改进的多变量哈希函数

邹又姣[1,2] 马文平[1] 冉占军[2] 陈和风[1]
(西安电子科技大学ISN国家重点实验室 西安 710071)[1] (西安理工大学理学院数学系 西安 710048)[2]

摘 要 针对基于MI算法提出的一种多变量哈希函数进行研究,对该算法的安全性进行分析,找到其破解方法,并在此基础上对该算法进行改进。改进算法在保持了原有算法的所有优点的基础上对这种碰撞攻击免疫。还对该改进算法进行了原像攻击、第二原像攻击、差分攻击和代数攻击方面的安全性分析。同时建立数学模型,并通过实验测试了该改进算法的雪崩效应及其稳定性。实验结果表明,该算法满足严格雪崩效应原则,具有理想的、稳定的雪崩效应。

关键词 MQ问题,多变量,哈希函数,雪崩效应
中图法分类号 TN918.1 **文献标识码** A

Improved Multivariate Hash Function

ZOU You-jiao[1,2] MA Wen-ping[1] RAN Zhan-jun[2] CHEN He-feng[1]
(State Key Laboratory of Integrated Service Networks, Xidian University, Xi'an 710071, China)[1]
(Science of College, Xi'an University of Technology, Xi'an 710048, China)[2]

Abstract A multivariate hash function based on multivariate public key cryptographic algorithm MI was researched and its security was analyzed so that the method to broke the hash function was found. For this reason, the improved hash function was proposed. The improved hash function can keep all advantages of the original hash function. Further more it is immune on the collision. Its preimage attack, second preimage attack, differential attack and algebraic attack were also analyzed in this paper. The avalanche effect and its stability of the improved hash function were tested on the basis of a mathematical model. The experiment data shows the improved hash function meets the strict avalanche criterion and its avalanche effect is perfect and stable.

Keywords MQ problem, Multivariate polynomials, Hash function, Avalanche effect

1 引言

哈希函数是一类特殊的密码基本算法,在密码学中扮演着非常重要的角色,它被广泛应用于数字签名、完整性检验、身份认证和动态口令鉴别等现代信息安全技术中,起到保护和认证信息的作用。自2004年Wang等人陆续宣布成功破译MD5和SHA-1后[1-3],构造安全的哈希函数成为密码学界迫切要解决的问题之一。为此,NIST开始了为期2年的对目前哈希函数的安全性状况评估,并于2007年开始公开征集SHA-3[4]。目前哈希函数候选算法的遴选正处于紧张的最后阶段。在征集过程中,涌现出了许多构造哈希函数的方法,如Grøstl算法[5],其基于分组密码AES的S盒;Skein算法[6],其基于分组密码Threefish;BLAKE算法[7],其基于ChaCha流密码。这些算法的设计思路主要有两种:基于以分组密码为基本模块的设计和基于以ARX(Addition, Rotation and Xor)为基本模块的设计。基于分组密码设计的哈希函数,其安全性依赖于分组密码的安全性。文献[8]从生日攻击的角度分析了基于分组密码设计的哈希函数的安全性;而基于ARX设计的哈希函数,其安全性也受到了严重的质疑[1-3]。

基于数学困难性问题——MQ问题的多变量多项式是构造哈希函数的一种新方向、新思路。文献[9]中提出改进的M-D结构的二次多变量哈希函数,然而这种基于二次多变量多项式构造的哈希函数容易受到差分分析的攻击[10]。文献[11]中提出用随机选择三次多变量多项式系数的方法来构造哈希函数,该方法在计算过程中对计算机的内存要求较高。文献[12]基于MI思想,采用三次多变量多项式构造多变量哈希函数,内存要求比文献[11]低。本文针对文献[12]中提出的多变量哈希函数的安全性进行分析,找到攻击方法,得出碰撞概率至少为25%的结论。为此,在原算法思想的基础上进行改进,提出了改进算法。改进算法不仅保持了原算法的所有优点,而且对该攻击方法免疫。本文还对改进算法在原像攻击、第二原像攻击、差分攻击和代数攻击方面的安全性进行了分析,并在建立数学模型的基础上,对改进算法的雪崩效应及其稳定性进行了实验测试。测试结果表明,该改进算法满足严格雪崩效应原则,并具有良好的、稳定的雪崩效应。

本文第2节介绍基础理论知识及文中所涉及的主要符号;第3节介绍文献[12]中提出的基于MI思想构造的三次多变量哈希函数,并对其安全性进行分析;第4节针对原多变

到稿日期:2012-08-15 返修日期:2012-11-23 本文受国家自然科学基金(61003214,61173192,60773002),高等学校博士学科点专项科研基金(20100203110003),高等学校创新引智计划项目(B08038)资助。
邹又姣(1975—),女,博士生,讲师,主要研究方向为多变量公钥密码体制、哈希函数、云计算,E-mail:yjzou2009@126.com;马文平(1963—),男,教授,博士生导师,主要研究方向为密码算法分析与设计、信息安全。

量哈希函数的弊端进行改进，提出改进算法，并对改进算法的安全性进行分析；第5节建立数学模型，通过实验对改进算法的雪崩效应及其稳定性进行测试，给出实验数据及实验结论；最后总结全文。

2 预备知识

2.1 基础理论

哈希函数是一类特殊的单向函数，其输入值为任意长度的消息序列，并输出固定长度的哈希函数值，它必须满足下列3个基本的密码学性质：

1)单向性：对任意给定的 y，寻找一个 x'，使得 $H(x')=y$ 在计算上是困难的。

2)弱抗碰撞性：给定 x，寻找 x'，使得 $H(x')=H(x)$ 在计算上是困难的。

3)强抗碰撞性：寻找两个不同的 x_1, x_2，使得 $H(x_1)=H(x_2)$ 在计算上是困难的。

密码学中许多密码算法都是基于某一个数学NPC问题，如公钥密码体制RSA算法基于大整数分解的困难性哈希函数的安全性。本文讨论的三次多变量哈希函数基于MQ问题，即在有限域上解多变量方程组是困难的。

MQ Question：设

$$p_k(x_1,x_2,\cdots,x_n)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ijk}x_ix_j+\sum_{i=1}^{n}b_{ik}x_i+c_k \tag{1}$$

式中，$a_{ijk}, b_{ik}, c_k\in F_q, k=1,2,\cdots,m$，$F_q$ 是势为 q 的有限域。解方程组 $P=(p_1,p_2,\cdots,p_m)=0$ 称为MQ问题。

文献[13]证明了MQ问题是NPC困难性问题，同时在文献[14]中给出了结论。

定理1 对任意给定的 $\delta=(\delta_1,\delta_2,\cdots,\delta_n)$，在多项式时间 $O(mn^2)$ 内总能找到一对值 $x=(x_1,x_2,\cdots,x_n)$，$y=(y_1,y_2,\cdots,y_n)$ 满足 $P(x)=P(y)$ 且 $y-x=\delta$。

上述定理说明，由有限域上二次多变量多项式构造而成的哈希函数不满足强抗碰撞性。文献[15]也证明了这类哈希函数不能抵御差分攻击。但由三次及三次以上的多变量多项式构成的哈希函数却没有类似的结论。因此基于MQ问题构造由三次多变量多项式构成的安全的哈希函数是可能的。

2.2 主要符号介绍

C：压缩函数；

m_c：哈希函数链路值长度；

m：消息摘要长度；

M：消息序列，且 $M=M_1M_2\cdots M_L$；

n：压缩函数的分组长度；

IV：压缩函数的初始值；

h_i：第 i 个消息块的链路值；

$\#bits$：已处理的消息比特数；

$salt$：变量。

3 多变量哈希函数及安全性分析

文献[12]中多变量哈希函数的压缩函数为一组三次多变量多项式，它由二次多变量多项式与线性多变量多项式相乘而得到。二次多项式的生成运用了多变量公钥密码算法MI的思想，而线性多项式采用了随机生成。同时，构造的多变量哈希函数采用HAIFA结构。

多变量哈希函数采用反馈模式，输入为两部分，一部分由消息块与变量 $salt$ 异或而成，另一部分为当前链路值与 $\#bits$ 的异或。设其输入为 X, Y，则 $X=h_{i-1}\oplus\#bits$，$Y=M_i\oplus salt$。在压缩函数的运算中，对其两个输入变量分别进行运算，记为 $F_1(X)$ 和 $F_2(Y)$，然后将二者结合生成下一轮的链路值。

3.1 多变量哈希函数

3.1.1 $F_1(X)$ 的构成

设 k 为含有 $q=2^s(s\in Z^+)$ 个元素的有限域，$g(x)\in k[x]$ 为 $l=n/s$ 次不可约多项式，$K=k[x]/g(x)$ 为 k 的 l 次扩张。

定义同构映射 $\varphi: K\to k^l$

$$\varphi(a_0+a_1x+\cdots+a_{l-1}x^{l-1})=(a_0,a_1,\cdots,a_{l-1}) \tag{2}$$

及 K 上的函数 $\overline{F}_1(X)$：

$$\overline{F}_1(X)=X^{q^\theta+1} \tag{3}$$

式中，θ 为满足 $\theta<l$ 的任意正整数。

构造 k^l 上的函数 $F_1(X)$

$$\begin{aligned}F_1(x_1,x_2,\cdots,x_l)&=L_1\circ\varphi\circ\overline{F}_1\circ\varphi^{-1}(x_1,x_2,\cdots,x_l)\\&=(f_1,f_2,\cdots,f_l)\end{aligned} \tag{4}$$

式中，$X=(x_1,x_2,\cdots,x_l)$，L_1 为基域 $GF(2)$ 上的任意可逆的三角矩阵。因此 $f_1,f_2,\cdots,f_l\in k[x_1,x_2,\cdots,x_l]$ 为二次多变量多项式，且具有形式：

$$\sum_{i=1}^{l}\sum_{j=1}^{l}a_{ij}x_ix_j+\sum_{i=1}^{l}b_ix_i+d \tag{5}$$

式中，$a_{ij}, b_i, d\in k, i,j=1,2,\cdots,l$。

3.1.2 $F_2(Y)$ 的构成

定义

$$F_2(Y)=L_2(Y)+C=(t_1,t_2,\cdots,t_l) \tag{6}$$

式中，L_2 为 $GF(2)$ 上的任意可逆三角矩阵，C 为域 k 上任意的 l 维向量，$Y=(y_1,y_2,\cdots,y_l)$。因此 $t_1,t_2,\cdots,t_l\in k[y_1,y_2,\cdots,y_l]$ 为线性多项式，且具有形式：

$$\sum_{i=1}^{l}g_iy_i+e \tag{7}$$

式中，$e\in k, g_i\in GF(2), i=1,2,\cdots,l$。

3.1.3 压缩函数 $C(X,Y)$

定义压缩函数 $C(X,Y)$ 如下：

$$\begin{aligned}C(X,Y)=\langle X,Y\rangle&=(f_1t_1,f_2t_2,\cdots,f_lt_l)\\&=(c_1,c_2,\cdots,c_l)\end{aligned} \tag{8}$$

因此 $c_1,c_2,\cdots,c_l$ 为 k 上关于 $x_1,x_2,\cdots,x_l,y_1,y_2,\cdots,y_l$ 的三次多变量多项式，并且具有如下形式：

$$\begin{aligned}&\sum_{i=1}^{l}\sum_{j=1}^{l}\sum_{k=1}^{l}a_{ijk}x_ix_jy_k+\sum_{i=1}^{l}\sum_{j=1}^{l}b_{ij}x_ix_j+\sum_{i=1}^{l}\sum_{k=1}^{l}d_{ik}x_iy_k+\\&\sum_{i=1}^{l}c_ix_i+\sum_{k=1}^{l}g_ky_k+e\end{aligned} \tag{9}$$

式中，$a_{ijk}, b_{ij}, d_{ik}, c_i, g_k, e\in k$。

3.1.4 参数 $salt$ 的生成

该算法中参数 $salt$ 通过递归生成，第 i 轮的值由前两轮的链路值生成，即 $salt[j]:=h_i[j]\oplus h_{i-1}[m_c+1-j], j=1,2,\cdots,m_c$。

3.2 安全性分析

文献[12]中根据HAIFA结构的特点和MQ的困难性问题分析了该多变量哈希函数是满足抗原像攻击的。但从压缩函数的构成上，我们通过进一步的分析研究，发现这种结构对碰撞攻击是不免疫的。

结论1 针对文献[12]中提出的多变量哈希算法，能找到与原消息等长、相同哈希值、不同初始值的消息的概率约为

50%，而找到与原消息等长、相同哈希值、相同初始值的消息的概率约为25%。

证明：记第 i 轮中 F_1、F_2、C 的输出分别为 α_i、β_i 和 h_i，即 $\alpha_i=(f_{1i},f_{2i},\cdots,f_{li})$，$\beta_i=(t_{1i},t_{2i},\cdots,t_{li})$，$h_i=\langle\alpha_i,\beta_i\rangle=(c_{1i},c_{2i},\cdots,c_{li})=(f_{1i}t_{1i},f_{2i}t_{2i},\cdots,f_{li}t_{li})$。设消息 M 已知，寻找与 M 具有相同的哈希值的另一消息 M'，其对应的分块分别记为 M_i 和 M_i'。由有限域的知识知，给定 h_i，对任意 α_i，总存在唯一的一个解 β_i 满足 $h_i=\langle\alpha_i,\beta_i\rangle$。

设 j 是满足 $h_{j-1}\neq h_{j-1}'$ 的最大值。将任意取定的 h_{j-1}' 与 $\#bits$ 异或，并代入 F_1 得到 α_j'，从而在 $h_j'=h_j$ 的条件下，解得唯一值 β_j'，又 L_2 可逆，从而唯一解得 $M_j'\oplus salt_j'$，且 $M_j'\oplus salt_j'\neq M_j\oplus salt_j$。任取 $salt_j'$，从而得到 M_j' 且 $M_j'\neq M_j$。由 $salt_j'$ 的生成算法及 h_{j-1}' 已知，可求解出 h_{j-2}'。将 h_{j-2}' 与 $\#bits$ 异或，代入 F_1 得到 α_{j-1}'。在 h_{j-1}' 已知的条件下计算得到 β_{j-1}'。任取 $salt_{j-1}'$，进一步得到 M_{j-1}'。如此继续，找到 $M_{j-2}',\cdots,M_3'$ 及 $h_{j-4}',\cdots,h_1'$。

若初始变量 IV 未知，继续采用上述方法推导出 M_2' 及 h_0'。而在第一轮的推导中，取 $salt_1'=h_0'$，解得唯一值 M_1'。由 h_0' 和 m 已知及 $salt_0'=0$，最后解得初始设置 IV'，从而找到与消息 M 具有相同长度、相同哈希值但不具有相同的初始值 IV 的消息 M'。

若初始变量 IV 已知，则在第二轮的推导中，先采用上述方法计算出 β_2'，然后由 m 和 IV 计算出初始哈希值 h_0，又由 h_1' 已知，得到 $salt_2'$，从而求解出 M_2'。取 $salt_1'=salt_1=h_0$，解得唯一值 $M_1'=M_1$，从而找到与消息 M 具有相同长度、相同哈希值、相同初始值的消息 M'。

由于 $h_{j-1}\neq h_{j-1}'$，$h_j=h_j'$，因此第 $j+1$ 轮中参数 $salt_{j+1}\neq salt_{j+1}'$，$F_1$ 的输出 $\alpha_{j+1}=\alpha_{j+1}'$，从而在 $h_{j+1}=h_{j+1}'$ 的条件下，$\beta_{j+1}=\beta_{j+1}'$，进一步得到 M_{j+1}' 且 $M_{j+1}'\neq M_{j+1}$。在第 $j+2$ 轮，第 $j+3$ 轮，…，由于前两轮的链路值相同，故 $M_i'=M_i$，$i=j+2,j+3,\cdots$。由此得到了与 M 具有相同的哈希值的另一消息 $M'=M_1'M_2'\cdots M_{j+1}'M_{j+2}M_{j+3}\cdots$。

从推导过程可以看出，满足第 j 轮开始有相同的链路值和哈希值的等长、不同初始值的消息有 $2^{n(j-1)}-1$ 个，而满足第 j 轮开始有相同的链路值和哈希值的等长、相同初始值的消息有 $2^{n(j-2)}-1$ 个。对分组长度为 L 的消息来说，找到与消息 M 具有相同长度、相同哈希值、不同初始值的消息 M' 共有 $2^{n(L-1)}-1$ 个，发生碰撞的概率约为50%，而找到与消息 M 具有相同长度、相同哈希值、相同初始值的消息 M' 共有 $2^{n(L-2)}-1$ 个，发生碰撞的概率约为25%。

虽然从理论上，这类哈希函数对代数攻击方法是免疫的，但由证明过程可以看出，这种特定的攻击方法是有效的。为此，我们在遵从原有思想的基础上对该算法进行改进，使其对这种攻击方法免疫。

4 改进算法及其安全性、性能分析

4.1 改进算法

保持原算法的所有思想，仍然采用分别计算 $F_1(X)$、$F_2(Y)$ 的方式生成由三次多变量多项式组成的压缩函数，并且每一部分的计算方法保持不变。我们交换变量 X 和 Y，即取 $X=M_i\oplus salt$，$Y=h_{i-1}\oplus\#bits$。

4.2 改进算法的安全性分析

4.2.1 抗原像攻击

在新构造的哈希函数中压缩函数是由 $2l$ 个变量、l 个三次多变量多项式构成的，而解这些多项式构成的方程组是一个MQ问题[16]。对攻击者来说，即使他已知第 $i-1$ 轮、第 i 轮的链路值 h_{i-1}、h_i 及 β_i，可以求解出 $\alpha_i=F_1(M_i\oplus salt)$，但是由此计算出 $M_i\oplus salt$ 仍是一个MQ问题，因此很难还原得到消息 M。故算法能抵御第一原像攻击。

4.2.2 碰撞攻击

寻找碰撞攻击等价于在有限域上求解差分方程 $C(X+\Delta X,Y+\Delta Y)-C(X,Y)=0$。而在 $\Delta X\neq 0,\Delta Y\neq 0$ 的情形下，该方程组的次数至少为2。解该方程组是一个MQ问题，并且与第一原像攻击具有相同的复杂度。

针对上述攻击方法，该改进算法免疫。

由于解二次多变量方程组是一个MQ问题，因此上述攻击方法若能成功，必须首先设定二次多变量多项式这一部件的输入值，然后求解线性多项式这一构成部件的值。

同样地，我们设 j 是满足 $h_{j-1}\neq h_{j-1}'$ 的最大值。在改进方案中，任取 $M_j'\oplus salt_j'$，并代入 F_1 得到 α_j'，在 $h_j'=h_j$ 的条件下，解得唯一值 β_j'。又 L_2 可逆，唯一解得 $h_{j-1}'\oplus\#bits$，从而得到上一轮的链路值 h_{j-1}'。用如下表示方法形象地说明推导过程：

$$\forall(M_j'\oplus salt_j')\Rightarrow\alpha_j'\overset{h_j'=h_j}{\Rightarrow}\beta_j'\Rightarrow h_{j-1}'\oplus\#bits\Rightarrow h_{j-1}'$$

若仍然任取 $salt_j'$，虽然可以求得 M_j' 及由 $salt_j'$ 的生成算法得到 h_{j-2}'，但在第 $j-1$ 轮推导过程中，不能得到 F_1 这部分函数的输入值，因为尽管由 h_{j-2}' 可以计算得到 β_{j-1}'，再结合 h_{j-1}' 可计算出 α_{j-1}'，但是 α_{j-1}' 是由二次多变量多项式计算得到，求解 $M_{j-1}'\oplus salt_{j-1}'$ 仍是一个MQ问题。

若第 $j-1$ 轮推导过程中采用与上一轮相同的方法推导得到第 $j-2$ 轮的链路值 h_{j-2}'，即

$$\forall(M_{j-1}'\oplus salt_{j-1}')\Rightarrow\alpha_{j-1}'\overset{h_{j-1}'}{\Rightarrow}\beta_{j-1}'\Rightarrow h_{j-2}'\oplus\#bits\Rightarrow h_{j-2}'$$

则由 h_{j-1}' 和 h_{j-2}' 得到 $salt_j'$，将结果代入 $M_j'\oplus salt_j'$，求得 M_j'，…，以此类推到第二轮：

$$\forall(M_2'\oplus salt'_2)\Rightarrow\alpha'_2\overset{h_{j-1}'}{\Rightarrow}\beta_2'\Rightarrow h_1'\oplus\#bits\Rightarrow h_1'$$

得到 M_3'。

若初始变量 IV 未知，则采用同样方法推导第一轮：

$$\forall(M_1'\oplus salt_1')\Rightarrow\alpha_1'\overset{h_1'}{\Rightarrow}\beta_1'\Rightarrow h_0'\oplus\#bits\Rightarrow h_0'$$

结合上一轮的结果得到 M_2'。虽然由 $salt_1'=h_0'$，可以得到 M_1'，然而在有限域上解关于 IV 的二次多变量方程组是一个MQ问题，其安全性得以保证。

若初始变量 IV 已知，则取 $salt_1'=salt_1=h_0$，结合上一轮结果得到 M_2'。在第一轮的推导中可由 $salt_1'=h_0$ 计算得到 α_1'，然而 M_1' 是二次多变量多项式方程组的变量。算法的安全性基于MQ问题得以保证。

以上所述证明了这种改进方案对上述代数攻击方法免疫，故该算法对碰撞攻击免疫。

4.2.3 抗第二原像攻击

改进算法对第二原像攻击免疫。由下面的定理可以给与证明。

定理 1[10] 对任意的三次多变量多项式 $C:k^{2n}\to k^n$，有下面结论：

(1)在多项式时间内找到第二原像是不可能的；

(2)在多项式时间内找到另一个解是不可能的。

该算法的中心映射为三次多变量多项式，上述定理 1 充分证明了该算法抗第二原像攻击。

4.2.4 其他攻击

本算法由于采用 HAIFA 结构，添加了变量 # *bits* 和 *salt*，因此对定点攻击、多碰撞攻击免疫。

在中心映射中运用了在域 K 上的高次单变量单项式来进行计算，使得攻击者无法通过追踪比特异或来寻找碰撞，因此差分分析对该算法攻击无效。

本算法尽管基于代数方法进行构造，但是对代数攻击也是免疫的。通过对多变量多项式方程的研究，人们已经提出了许多的代数攻击方法，比如 Gröbner 基[17]、XL[18]、MXL 族[19-21]等等。这些攻击方法对由确定系数多变量多项式构成的算法是有效的，而本算法采用 MI 思想构造，其多变量多项式的系数是不定的，因此对代数攻击免疫。

5 雪崩效应分析及测试

改进算法仅交换了变量 X、Y，因此在内存需求上与原算法相同。本节主要通过实验来讨论雪崩效应及其稳定性。

为此，利用下述数学模型来对雪崩效应及其稳定性进行评价：

$$\begin{cases}\overline{B}=\dfrac{1}{N}\sum\limits_{i=1}^{N}B_i\\ p=\overline{B}/m\\ \Delta B=\sqrt{\dfrac{1}{N-1}\sum\limits_{i=1}^{N}(B_i-\overline{B})^2}\\ \Delta p=\sqrt{\dfrac{1}{N-1}\sum\limits_{i=1}^{N}(\dfrac{B_i}{m}-p)^2}\end{cases}\tag{10}$$

式中，B_i 表示第 i 次测试中哈希值改变的比特数，N 为测试次数，$\overline{B}$ 为平均改变的比特数，m 为消息哈希值长度，p 为平均改变概率，ΔB 和 Δp 分别表示改变比特数和改变概率的标准差。实验过程中，我们取测试次数为 200，消息长度任意，哈希值分别为 256，480，512 和 1024，域 k 分别为 $GF(2^8)$ 和 $GF(2^{16})$，改变原始消息中任意一比特，测试结果如表 1 所列。

表 1 改进算法的统计性能

k	m	B_{min}	B_{max}	$\overline{B}$	p(%)	ΔB	Δp(%)
2^8	256	112	148	127.65	49.86	7.28	2.84
	480	218	266	243.87	50.81	11.55	2.41
	512	228	296	255.81	49.96	12.64	2.47
	1024	481	562	514.73	50.27	14.94	1.46
2^16	256	109	142	126.86	49.55	6.71	2.62
	480	212	268	241.30	50.27	11.05	2.30
	512	232	286	255.37	49.88	9.72	1.90
	1024	460	556	510.49	49.85	16.31	1.59

通过实验数据，我们得到：改变消息中任一比特后哈希值有 50%左右的概率发生改变，满足严格雪崩效应原则[22]，达到了哈希函数雪崩效应的理想状态；同时 ΔB 和 Δp 的值都非常小，充分说明了该改进算法的雪崩效应具有良好的稳定性。

结束语 本文针对文献[12]中提出的多变量哈希函数的碰撞性概率进行分析，通过分析得出该算法是不安全的。为此，在原算法的基础上进行改进，提出了新的改进算法，该改进算法不仅保持了原有算法的所有优点，而且对这种碰撞攻击是安全的。同时，本文还建立了数学模型，通过实验测试了该改进算法的雪崩效应及其稳定性。实验结果表明，该算法具有理想的、稳定的雪崩效应。

参 考 文 献

[1] Wang Xiao-yun, Yu Hong-bo. How to break MD 5 and other hash functions[C]// Proceedings of EUROCRYPT 2005, LNCS 3494. Berlin: Springer-Verlag, 2005: 19-35

[2] Wang Xiao-yun, Yao A C, Yao F. Cryptanalysis of SHA-1 Hash Function[R]. Cryptographic Hash Workshop, Invited Report. 2005

[3] Wang Xiao-yun, Yu Hong-bo, Wang Wei, et al. Cryptanalysis on HMAC/NMAC-MD5 and MD5-MAC[C]// Proceedings of EUROCRYPT 2009. Berlin: Springer-Verlag, 2009: 121-133

[4] Federal Register. Government Printing Office[J]. 2007, 72 (212): 62212-62220

[5] Gauravaram P, Knudsen L R, Matusiewicz K, et al. groestl_FinalRnd. zip[OL]. http://ehash. iaik. tugraz. at /wiki/groestl

[6] Ferguson N, Lucks S, Schneier B, et al. Skein_FinalRnd. zip [OL]. http://ehash. iaik. tugraz. at/wiki/Skein

[7] Aumasson J-P, Henzen L, Meier W, et al. Blake_FinalRnd. zip [OL]. http://ehash. iaik. tugraz. at/wiki/BLAKE

[8] Yuan Z, Wang W, Jia K T, et al. New birthday attacks on some MACs based on block ciphers[C]// Proceedings of CRYPTO 2009. Berlin: Springer-Verlag, 2009: 209-230

[9] 王尚平，任娇霞，张亚玲，等. 改进 M-D 结构的二次多变量 hash 函数[J]. 哈尔滨工业大学学报，2011，32(4)：464-470

[10] Ding J T, Yang B Y. Multivariates polynomials for hashing [OL]. Cryptology ePrint Archive. http://eprint. iacr. org/2007/137. pdf, 2007

[11] 王后珍，张焕国，杨飚. 多变元 Hash 函数的构造与分析[J]. 电子学报，2011，39(1)：237-241

[12] Zou You-jiao, Ma Wen-ping, Ran Zhan-jun, et al. A New Multivariate Hash Function With HAIFA Construction[C]// 10th IEEE Int. Conf. on Trust, Security and Privacy in Computing and Communications, TrustCom 2011, 8th IEEE Int. Conf. on Embedded Software and Systems, ICESS 2011, 6th Int. Conf. on FCST 2011. 2011: 884-888

[13] Ding J T. Multivariate Public Key Cryptosystems[M]. Berlin: Springer-Verlag, 2006: 11-190

[14] 王后珍，张焕国，伍前红，等. 多变量 Hash 函数的构造理论和方法[J]. 中国科学：信息科学，2010，40(10)：1299-1311

[15] Biham E, Dunkelman O. A framework for iterative hash functions-HAIFA[C]// Proceedings of Second NIST Crypto-graphic Hash Workshop. 2006

[16] Kelsey J, Schneier B. Second Preimages on n-Bit Hash Functions for Much Less than 2n[C]// Cryptology, Proceedings of EUROCRYPT 2005, Lecture Notes in Computer Science 3494. Springer-Verlag, 2005: 474-490

[17] Adams W W, Loustaunau P. Introduction to Gröbner Bases[D]. Graduate Studies in Mathematics, American Mathematical Society, Providence, R. I., 1994

(下转第 75 页)

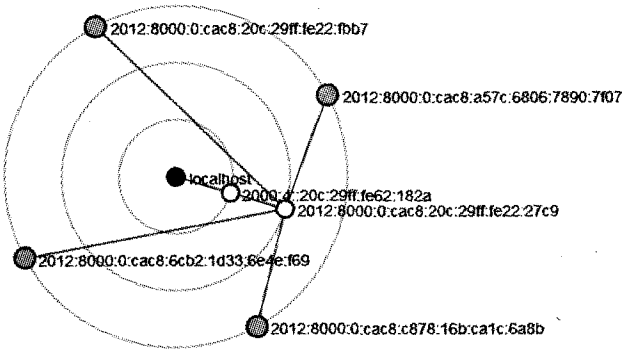

图 6　启用 ON-NTPM6 前 NMAP 拓扑探测结果

由结果可知，在未采取保护措施时，攻击者在获得站点内所有节点信息的情况下，通过探测得到的拓扑结构与网络的真实情况相同。

接着采用 ON-NTPM6 对站点网络进行保护。由于实验节点较少，设置欺骗拓扑深度 $h=2$，生成路标节点的概率为 0.25。再次使用 NMAP v5.91 对目标站点进行探测，探测所得拓扑如图 7 所示。

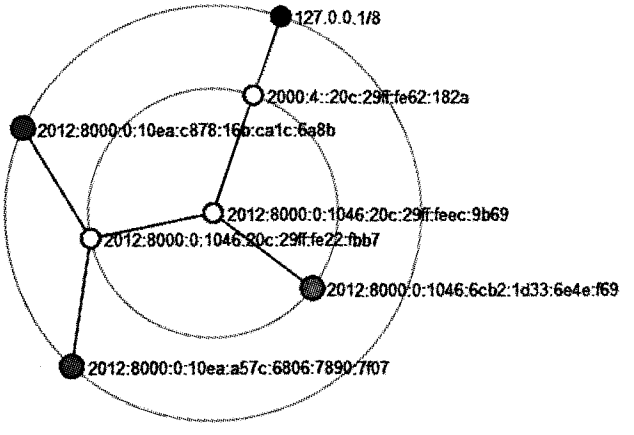

图 7　启用 ON-NTPM6 后 NMAP 拓扑探测结果

由结果可知，当采用 ON-NTPM6 对网络进行保护后，即使攻击者获得了当时全部存活主机的位置信息，但仍然只能得到一个虚假的拓扑结构，如主机节点（地址为 2012:8000:0:cac8:20c:29ff:fe22:fbb7）在启用 ON-NTPM6 后，攻击者通过拓扑探测，认为其是一个路由节点（地址为 2012:8000:0:1046:20c:29ff:fe22:fbb7）。也就是说，ON-NTPM6 能够有效地欺骗攻击者，从而保护目标站点的网络结构安全。

结束语　IPv6 的部署已全面展开。与 IPv4 相比，IPv6 并不是一次简单的协议升级，而是具有了许多改进的特性，如实现了网络层的身份认证、恢复了端到端通信等。这些新特性使得现有的以 NAT 机制为主的网络拓扑保护技术已不再适用。

借鉴“隐真”和“示假”的思想，本文提出了基于重叠网的 IPv6 网络结构保护模型（ON-NTPM6），首先提出“重叠隐蔽网”的设计，通过构建一个具有真实网络前缀的虚拟子网来实现对真实拓扑的隐藏，然后给出重叠隐蔽网拓扑动态生成算法，实现了重叠隐蔽网的拓扑结构的动态变化。理论分析与实验结果表明，所提出的模型可有效隐蔽网络真实拓扑结构，而且能够以虚假的复杂拓扑结构欺骗攻击者，消耗其攻击资源。

参 考 文 献

[1] Srisuresh P, Egevand K. Traditional IP network address translator (traditional NAT)[M]. IETF RFC 3022, Network Working Group, Jan. 2001

[2] Groat S, Dunlop M, Marchany R, et al. IPv6: nowhere to run, nowhere to hide[C]// Proceeding of the 44th International Conference on System Sciences. Hawaii, 2011: 1-10

[3] Narten T, Draves R, Krishnan S. Privacy extensions for stateless address autoconfiguration in IPv6[M]. IETF, RFC 4941, Network Working Group, Sep. 2007

[4] Wasserman M, Baker F. IPv6-to-IPv6 network prefix translation [M]. IETF, RFC 6296, Network Working Group, June 2011

[5] De Velde G V, Hain T, Droms R, et al. Local network protection for IPv6[M]. IETF, RFC 4864, Network Working Group, May 2007

[6] Beitollahi H, Geert Deconinck G. An Overlay Protection Layer against Denial-of-Service Attacks[C]// Proceeding of IEEE International Symposium on Parallel and Distributed Processing. 2008: 1-8

[7] Keromytis A D, Misra V, Rubenstein D. SOS: An Architecture for Mitigating DDoS Attacks[J]. IEEE Journal on selected areas in communications, 2004, 22(1): 176-188

[8] Keromytis A D, Misra V, Rubenstein D. SOS: Secure Overlay Services[C]// Proceeding of ACM SIGCOMM. 2002: 61-72

[9] Stavrou A, Keromytis A D, Nieh J, et al. MOVE: An End-to-End Solution to Network Denial of Service[C]// Proceeding of the ISOC Symposium on Network and Distributed System Security. 2005

[10] 杨柳，李振宇，张大方，等. 冗余最小化的 IPv6 拓扑发现方法[J]. 计算机研究与发展，2007，44(6)：939-946

[11] Lyon G F. Nmap Network Scanning. The Official Nmap Project Guide to Network Discovery and Security Scanning[M]. USA, 2009

（上接第 48 页）

[18] Courtois N, Klimov A, Patarin J, et al. Efficient Algorithms for Solving Overdefined Systems of Multivariate Polynomial Equations[C]// Proceedings of International Conference on the Theory and Application of Cryptographic Techniques (EUROCRYPT). Lecture Notes in Computer Science. 1807, Belgium: Springer, 2000: 392-407

[19] Ding J, Buchmann J, Mohamed M S, et al. MutantXL[C]// Proc. Symbolic Computation and Cryptography. Beijing, 2008: 16-22

[20] Mohamed M S, Mohamed W S, Ding J, et al. MXL2: Solving polynomial equations over GF(2) using an improved mutant strategy[C]// J. Buchmann, J. Ding, eds. Post-Quantum Cryptography-PQCrypto 2008 (Cincinnati, 2008). Lect. Notes Comput. Sci. 5299. Berlin: Springer, 2008: 203-215

[21] Mohamed M S E, Cabarcas D, Ding Jin-tai, et al. MXL3: An efficient algorithm for computing Gröbner bases of zero-dimensional ideals[C]// Proceedings of ICISC 2009. Lect. Notes Comput. Sci. 5984. Springer, 2010: 87-100

[22] Webster A, Tavares F, Stafford E. On the design of S-boxes[C]// Advances in Cryptology-Crypto'85. Lecture Notes in Computer Science 218. New York, NY: Springer-Verlag, 1985: 523-534