单向累计函数技术分析*)

万国根　周世杰　秦志光

(电子科技大学计算机科学与工程学院　成都 610064)

摘　要　单向累计函数是与密码学密切相关的一门新兴技术。本文综述了单向累积函数及其相关的基本技术。详细分析了单向累积函数的构造方法,包括 RSA 单项累积函数和无冲突单向累积函数。在讨论了单向累积函数的技术发展方向之后,对其应用前景和应用领域做了详细介绍。

关键词　单向累积函数,信息安全,密码学,分布式计算,认证

An Overview of the One-Way Accumulator Technology

WAN Guo-Gen　ZHOU Shi-Jie　QIN Zhi-Guang

(School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu 610054)

Abstract　The one-way accumulator (OWA) technology is one of important cryptograph related technology. This paper gives a summary discussion about this new technology. The methods used to build the one-way accumulator are also addressed, which include the RSA based OWA and the Free-collision OWA. Then, in order to apply this technology in information technology, some special issues about this technique and its applicable fields are also touched on.

Keywords　One-way accumulator, Information security, Cryptography, Distributed computing, Authentication

1　单向累积函数技术

在讨论累积函数之前,我们先定义一些基本概念。一般而言,关于单向 Hash 函数的定义和论述仅需考虑一个参数的情况,但是本文需要考虑两个输入参数的情况,因此给出其定义:

定义 1　单向 Hash 函数类 $H=\{h;X\times Y\rightarrow Z\}$ 是一组满足以下性质的函数集合:

$P\{k\in Z_N,(x,y)\in X\times Y,y'\in Y::$

$$\exists x'\in X;h(x,y)=h(x',y')\}<\frac{1}{A(k)}$$

其中 $A(\cdot)$ 表示多项式时间算法。上述定义表明,给定 $x,y,z=h(x,y)$ 的计算可以在多项式时间内完成;而如果仅知道 x,y,y',寻找 x' 并使其满足 $h(x,y)=h(x',y')$ 的计算,在多项式计算时间内完成的概率极小。但是,该定义没有否认满足条件的 x' 的存在性,只是表明寻找这样的值在整个集合中很难,即计算困难。因此,上述单向函数的定义可以理解为对不同的值对 (x,y) 进行 $h(x,y)$ 运算得到相同值的冲突(Collisions)很小。

尽管上述定义没有指明函数 H 输出值和输入值的长度关系,但在本文中认为 $|X|\approx|Y|\approx|Z|\approx k$($k$ 是一个与安全需要相关的参数)。

定义 2　函数 $f:X\times Y\rightarrow X$ 具有半交换率(Quasi-commutative)是指对 $\forall x\in X,\forall y_1,y_2\in Y$,以下等式成立:

$f(f(x,y_1),y_2)=f(f(x,y_2),y_1)$

一个单向 hash 函数满足半交换率,则首先根据单向性,正向计算容易,而反向计算困难。其次,满足半交换率意味着在给定初始值(Seeds)的条件下,多次 Hash 运算所得到的结果不会因计算顺序的不同而发生改变。

具有半交换率性质的单向 Hash 函数可以用来验证某个值 y_i 是否在指定的集合 $Y=\{y_i\}$ 中。具体做法是,Y 的累积计算结果 z 可以用单向累积函数 $h\in H$ 采用以下公式来计算:

$z=h(h(\cdots h(h(h(x_0,y_1),y_2),y_3),\cdots,y_{l-1}),y_l)$

除去 y_i 以外的其他值 $y_i=\{y|y\in Y,y\neq y_i\}$ 的累积值(称之为部分累积值)也可以用单向累积函数来计算:

$z_i=h(h(\cdots(h(h(\cdots h(h(h(x_0,y_1),y_2),y_3),\cdots,y_{i-1}),$
$y_{i+1}),\cdots,y_{l-1}),y_l)$

需要验证 $y_i\in Y$ 时,使用以下公式计算 z':

$z'=h(z_i,y_i)$

如果　$z'=z$

则　$y_i\in Y$

上述结论成立是因为如果攻击者不知道 y_i,根据对单向函数的定义,则它将面临构造 y' 使得 $z=h(z_i,y')$ 成立的计算困难。因此,(y_i,z_i) 可以视为 $y_i\in Y$ 的见证(Witness)。在以下讨论中,Z_N 表示所有的正整数集合,Z_n 表示长度在 n 以内的正整数集合(本文中,在某些条件下,Z_n 和 Z_n 也可能是指足够大的整数集合,根据上下文应可分辨其含义,因此本文不再做额外说明)。

定义 3　(y_i,z_i) 是 $y_i\in Y$ 的见证,是指它满足以下条件:

$p\{k\in Z_N,z_i,y_i::$

$$\exists y'\in Y,y'\neq y_i;h(z_i,y_i)=h(z_i,y')\}<\frac{1}{A(k)}$$

但是,上述分析存在一个明显的问题:我们总是假定攻击者只能在给定集合 Y 中随机选取预测值 y'。而实际上攻击者完全可能在值域集合 Y 之外,容易地找到一个 y' 满足 $z=h(z_i,y')$,从而破坏上述关于见证的定义。如果将攻击者可选择的预测值的范围扩展到指定集合 Y 之外,则可以得到强单

*)本课题得到国家高新技术研究发展引导计划(863 引导计划)网络化系统容灾技术(项目编号:2002AA001042),国家高新技术发展计划(863 计划)战略预警与监管体系结构研究(项目编号:2002AA142040)资助。**万国根**　博士研究生,主要研究方向:中文信息处理技术、网络安全理论与技术。**周世杰**　博士,主要研究方向:中间件、网络安全理论与技术。**秦志光**　博士,博士生导师,主要研究方向:中间件、网络安全理论与技术。

向累积函数的定义。

定义 4 强单向 Hash 函数类 $H=\{h:X\times Y\to Z\}$ 是一组满足以下性质的函数集合：

$P\{k\in Z_N,(x,y)\in X\times Y::$

$$\exists x',y':h(x,y)=h(x',y')'<\frac{1}{A(k)}$$

上述定义与普通单向函数的区别在于，前者允许攻击者在已知集合和该集合之外选择预测值，而后者则要求攻击者只能在已知的集合内选择。同样，我们可以得到关于无冲突(Free-Collision)强见证的定义：

定义 5 (y_i,z_i) 是 $y_i\in Y$ 的强见证，是指满足以下条件：

$p\{k\in Z_N,z_i,y_i::$

$$\exists y',y'\neq y_i:h(z_i,y_i)=h(z_i,y')\}<\frac{1}{A(k)}$$

具有"见证"一个指定的值是否属于一个指定输入集合功能的特殊单向函数，被称之为单向累积函数(OWA:One-Way Accumulator)[1,2]，以示与其他单向函数的区别。为了今后使用方便，我们给出单向累积函数的定义：

定义 6 单向累积函数(OWA)是指具有以下性质的函数：

(1)一个与安全参数 s 有关的 N 个关键字集合 $K=\{k_i\mid i=1,2,\cdots,N\}$ 及一个概率多项式时间算法满足：

$DOWA\_GenKey_s(K):K\to k$

且：

$$P\{DOWA\_GenKey(K):K\to k\}=\frac{1}{N}$$

该算法根据安全参数，随机生成一个安全关键字。该过程可以用一般的随机函数来完成，在以下讨论中，这样的随机函数用 $Random(X)$ 表示。

(2)对 $\forall k\in K$，存在一个多项式时间算法生成一组关于 k 的合适输入值 Y：

$DOWA\_GenRep_k(Z_N):Z_N\to Y,Y=\{y_i\mid i=1,2,\cdots,N\}$

该算法表明，可以在整数集合中找到一组适合于安全关键字 k 要求的输入。

(3)一个满足半交换率的单向 hash 函数 $h\in H,H:X\times Y\to X$ 是满足半交换率的单向 hash 函数类，根据该 hash 函数 h 可以构造以下多项式时间算法：

(a)累积值算法：$z(Y)=DOWA\_AccTot(Y)$

(b)部分累积值算法：$z_i(Y-\{y_i\})=DOWA\_AccPar(Y-\{y_i\})$

(c)验证算法：$z'(z_i,y_i)=DOWA\_AccAut(z_i,y_i)$

条件(1)和(2)说明对于这个特殊的单向函数的输入集合和输出集合容易构造，而条件(3)则说明该单向函数必须满足半交换率，且容易构造出用于"见证"目的的算法。简而言之，单向累积函数是一个满足半交换率的单向函数，该函数对一个输入集合进行运算之后，仍保持单向性；其次，根据该函数，可以构造一组算法，用于见证一个指定的值是否属于预定的输入集合。

2 RSA 单向累积函数

文[1]不仅提出了单向累计函数的概念，而且还构造了一个基于 RSA 假设的单向累计函数。在讨论 RSA 单向累计函数的构造之前，我们将与本文相关的密码学中常见的定义、概念和定理描述如下。

定义 7 p 为强素数是指 $p=2p'+1$，且 p' 为奇素数。

定义 8 n 为严格整数是指 $n=pq$，其中 p,q 均为强素数且 $p\neq q$。

在有关 RSA 密码理论的书籍和文献中，将 n 称之为 RSA 模数，而且对 n 的计算也有不同的定义。为了将我们关于严格整数的定义与 RSA 模数的定义区别开，我们用 Z_s 来表示所有严格整数的集合。

定义 9 二元强泛 hash 函数类 $U=\{u:A\to B\}$ 是满足如下性质的一组无限函数集合：

$$P_{h\in H}\{\forall a_1,a_2\in A,a_1\neq a_2:u(a_1)=u(a_2)\}\leqslant\frac{1}{|B|}$$

定义 10 RSA 难题是指已知 $\forall y,z,n\in Z_n^1$，$\exists x\in Z_n:z=x^y \bmod n$

定义 11 RSA 难题猜想是指对所有多项式时间算法 A，RSA 难题在计算上不可行，即：

$$P\{y,z,n\in Z_n::\exists x:z=x^y \bmod n\}\leqslant\frac{1}{A(n)}$$

根据 RSA 难题假设，函数 $z=x^y \bmod n$ 满足单向性。其次，函数 $z=x^y$ 满足半交换率。即 $\forall y_1,y_2:z(z(x,y_1),y_2)=(x^{y_1})^{y_2}=x^{y_1y_2}=(x^{y_2})^{y_1}=z(z(x,y_2),y_1)$ 成立。由此我们可以得到：

定义 12 如果 n 为严格整数，对于输入集合 $Y=\{y\mid y\in Z_N\},|Y|<n$，函数 $z=x^Y \bmod n$ 在 RSA 难题猜想下被定义为 RSA 单向累积函数(RSA-OWA:RSA One-Way Accumulator)[1]。

其中 $z=x^y \bmod n\leftrightarrow z=x^{\prod_{y\in Y}r^y}$。

RSA 累积函数的特点是简单、易于实现，且它以一般的密码学理论为基础，安全性容易得到证明。此外，文[1]还证明，如果求解元根在计算上困难，则 RSA 单向累计函数可以抵御攻击者主动伪造输入值对算法的攻击。但是，由于 RSA 累积函数的单向性是以 RSA 难题猜想为基础的，因此在指定集合内可能存在较小的冲突，属于弱单向函数的定义。相应的 RSA 累积函数也属于弱单向累积函数，即在指定输入集合中冲突的几率很小。RSA 单向累积函数的构成如下：

(1) $DOWA\_GenKey_s(K):=Random(Z_n)$，其中 $Rand(Z_N)$ 是一个大整数随机选择函数。

(2) $DOWA\_GenRep(Z_N):=Random(Z_N)$，其中 $Rand(Z_N)$ 是大整数随机选择函数。

(3) $h(x,y)=x^y$，n 为严格整数：

(a)累积值算法：$DOWA\_AccTot(Y):=x^{\prod_{y\in Y}r^y} \bmod n$

(b)部分累积值算法：$DOWA\_AccPar(Y-\{y_i\}):=x^{\prod_{y\in Y-\{y_i\}}y} \bmod n$

(c)验证值算法：$DOWA\_AccAut(z_i,y'):=z_i^{y'} \bmod n$

3 无冲突 RSA 单向累积函数

文[2]在 RSA 单向累计函数的基础之上，提出了所谓的无冲突 RSA 单向累计函数的概念。RSA 单向累计函数假定攻击者只能在制定的范围内选择输入进行攻击。而在实际应用环境中，面临攻击者可以伪造任意的输入值。在此条件下，RSA 出现冲突的概率增大，从而破坏单向累计函数的性质。Nibo Baric 等人以 RSA 强难题猜想为基础，提出了所谓的无冲突单向累计函数的概念。

定义 13 RSA 强难题是指已知 $\forall z,n\in Z_s$，$\exists x\in Z_p,y:z=x^y \bmod n$，其中 Z_p 是素数集合。

定义 14 RSA 强难题猜想是指对所有的多项式时间算

[1] 这里的 Z_n 表示大整数集合

法 A，强 RSA 难题在计算上不可行，即：

$$P\{z\in Z_n, n\in Z_s :: \exists x\in Z_n, y\in Z_p : z=x^y \bmod n\}\leqslant\frac{1}{A(n)}$$

与一般的 RSA 相比，强 RSA 难题猜想允许自由选择(x, y)的组合，即攻击者不仅可以选择指数函数的底数，也可以选择指数。此外，强 RSA 难题猜想也要求指数为素数，而一般的 RSA 难题假设对指数没有特殊要求。对于强 RSA 难题猜想没有严格的证明表明其计算上可行；同样，也没有严格的理论证明能表明其在计算上可行。

定义 15 在强 RSA 难题猜想的条件下，对 $n\in Z_s$，$\gcd(n,x)=1$，输入集合 $Y=\{y|y\in Z_p\}$，$|Y|<n$，函数 $z=x^Y \bmod n$ 被称之为无冲突 RSA 单向累积函数（FRSA-OWA：Free-collision RSA One-Way Accumulator）[2]。

与 RSA 单向累积函数不同，强 RSA 单向累积函数要求所有的输入为素数，从而保证攻击者在自由选择 x, y 的条件下，产生冲突的概率很小。关于强 RSA 单向累计函数的无冲突性由以下定理说明：

定理 1 在强 RSA 难题猜想成立的条件下，强 RSA 单向累计函数是无冲突的。

该定理的证明参见文[2]。无冲突 RSA 单向累积函数的组成如下：

(1) $DOWA\_GenKey_s(K)=Random(Z_N)$，其中 $Rand(Z_N)$是一个大整数随机选择函数。

(2) $DOWA\_GenRep(Z_N):=Random(Z_p)$，其中 $Rand(Z_p)$是大素数随机选择函数。

(3) $h(x,y)=x^y$，n 为严格整数；

(a)累积值算法：$DOWA\_AccTot(Y):=x^{\prod_{y\in Y}r^y} \bmod n$

(b)部分累积值算法：$DOWA\_AccPar(Y-\{y_i\}):=x^{\prod_{y\in Y-\{y_i\}}y} \bmod n$

(c)验证值算法：$DOWA\_AccAut(z_i,y'):=z_i^{y'} \bmod n$

此外，依据数论相关知识，在强 RSA 单向累积函数中，可构造出增加值和删除值算法：

(d)增加值算法：$DOWA\_AccAdd(z,y):=z^y \bmod n$

(e)删除值算法：$DOWA\_AccDel(z,y):=z^{y^{-1} \bmod \varphi(n)} \bmod n$

4 动态单向累计函数

文[3,5]分别讨论了动态单向累计函数。从概念上来看，动态单向累计函数是强 RSA 单向累计函数的拓展。强 RSA 单向累计函数是通过将输入限定为素数，而在实际应用环境中，输入并不一定满足该要求。如果输入不是素数，则上述构造函数的删除操作效率很低。为此，动态单向累计函数要求构造出来的算法必须具有额外的一个特性：存在高效删除算法，使得删除原有输入值集合中的任意一个值可以进行高效的验证操作。

定义 16 动态单向累计函数是指满足以下性质的单向累计函数：如果 $v=f(u,x')$，$x,x'\in X$，并且 $f(w,x)=v$，那么存在高效算法 D 和 G 使得：(1) $D(t_f,v,x')=v'\rightarrow v'=f(u,X-\{x'\})$；(2) $W(f,v,v',x,x')=w'\rightarrow f(w',x)=v'$。

其中 t_f 是一个陷门值，如果知道该陷门值，则可以利用算法 D 从输入集合中删除任何一个值。删除一个值之后，利用算法 W 可以对其他值进行正常认证。根据对动态单向累计函数的定义，文[3,5]讨论了其构造，并证明在强 RSA 难题猜想成立的条件下，该过程满足动态单向累计函数的定义。

5 单向累积函数技术发展及其应用分析

单向累计函数存在的条件是单向函数存在。单向函数的存在性以及寻找新的单向函数，一直是密码学领域的研究热点。RSA 单向累积函数由于与 RSA 密码体制有着密切的关系，固然具有其优越性，这也是现在所有的单向累计函数均建立在 RSA 假设之下。但是，以 RSA 单向累计函数为代表的现有单向累计函数只有在输入为大整数（或者大素数）的条件下，其安全性才能得到保证。这就大大限制了单向累计函数在信息安全领域（尤其是无线安全领域和对等计算[9]领域）的推广应用。在上述环境中，某些节点的存储能力和计算能力均受限，从而无法完成需要的大量计算工作。文[7]分析了单向累计函数在无线感知网络中的应用，分析结果表明，在输入为 1024 位的条件下，节点的存储开销总计 4K 字节。这在一般的环境中完全可以被接受，但是在无线环境中，这也是不小的一个开销。此外，RSA 对计算的要求，也使得节点计算开销很大。

因此，如何构造出输入值很小，但安全强度又足够高的单向累计函数将是该领域的重要研究方向之一。文[7]提出了构造基于椭圆曲线密码算法（ECC）的单向累计函数，这样在同等安全强度条件下，可减少对输入长度的要求，从而节约存储空间。此外，由于 ECC 计算速度在同等条件下也较 RSA 快，因而也可以节约一定的计算开销。

应用是检验技术是否科学的重要依据。如何将单向累计函数技术应用到信息安全领域，也是该领域的重要研究方向之一。目前，单向累计函数在认证和数字签名等领域[3,6,7]的研究工作也处于起步阶段，尚有大量工作需要开展。

总结 单向累计函数是与密码算法紧密相关的新兴技术，它在信息安全领域有着广泛的需求。本文从理论的角度分析了单向累计函数技术。但是单向累计函数与密码学有着千丝万缕的联系，因而其发展也与密码学进展有很大的关系。必须结合密码学知识，才能构造出满足需要的单向累计函数来。

参 考 文 献

1 Benaloh J, Mare M D. One-Way Accumulators: A decentralized alternative to digital signatures. In: advances in ctryptology Eutocrypt'93, Proc. of the workshop on the theory and applications of cryptographic techniques, LNCS 765, Springer-Verlag, Berlin, 1994

2 Maric N, Piftzmann B. Collision-Free Accumulators and Fail-Stop Signature Schemes Without Trees. Eurocrypt'97, 1997

3 Camenisch J, Lysyanskaya A. Dynamic accumulators and application to efficient revocation of anonymous credentials. Crypto'02, 2002

4 Sander T, Ta-Shma A, Yung M. Blind, auditable membership proofs. In: proc. Financial cryptograph(FC 2000), LNCS 1962, Springer-Verlag, 2001

5 Goodrich M T, Tamaassia R, Hasic J. An efficient dynamic and distributed cryptographic accumulator. Lecture Notes In Computer Science. In: Proc. of the 5th Intl. Conf. on Information Security, Springer-Verlag, 2002

6 Gokhale S, Dasgupta P. Distributed authentication for peer-to-peer networks. In: Symposium on Applications and the Internet Workshops (SAINT'03 Workshops), Orlando, Florida, 2003

7 Zachary J. A decentralized protocol for secure Node-to-Node authentication in distributed sensor networks. http://www.cse.sc.edu/~jmz/milcom-draft.pdf, 2003

8 Striner, Micharl, Tsudik G, et al. CLIQUES: a new approach to group key agreement. IEEE Transactions on parallel and distributed systems, 2000

9 Milojicic D S, et al. Peer-to-Peer Computing. HP Laboratories palo alto, HPL-2002-57, March 2002