j\in[1,\ldots,m]$,$i<j$,Pr[$X[i]=X[j]$]=Pr[$H(L,M[i],i)=H(L,M[j],j)$],由于 L 是一个随机选取的整数,由引理 1 可知,这个概率小于等于 2^{-126}.

Case(D_2,D_3):Pr[$X[i]=\Sigma$],其中,$i\le m$,由于$\Sigma=Y[1]\oplus Y[2]\oplus\ldots\oplus Y[m]$是一个随机且独立于 L 的 n 比特串,且 $i\neq 0$,由引理 1 可知,这个概率小于等于 2^{-126}.

综上可知:在集合 $D_1\cup D_2\cup D_3$ 中的 $m+2$ 个点中,任意两个碰撞的概率最多是 2^{-126},因此可知不等式 $Mcoll_n(m)\le C_2^{m+2}\frac{1}{2^{126}}$ 成立.

考虑碰撞发生在两个消息间的情况:

Case(D_2,D_4):因为 $i\in[1,\ldots,m]$,并且 $j\in\{j:\overline{M}[j]\neq M[j]\}$,考虑 $\Pr[X[i]=\overline{X}[j]]=\Pr[H(L,M[i],i)=H(L,\overline{M}[j],j)]$. 如果 $i\neq j$,由引理 1 可知,这个概率为 2^{-126};如果 $i=j$,因为 $M[i]\neq\overline{M}[j]$,必有 $H(L,M[i],i)\neq H(L,\overline{M}[j],j)$,因此这个概率为 0.

Case(D_2,D_5):Pr[$X[i]=\overline{\Sigma}$],由于 $\overline{\Sigma}$ 是一个随机且独立于 L 的 n 比特串,由引理 1 可知,这个概率小于等于 2^{-126}.

Case(D_3,D_4):Pr[$\Sigma=\overline{X}[j]$],由于Σ是一个随机且独立于 L 的 n 比特串,由引理 1 可知,这个概率小于等于 2^{-126}.

Case(D_3,D_5):Pr[$\Sigma=\overline{\Sigma}$],如果 $m>\overline{m}$,那么 Pr[$\Sigma=\overline{\Sigma}$]=2^{-n}.因为Σ含有一个随机变量 $Y[m]$而 $\overline{\Sigma}$ 没有,因此,二者相等的概率为 2^{-n};如果 $m<\overline{m}$,情况类似;当 $m=\overline{m}$ 时,因为 $M\neq\overline{M}$,必然存在 $i\le m$, $M[i]\neq\overline{M}[i]$,因此有 $Y[i]\neq\overline{Y}[i]$,概率 Pr[$\Sigma=\overline{\Sigma}$]$\le 2^{-128}$.综上可知,Pr[$\Sigma=\overline{\Sigma}$]$\le 2^{-126}$.

综上,集合 $D_2\cup D_3$ 中任意一点和集合 $D_4\cup D_5$ 中任意一点碰撞的概率最多是 2^{-126},因为有$|D_2\cup D_3|\cdot|D_4\cup D_5|\le(m+1)(\overline{m}+1)$,所以,不等式 $MMcoll_n(m,\overline{m})\le\frac{(m+1)(\overline{m}+1)}{2^{126}}$ 成立.

所以,引理 2 得证.以下证明定理 1.由以上引理可知,

$$
\begin{aligned}
Adv^{prf}_{\mathrm{DPMAC}[Perm(n),n]}(A) &\le \max_{\substack{m_1,\ldots,m_q\\ \sigma=\sum m_i\\ m_i\ge 1}}\left\{\sum_{1\le s\le q}Mcoll_n(m_s)+\sum_{1\le s<r\le q}MMcoll_n(m_s,m_r)\right\}+\frac{(\sigma+q+1)^2}{2^{n+1}}\\
&\le \max_{\substack{m_1,\ldots,m_q\\ \sigma=\sum m_i\\ m_i\ge 1}}\left\{\sum_{1\le s\le q}Mcoll_n(m_s)\right\}+\max_{\substack{m_1,\ldots,m_q\\ \sigma=\sum m_i\\ m_i\ge 1}}\left\{\sum_{1\le s<r\le q}MMcoll_n(m_s,m_r)\right\}+\frac{(\sigma+q+1)^2}{2^{n+1}}\\
&\le \max_{\substack{m_1,\ldots,m_q\\ \sigma=\sum m_i\\ m_i\ge 1}}\left\{\sum_{1\le s\le q}C_2^{m_s+2}\frac{1}{2^{126}}\right\}+\max_{\substack{m_1,\ldots,m_q\\ \sigma=\sum m_i\\ m_i\ge 1}}\left\{\sum_{1\le s<r\le q}\frac{(m_s+1)(m_r+1)}{2^{126}}\right\}+\frac{(\sigma+q+1)^2}{2^{n+1}}\\
&\le \frac{(\sigma+2q)^2/2}{2^{126}}+\frac{(\sigma+2q)^2}{2^{126}}+\frac{(\sigma+q+1)^2}{2^{n+1}}\\
&\le \frac{2(\sigma+2q)^2}{2^{126}}.
\end{aligned}
$$

定理 1 得证. □

与 XOR MAC 和 PMAC 相比,DPMAC 优势函数的界虽然稍大于它们,但这一界仍然是可忽略的.因此,我们说 DPMAC 是可证明安全的.以上是从可证明安全的角度对 DPMAC 进行的分析,下面从现有的针对消息认证码的攻击的角度来进行分析.针对消息认证码的攻击主要有两种类型,一种是伪造攻击;另一种是密钥恢复攻击.密钥恢复攻击要强于伪造攻击,因为一旦能够恢复密钥,就可以任意进行伪造.对于伪造攻击,上述安全性的证明已经给出了最好的这种攻击的界,因此,对于这种攻击,DPMAC 是安全的.而要发起密钥恢复攻击,首先要找到长度小于 128-bit 的两块不同消息的碰撞(也是一种伪造攻击),然后再使用穷尽密钥搜索的方法来恢复密钥.显然,这种攻击的复杂度要远大于伪造攻击,因此,对于密钥恢复攻击,DPMAC 也是安全的.

5 结束语

本文提出了一种新的可并行的消息认证码算法 DPMAC.该算法只需要一个密钥,而且是确定性的.本文不仅给出了具体的构造方法,而且使用 Game-Playing 技术证明了它的安全性.尽管它的安全性的界要大于 XOR MAC 和 PMAC,但它仍然是可忽略的.

由于 DPMAC 避免了消息的扩展,而且不需要维持随机值,因此在性能上要明显优于 XOR MAC 和 XOR MAC 的改进版 PCS(protected counter sums)[12].与 XECB MAC 相比,DPMAC 不需要维持随机值,而且在调用分组密码的次数上要小于等于 XECB MAC;而二者除了使用分组密码之外都使用了模运算,因此,DPMAC 的性能要略高于 XECB MAC.与 PMAC 相比,虽然 DPMAC 在性能上略低于它,但 PMAC 受专利保护.与 CBC MAC 相比,虽然在串行计算时,DPMAC 的性能要低一些,但是,DPMAC 不仅克服了 CBC MAC 只能处理定长消息的缺点,而且还是可并行的.

DPMAC 的性能主要受函数 $H(L,M,i)=i\times L \bmod (2^{128}+51)+M \bmod 2^{128}$ 的影响,而这种模运算在现代 CPU 上能够很快地加以计算,因此,DPMAC 能被快速地实现.此外,素数 $2^{128}+51$ 也可根据需要更换.由前面的证明可知,这并不影响其安全性.在实际实现中,如果允许查表,也可以让 $H(L,M,i)=i\cdot L\oplus M$(其中,“·”运算是 $GF(2^{128})$上模不可约多项式 $x^{128}+x^7+x^2+x+1$ 的乘法),这样,按照文献[13]中的方法使用查表运算也能很快地计算该函数.而 DPMAC 在 $H(L,M,i)=i\cdot L\oplus M$ 时的安全性证明与上述证明类似.

DPMAC 具有可并行性、确定性、支持在线、效率高、无专利保护以及可证明安全等诸多优点,其缺点是,当明文长度是分组长度的整数倍时,需要多调用一次分组密码.而这也是大多数消息认证码都具有的缺点.

References:

[1] Bellare M, Canetti R, Krawczyk H. Keying Hash functions for message authentication. In: Koblitz N, ed. Advances in Cryptology—CRYPTO'96. LNCS 1109, Berlin, Heidelberg: Springer-Verlag, 1996. 1–19.

[2] Bellare M, Kiliany J, Rogaway P. The security of the cipher block chaining message authentication code. Journal of Computer and System Sciences, 2000,61(3):362–399.

[3] Wu WL, Feng DG. The state-of-the-art of research on block cipher mode of operation. Chinese Journal of Computers, 2006, 29(1):21–36 (in Chinese with English abstract).

[4] Bellare M, Guerin R, Rogaway P. XOR MACS: New methods for message authentication using finite pseudorandom functions. In: Coppersmith D, ed. Advances in Cryptology—CRYPTO'95. LNCS 963, Berlin, Heidelberg: Springer-Verlag, 1995. 15–28.

[5] Black J, Rogaway P. A block-cipher mode of operation for parallelizable message authentication. In: Knudsen L, ed. Advances in Cryptology—EUROCRYPT 2002. LNCS 2332, Berlin, Heidelberg: Springer-Verlag, 2002. 384–401.

[6] Gligor VD, Donescu P. Fast encryption and authentication: XCBC encryption and XECB authentication modes. In: Matsui M, ed. FSE 2001. LNCS 2355, Berlin, Heidelberg: Springer-Verlag, 2002. 92–108.

[7] Goldwasser S, Bellare M. Lecture notes on cryptography. 2001. http://www.cse.ucsd.edu/users/mihir/crypto-lecnotes.html

[8] Goldreich O, Goldwasser S, Micali S. How to construct random functions. Journal of the ACM, 1986,33(4):792–807.

[9] Bellare M, Rogaway P. The game-playing technique. Cryptology ePrint Archive, Report, 2004/331. http://eprint.iacr.org/

[10] Shoup V. Sequences of games: A tool for taming complexity in security proofs. Cryptology ePrint Archive, Report, 2004/332. http://eprint.iacr.org/

[11] Hong DW, Kang JS, Preneel B. A concrete security analysis for 3GPP-MAC. In: Johansson T, ed. FSE 2003. LNCS 2887, Berlin, Heidelberg: Springer-Verlag, 2003. 154–169.

[12] Bernstein D. How to stretch random functions: The security of protected counter sums. Journal of Cryptography, 1999,12(3): 185–192.

[13] McGrew D, Viega J. The Galois/counter mode of operation (GCM). Submission to NIST Modes of Operation Process, 2004. http://csrc.nist.gov/CryptoToolkit/modes/proposedmodes/

附中文参考文献:

[3] 吴文玲,冯登国.分组密码工作模式的研究现状.计算机学报,2006,29(1):21–36.

王大印(1977－),男,博士,主要研究领域为密码学与信息安全.

吴文玲(1966－),女,博士,研究员,博士生导师,主要研究领域为密码学与信息安全,理论密码学及密码学中的数学问题.

林东岱(1964－),男,博士,研究员,博士生导师,主要研究领域为密码学与信息安全,网络分布式计算.