不可靠信道上抗主动攻击的组播认证

何永忠[1,2], 韩 臻[1], 李晓勇[1]
(1. 北京交通大学计算机学院信息安全实验室, 北京 100044; 2. 中国科学院信息安全国家重点实验室, 北京 100080)

摘 要: 组播是视频会议、协同工作等各种群组应用的基本通讯模式, 组播安全性的研究具有重要意义. 组播通常构建在不可靠的通讯协议上, 因此存在数据包的丢包现象. 大多数的组播认证方案不能用于这种环境, 其他一些方案的主要目标是针对网络通讯故障引起的随机包丢失情况, 而不能抵抗主动攻击. 本文提出了抗部分碰撞哈希函数簇的思想, 然后利用哈希图和纠错码技术提构造一种在不可靠信道上新的组播认证方案. 该方案不仅具有很高的通讯性能和计算性能, 并且在存在部分数据包丢失的情况下也可以抵抗主动攻击. 本文提出了一种针对该方案特性的不可靠信道组播认证的形式安全模型, 并在此安全模型下基于规约技术证明了该方案的安全性.

关键词: 组播; 不可否认性; 不可靠信道; 主动攻击
中图分类号: TP393. 08 文献标识码: A 文章编号: 0372-2112 (2008) 07-1249-07

Multicast Authentication over Lossy Channels Against Active Attack

HE Yong-Zhong[1,2], HAN Zhen[1], Li Xiong-Yong[1]
(1. *Laboratory of Information Security, School of Computer and Information Technology, Beijing Jiaotong University, Beijing* 100044, *China*;
2. *State Key Laboratory of Information Security, Chinese Academy of Sciences, Beijing* 100080, *China*)

Abstract: Multicast is the fundamental communication mechanism for all sorts of group oriented application sush as video conference or cooperative work. It is important to study and improve the security of multicast. As multicast is layered on top of unreliable communication protocols such as UDP in TCP/IP protocol suites, data packets are lost possibly. Most multicast authentication schemes do not tolerate packets loss, some others may only work in random packets loss while vulnerable to active attack. In this paper, a new concept of partial collision resistant hash function is introduced, base on which, together with hash graph and error correcting code, an efficient multicast authentication scheme over lossy channel is presented. The scheme is not only very efficient in computation and communication complexity, but also secure against active attack. After a formal security modeling of multicast authentication schemes over lossy channel, the security of the proposed scheme is proved under this definition.

Key words: multicast; nonrepudiation; lossy channels; active attack

1 引言

组播是计算机网络中一种常用的通讯方式. 在组播通讯中, 一个用户可以同时向组中其他用户发送消息. 组播是支持付费电视、视频会议和协同工作等各种群组应用的基本通讯模式. 由于组播层的安全机制可以为所有上层的群组应用提供安全服务, 从而大大简化群组应用的安全功能设计和实现, 因此组播安全性的研究具有重要意义. 目前, 组播安全研究的方向主要有组播密钥管理、组播源认证、接收者访问控制和组播数字指纹等等[1], 其中, 组播认证机制可以支持接收者对组播信息的完整性进行验证. 组播认证方案的一般应该满足下列安全需求: 可认证性, 数据接收者可以验证数据发起者的身份; 完整性, 数据接受者可以验证接收的数据没有被修改过; 不可否认性, 数据发起者不能否认相应的数据是他发出的.

组播认证与一般的消息认证具有不同的特点. 首先, 组播通讯是一对多的关系, 即一个发送者对应多个接收者, 而消息认证的一般模型是一个发送者和一个接收者; 其次, 组播通讯往往具有实时性特征, 发送者需要及时生成验证信息, 接收者需要及时验证数据的完整性; 最后, 在互连网中, 实时的组播通讯通常采用面向无连接的UDP协议实现, 数据包丢失不会重传. 由于这些特点, 导致组播认证比普通的消息认证问题更为复杂和困难: (1) 不能直接使用消息认证码(MAC). 由于组播通讯中存在多个接收者, 如果直接使用消息认证码对每一

收稿日期: 2007-06-25; 修回日期: 2008-01-11
基金项目: 国家863高技术研究发展计划(No. 2007AA01Z410); 国家973重点基础研究发展规划(No. 2007CB307100, 2007CB307106); 信息安全国家重点实验室开放课题

个组播消息进行认证就不能满足不可否认性的要求,因为对每个用户认证密钥都是一样的,所以除了发送者其他用户都可以伪造认证码. (2) 不能直接使用消息数字签名方案. 由于组播通讯的数据量很大,直接对每一个数据包签名和验证会给发送者和验证者带来极大的计算负担,因此是不可行的.

基于组播认证的需求和特点,研究者提出了各种组播认证方案. 针对单一 MAC 的内部人伪造攻击,文献[1]采用多重 MAC 技术进行组播认证,对每个数据包用多个对称密钥生成多个 MAC,从而防止一定数量的内部人合谋攻击. 此方案的主要问题是可扩展性差,且不具有不可否认性. Gennaro 和 Rohatgi 等人[2]提出了流签名问题,并给出了一个基于签名和哈希函数链的流签名方案,其思想成为此后很多组播认证的基础,但是如果出现数据包丢失,接收者就不能对丢失的数据包后面的数据包进行验证. 各种针对网络传输存在丢失数据包的情况,文献[3]提出的基于 Merkle 哈希树的组播认证方案,其主要缺点是通讯代价较大. 基于哈希链的思想,出现了多种基于哈希图的组播认证方案[4][5],主要针对随机的数据包丢失或者突发性数据包丢失;李等[6]提出基于随机函数簇和认证矩阵的组播认证方案,具有较高的性能,可以很好的应对随机的包丢失情况,但是不能抵抗主动的伪造攻击[7]. 文献[8]提出了基于纠错码的方案,以及文献[9]提出的基于擦除容忍的消息认证方案可以抵抗主动攻击,但是通讯效率较低.

所谓主动攻击,是指攻击者对组播认证的伪造攻击. 与对普通的认证方案的伪造攻击不同,攻击者除了具有自适应选择密文攻击能力外,还具有选择性丢包的能力. 本文首先提出了抗部分碰撞哈希函数簇的思想,然后利用哈希图和纠错码技术提构造一种在不可靠信道上新的组播认证方案(MALC: Multicast Authentication over Lossy Channels). 该方案不仅具有很高的通讯性能和计算性能,并且在存在部分数据包丢失的情况下也可以抵抗主动攻击. 本文提出了一种针对该方案特性的不可靠信道组播认证的形式安全模型,并在此安全模型下基于规约技术证明了该方案的安全性.

2 基础知识

本文提出组播认证方案 MALC 是基于哈希函数和签名算法等基本密码模块构造的. 本节将介绍这些相关概念并给出其形式定义. 另外,在对 MALC 进行安全性证明时用到概率论中的一个经典不等式 Hoefding 不等式也在本节给出.

哈希函数要求输入长度任意,输出长度固定,并且要求对于任何输入,计算哈希函数值容易(正向计算容易),攻击者找到两个不同输入且具有相同哈希值很难(抗碰撞性). 下面的定义 1 给出了哈希函数的严格数学定义.

定义 1 (抗碰撞的哈希函数)函数 $H: \{0,1\}^* \rightarrow \{0,1\}^k$ 是哈希函数,如果满足下列性质

(1)对于所有 $x \in \{0,1\}^*$,可以在以 $|x|+k$ 为变量的多项式时间内计算 $H(x)$

(2)对于任意概率多项式时间(PPT)算法 A,任意多项式 q 和足够大的 k 有

$$Pr[(x_1, x_2) \leftarrow A: x_1 \neq x_2 \wedge H(x_1) = H(x_2)] < \frac{1}{q(k)}$$

根据攻击者的能力不同,有多种满足不同安全性的签名体制. 如果攻击者可以选择任何消息要求签名者签名,并且可以根据以前的签名结果来选择之后的签名消息,称为自适应选择消息攻击,是一种安全性最高的签名体制. 下面给出自适应选择消息攻击安全的签名体制.

定义 2 (抗自适应选择消息攻击的消息签名体制)[10]消息签名体制由算法三元组(gen, sig, ver)组成,其中 gen 是密钥生成算法,随即生成签名密钥和验证密钥; sig 是签名算法, ver 是验证算法. 签名体制是自适应选择明文攻击安全性的,如果对于任意概率多项式时间(PPT)算法 A,任意多项式 q 和足够大的 k 有:

$$Pr[\mathrm{ver}(PK, M) = 1: (SK, PK) \in \mathrm{gen}(1^k); M \leftarrow A^{\mathrm{sig}_{SK}}(PK)] \leqslant \frac{1}{q(k)}$$

其中, k 是安全参数,等于密钥的长度. 记号 $A^{\mathrm{sig}_{SK}}$ 表示算法 A 可以对签名密钥为 SK 的签名算法 sig 询问消息的签名,算法 A 最后输出的消息不能是曾经由 sig 签名的消息. □

定义 3 (Hoefding 不等式)[11]令 $X_1, X_2, \cdots, X_n$ 是 n 个独立的同分布随机变量,每个随机变量分布在区间 $[a, b]$,令 μ 为每个随机变量的数学期望值. 对于任何的 $\varepsilon > 0$,有

$$Pr\left[\sum_{i=1}^{n} X_i - n \cdot \mu > n \cdot \varepsilon\right] < e^{-\frac{2\varepsilon^2}{(b-a)^2 n}}$$

定义 4 (纠错码)纠错码编码算法 $C: \{0,1\}^n \rightarrow \{0,1\}^m$ 将长度为 n 的消息映射为长度为 m 的编码, $m > n$. 如果接收到的编码 Y 与 $C(X)$ 最多有 e 个不同,解码算法可以从 Y 恢复 X. □

3 抗部分碰撞的哈希函数簇

抗碰撞的哈希函数要求找到具有完全相同哈希值的两个不同输入消息不可行;而抗部分碰撞哈希函数簇要求找到具有部分哈希值相同的两个不同输入消息不可行. 因此利用抗部分碰撞哈希函数簇进行消息认证,当部分哈希值丢失也能利用剩余哈希值进行验证.

定义 **5**　(哈希函数簇抗部分碰撞性) 设 $F=\{f_j \mid 1 \leqslant j \leqslant n\}$ 是单比特输出的哈希函数簇. F 是 (v, n)-抗部分碰撞性哈希函数簇, 如果对任何概率多项式时间的算法, 足够大的 n, 找到一对不同的消息 (m', n'), v 部分碰撞成功的概率可以忽略, 即找到 $m, m' \in \{0,1\}^*$, $m' \neq m$ 且 $(f_{r1}(m), f_{r2}(m), \cdots, f_{rv}(m)) = (f_{r1}(m'), f_{r2}(m'), \cdots, f_{rv}(m'))$, 其中 $1 \leqslant r_1 < r_2 < \cdots < r_v \leqslant n$, 成功的概率可以忽略. □

下面给出一个抗部分碰撞性哈希函数簇及其充分条件.

定理 **1**　设 $F=(f_i \mid 2 \leqslant i \leqslant n)$ 是单比特输出的哈希函数簇, 其中 $f_i(m)=f(i, m)$, f 是单比特输出的安全哈希函数. F 是 (v, n)-部分抗碰撞性哈希函数簇, 如果 $v=n/2+n\cdot\varepsilon 0<\varepsilon \leqslant 1/2$.

证明　(反证法)

假设存在概率多项式时间的算法 A, 一个多项式 q, 以及存在无数多个 n, A 成功地计算部分碰撞的概率大于 $1/Q(n)$.

根据哈希函数的 Random Oracle 模型[12], 哈希函数簇 $F=\{f_i \mid 1 \leqslant i \leqslant n\}$ 中的 n 个哈希函数 f_i 都是在 $[0, 1]$ 上均匀分布的独立同分布随机变量. 我们定义随机变量 X_i: $f_i(m')=f_i(m)$ 时 $X_i=1$, 否则 $X_i=0$. X_i 数学期望为 $\mu=1/2$, 那么根据上述假设有:

$$Pr\left[\sum_{i=1}^{n} X_i \geqslant n/2 + n\cdot\varepsilon\right] > \frac{1}{q(n)} \qquad (1)$$

而根据 Hoeffding 不等式有: $X_1, X_2, \cdots, X_n$ 是 n 个独立的同分布随机变量, 每个随机变量分布在区间 $[0, 1]$, 每个随机变量的数学期望值 $\mu=1/2$, 对于任何的 $\varepsilon>0$, 有

$$Pr\left[\sum_{i=1}^{n} X_i \geqslant n/2 + n\cdot\varepsilon\right] < e^{-2\varepsilon^2 n} \qquad (2)$$

式(1)和式(2)矛盾, 所以假设不成立.　　证毕.

需要注意的是, 上述哈希函数簇的实际安全性取决于 n. 一般来说, n 要和实用哈希函数输出长度相当(一般要大于 160 位).

4　组播认证方案 MALC

组播通讯是一种面向多个用户的群组协议, 其中存在一个发送者发送消息, 其他的用户都是作为接收者接收消息. 网络安全模型中, 假设攻击者完全控制网络, 所有的消息都发送给攻击者, 所有的接收的消息都通过攻击者转发, 攻击者可以截留、窜改、伪造或者延迟消息. 一般的网络安全协议不允许通讯消息丢失, 如果消息丢失, 协议就无法正常进行, 从而导致异常中止. 基于组播协议的视频音频通讯对实时性要求较高, 因此往往采用 UDP 协议传输数据包. 由于 UDP 协议对丢失数据包没有重传机制, 因此与普通网络安全协议不同, 组播协议需要在部分数据包丢失的情况下依然可以正常进行. 在这样的环境中, 接收者接收到的消息数量上可能少于发送的消息数量, 其中还可能包含伪造或者窜改的消息. 为了组播认证方案的正常工作, 我们假设网络丢失数据包概率存在一个最大上界. 一方面这个假设在实际网络通讯运行中能够满足的; 另一方面对于少数极端情况, 当丢失数据包概率太大时, 接收质量达不到基本要求, 此时组播认证也失去意义. 因此该假设是合理的. 另外需要说明的是, 由于主动攻击的存在, 数据包的丢失可能完全是攻击者恶意导致的, 因此不能假设数据包的丢失满足特定的概率分布, 比如随机分布或者突发式的单点分布等.

定义 **6**　(丢包率 γ) 如果组播发送的数据包数目是 n, 那么任何接收者至少可以接收到 $(1-\gamma)n$ 个数据包, 其中 $0 \leqslant \gamma \leqslant 1$. □

本方案的基本思想是, 首先将要组播发送的消息按照顺序分解为大小相同的多个数据块. 每个数据块进一步分解为多个数据包, 因此在本方案中数据块被称为数据包列. 其中数据包可以作为单独的 IP 包发送. 从最后一个数据包列开始, 分别用哈希函数簇来计算每一个数据包列中所有数据包的哈希值, 并将它们分散添加到前一个数据包列的数据包中. 可见, 每一数据包就会包含后一个数据包列中所有数据包的部分认证信息. 重复上述计算, 最后将第一个数据包列生成哈希函数进行签名, 将签名值和哈希值用纠错码编码后分散到第一个数据包列中各个数据包中传输. 采用纠错码将签名数据包 p_0 编码后分散发送的目的是保证接收者可以可靠的接收到 p_0, 虽然采用重复多次发送 p_0 的方法也能达到这个目的, 但是通讯性能较差.

下面是方案的具体描述.

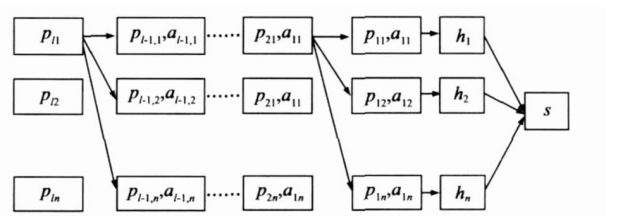

图 1　组播认证 MALC

MALC(组播认证方案)

初始化: $F=\{f_j \mid f_j(m)=f(j, m), 1 \leqslant j \leqslant n\}$ 是单比特输出的哈希函数簇, 其中 f 是单比特输出哈希函数. 基本消息签名算法为 $(\mathrm{gen}, \mathrm{sig}, \mathrm{ver})$, 安全参数为 k_1. $(sk, pk) \leftarrow \mathrm{gen}(2^{k_1})$, sk 是发送者签名密钥, p_k 是验证密钥. H 是 k_2 比特输出的哈希函数, 数据包丢包率 γ, 数据包集数目为 n. 纠错码编码算法的输入消息长度为 m_1 位, 输出长度为 m_2 位, 最多可纠错 γm_2 位. 发送者

签名密钥 s_k 为秘密参数, 仅发送者知道, 其他均为公开信息.

组播签名算法: 发送者对组播消息 M 生成签名. 发送者首先选择数据包集参数 l.

(1) 将 M 分解为 l 个数据包列 $P_1, P_2, \cdots, P_l$, 即 $M = \langle P_1, P_2, \cdots, P_l \rangle$. 每一个数据包列 P_i 由 n 个数据包组成.

(2) 从最后一个数据包列 P_1 到第二个数据包列 P_2, 依次生成 $P_l, P_{l-1}, \cdots, P_2$ 的认证信息.

设 $a_{\bar{i}\bar{j}} \in \{0,1\}^n$ 是部分认证信息, 令 $a_{\bar{i}\bar{j}} = 0, j = 1, 2, \cdots, n$.

计算 $b^i_{rj} = f_j(a_{i+1,r}, p_{i+1,r}), i = l-1, l-2, \cdots, 1; j = 1, 2, \cdots, n; r = 1, 2, \cdots, n$ 令 $a_{\bar{i}\bar{j}} = (b^i_{1\bar{j}} \| b^i_{2\bar{j}} \| \cdots \| b^i_{n\bar{j}}) \in \{0,1\}^n$, $a_{\bar{i}\bar{j}} (1 \leqslant j \leqslant n)$ 作为 P_{i+1} 的第 j 部分认证信息, 和 P_i 中 $p_{\bar{i}\bar{j}}$ 组成一个新的数据包 $(p_{\bar{i}\bar{j}}, a_{\bar{i}\bar{j}}, i, j)$.

(3) 生成 P_1 的哈希值: $h_1 \| h_2 \| \cdots \| h_n = H(p_{11}, a_{11}) \| H(p_{12}, a_{11}) \| \cdots \| (p_{1n}, a_{1n})$

(4) 计算签名: $s = \mathrm{sig}_{sk}(h_1 \| h_2 \| \cdots \| h_n \| l \| n)$

(5) 令 $p_0 = (h_1 \| h_2 \| \cdots \| h_n \| l \| n, s)$, 通过纠错码编码算法将 p_0 编码为 $p'_0 = s_1 \| s_2 \| \cdots \| s_n$, 分别附加在第一个数据包列 P^s_1 各个数据包中. 算法输出为 $M^s = \langle p^s_1, p^s_2, \cdots, p^s_l \rangle$, 其中 $P^s_1 = \{(p_{1j}, a_{1j}, 1, j, s_j) \mid j = 2, \cdots, n\}$, $P^s_i = \{(p_{\bar{i}\bar{j}}, a_{\bar{i}\bar{j}}, i, j) \mid j = 2, \cdots, n\}$

组播验证算法: 输入为签名的组播消息; 输出为通过验证的数据包集合或者验证失败. 假设接收到的签名组播消息为 $M^{\bar{s}}_1 = \langle P^{\bar{s}}_1, P^{\bar{s}}_2, \cdots, P^{\bar{s}}_{\bar{l}} \rangle$, $P^{\bar{s}}_1 = \{(\bar{p}_{1j}, \bar{a}_{1j}, 1, j, \bar{s}_j) \mid j \in \{1, 2, \cdots, n\}\}$, $P^{\bar{s}}_{\bar{i}} = \{(\bar{p}_{\bar{i}\bar{j}}, \bar{s}_{\bar{i}\bar{j}}, i, j) \mid j \in \{1, \cdots, \bar{n}\}\}$. 令 $J_i = \{j \mid (\bar{p}_{ij}\bar{a}_{ij}, i, j) \in P^{\bar{s}}_i\}$, 表示接收到的第 i 个数据包列中数据包的序号集合, 由于可能存在丢包现象, 因此 J_i 是 $\{1, \cdots, \bar{n}\}$ 的子集. 输出为成果或者失败标志 $\{0,1\}$, 以及通过验证的数据包列的集合 $\bar{M} = \langle \bar{P}_1, \bar{P}_2, \cdots, \bar{P}_{\bar{l}} \rangle$, 算法开始时各数据包集 $\bar{P}_i$ 初始化为空集, 验证结果标志初始化为 0.

(1) 验证数字签名: 从 $P^{\bar{s}}_1$ 各个数据包中 $\bar{s}_j \in \{1, 2, \cdots, \bar{n}\}$ 的通过纠错码恢复 $p^{\bar{s}}_0$, 验证 $\mathrm{Ver}_{pk}(\bar{h}_1 \| \bar{h}_2 \| \cdots \| \bar{h}_n \| \bar{l} \| \bar{n}, \bar{s}) = 1$, 如果不成立则输出验证结果标志 1, 算法结束.

(2) 验证 $P^{\bar{s}}_1$ 的哈希值. 对收到第一个数据包列 $P^{\bar{s}}_1$ 中的每个数据包, 验证 $\bar{h}_j = H(\bar{p}_{1j}, \bar{a}_{1j})$. 如果成立, 则 $\bar{P}_1 = \bar{P}_1 \cup \{(\bar{p}_{1j}, 1, j)\}$, 否则验证失败标志设置为 1, 算法继续进行.

(3) 依次验证 $P^{\bar{s}}_2, \cdots, P^{\bar{s}}_{\bar{l}}$ 的认证信息. 在验证第 i 个数据包列 $P^{\bar{s}}_i$ 时, 先计算前一个数据包列 $P^{\bar{s}}_{i-1}$ 中通过验证输出数据包的序号集合 $J'_{i-1} = \{j \mid (\bar{p}_{i-1,j}, i-1, j) \in \bar{p}_{i-1}\}$, 并验证 $(1-\gamma)n \leqslant |J'_{i-1}|$ 是否成立, 如不成立输出失败标志 1, 验证算法结束. 每一个 $j \in J_i$, 如果对于所有 $r \in J'_{i-1}$ 式 $(\bar{a}_{\bar{i}\bar{j}})_r) = f_j(\bar{a}_{i+1,r}, \bar{p}_{i+1,r})$ 都成立, 则 $\bar{P}_i = \bar{P}_i \cup \{(\bar{p}_{\bar{i}\bar{j}}, i, j)\}$; 否则验证结果标志设置为 1, 算法继续进行.

(4) 最后输出验证结果标志和 $\bar{M} = \langle \bar{P}_1, \bar{P}_2, \cdots, \bar{P}_{\bar{l}} \rangle$. □

说明: 在上述验证算法中, 即使部分数据包被窜改导致验证结果为失败, 但是接收者也可以获得通过验证的部分数据包 $\bar{M} = \langle \bar{P}_1, \bar{P}_2, \cdots, \bar{P}_{\bar{l}} \rangle$.

MALC 参数选择: 在 MALC 中涉及各种参数, 他们的选择直接影响协议的安全. 下面分别给出安全相关参数的选择原则.

(1) k_1 和 k_2 确定与具体的签名体制和哈希函数有关. 比如用 RSA 签名, k_1 至少 1024 位或者 2048 位; 用 SHA1 作为哈希函数时, k_2 至少 160 位.

(2) 假设组播消息 M 总数据包数目为 m, 将 M 分解为 l 个数据包列, 如果最后一个数据包列中数据包数可以少于 n, 需要用全 0 填充至 n 个. 因此 $m = nl$.

(3) 纠错码输入位长 m_1 等于 $p_0 = (h_1 \| h_2 \| \cdots \| h_n \| l \| n, s)$ 的位长, 因为其中签名 s 为 k_1, 哈希值 h_i 均为 k_2, l 和 n 各 8 位, 分隔符不计, 故 m_1 长度是 $k_1 + nk_2 + 16$ 位.

(4) 数据包丢包率为 γ, 为了保证能够纠错码解码算法可以正确的解码 p_0 数据包, 纠错码输出位长 m_2 和可纠错位数 e 应该满足 $e \geqslant \gamma m_2$.

(5) 数据包丢包率 γ 取决于实际网络服务质量, 且根据抗部分碰撞哈希函数簇定理 1, 必须满足 $\gamma \leqslant 1 - \frac{1}{2} - \varepsilon$, 其中 $0 < \varepsilon \leqslant \frac{1}{2}$ 是一个的正数. 因此, 丢包率应该小于 1/2.

(6) 数据包丢包率 γ、n 和部分碰撞函数的安全强度密切相关, 参照哈希函数的安全位数, $(1-\gamma)n$ 至少应该大于 160 位. 安全强度与成反比, 与 n 成正比. 当 n 足够大时, 仅仅要求数据包最大的丢包率小于 1/2, 对方案的安全性是足够. 然而, 需要指出的是对于较小的 n, 丢失率仅仅小于 1/2 是不能满足安全需要的. 比如, $n = 160$, 当丢包率为 0.4 时, 攻击者随机选择一个数据包, 伪造成功的概率大于 0.005, 最大丢包率为 0.3 时, 伪造成功概率大于 10^{-8} (采用正态分布进行估计). 可见选择合适的 n 对安全强度的影响很大. 另一方面, 从性能角度出发也不能任意选择过大的 n.

5 安全模型和安全证明

基于计算复杂性理论的可证明安全是目前保证密码体制安全性的重要方法. 为了证明方案 MALC 的安全性, 我们将给出一个基于计算复杂性理论的组播认证

安全定义, 该定义可以用于不可靠通讯时有数据包丢失的情况. 文献[3]中流签名安全性定义没有考虑数据包丢失的情况, 因此不适合我们的组播认证体制.

定义 **7**　(不可靠信道上的组播认证体制 MAS) 在不可靠信道通讯存在数据包丢失情况下, 一个组播认证方案所以可用一个三元(Mgen, Msig, Mver) 组表示, 每个组成元素是一个概率多项式时间的算法, 并且满足下性质:

(1)Mgen 是密钥生成算法. 输入为安全参数 1^k, 输出一对密钥 $(SK, PK) \in G(1^k)$, 其中 SK 为签名私钥, PK 为验证公钥.

(2) Msig 是组播认证签名算法. 输入为安全参数 1^k, 签名私钥 SK, 最大丢失率是 γ, 数据包列长度 n, 以及一个组播消息 M. 将输入消息表示为 l 个数据包列, $M = \langle P_1, P_2, \ldots, P_l \rangle$, 每个数据包列 P_i 由 n 个数据包组成, $P_i = \{p_{ij} \mid j = 1, 2, \ldots, n\}$. 输出一个签名的组播消息 $M^s = \langle P_1^s, P_2^s, \ldots, P_l^s \rangle$.

(3)Mver 是组播认证验证算法. 输入验证公钥 PK、组播签名消息 $\bar{M^s} = \langle \bar{P_1^s}, \bar{P_2^s}, \ldots, \bar{P_l^s} \rangle$, 其中可能丢失了部分数据包. 输出是 1(表示验证失败) 或者 0(表示验证成功).

定义 **9**　(不可靠信道上组播认证的安全性) 对于不可靠信道上组播认证体制 MAS, 如果攻击者可以选择任意多的消息并获得相应的组播签名, 但是攻击者不能在多项式时间内伪造另外一个可以通过验证的组播签名, 那么 MAS 是抗选择消息主动攻击安全的. 形式化地, 假设攻击者算法 A 的输入为公钥 PK, 并且可以多次询问组播认证签名算法 Msig 对不同组播消息的签名, 然后输出伪造的签名组播消息 M^s(该签名组播消息 M^s 不能包含于攻击者算法 A 询问问答机 Msig 输出过的任何签名消息 M_i^s, $M^s \not\subset M_i^s$,) 可以通过 Mver 验证成功, 那么一个组播认证体制 MAS 安全的条件是, 对于所有的伪造攻击算法 A, 所有的多项式 q, 所有足够大的整数 k, 满足:

$$Pr[\mathrm{Mver}(PK, M^s) = 1 : (SK, PK) \in \mathrm{Mgen}(l^k); M^s \leftarrow A^{\mathrm{Msing}0}(PK)] \leqslant \frac{1}{q(k)}$$

在假设 MALC 基于的基本密码体制是安全可以证明 MALC 的安全性.

定理**2**　如果哈希函数 H 是抗碰撞安全的, 单比特输出哈希函数向量 F 是抗部分碰撞哈希函数向量, 消息签名体制(gen, sig, ver) 是在选择消息攻击下不可伪造的, 那么 MALC 在选择流攻击下是不可伪造的.

证明　假设 MALC 是不安全的, 即存在一个攻击算法 A, 一个多项式 q, 无穷多的整数 k,

$$Pr[\mathrm{Mver}(PK, M^s) = 1 : (SK, PK) \in \mathrm{Mgen}(l^k); M^s \leftarrow A^{\mathrm{Msing}0}(PK)] > \frac{1}{q(k)}$$

即攻击算法 A 先询问组播签名问答机 Msig 并获得 ζ 个组播签名 $M_1^s, M_2^s, \ldots, M_\zeta^s$, 然后伪造出一个新的组播签名 $M^s \not\subset M_i^s$, $i = 1, 2, \ldots, \zeta$, 其中 $M^s = \langle P_1^s, P_2^s, \ldots, P_l^s \rangle$. 把伪造的组播签名与询问问答机 Msig 获得的组播签名比较, 那么存在三种情形: 第一种情形, 从伪造的组播签名 M^s 用纠错码恢复的第一个数据包 $p_0 = (h_1 \mid \ldots \mid h_n \mid l \mid n, s)$ 与 M_i^s, $i = 1, 2, \ldots, \zeta$ 中的对应数据包都不相同. 第二种情形, 伪造的组播签名 $\hat{M}^s$ 的第一个数据包与问答机输出的某个组播签名的第一数据包相同, 但是第一数据包列中有数据包 (p_{1v}, a_{1v}) 与 $\hat{M}^s$ 相同下标的数据包不同; 第三种情形, M^s 与问答机输出的某个组播签名 $\hat{M}^s$ 相比, 第一个数据包相同, 并且前 $u-1$ ($u > 1$) 个数据包列是 $\hat{M}^s$ 的之序列, 即, $P_i^s \subseteq \hat{P}_i^s$, $i = 1, 2, \ldots, u-1$. P_u^s 中的某个数据包 (p_w, a_{uw}) 与 $\hat{M}^s$ 中相同下标的数据包不相同. 注意到在 p_0 中包含了组播签名的长度(数据包集个数) l, 如果伪造的组播签名与某个 $\hat{M}^s$ 第一个数据包相同, 则长度 l 也相同, 所以不可能伪造出一个组播签名 M^s 长度 l 大于某个 $\hat{M}^s$ 的长度 $\hat{l}$ 且它们的前 $\hat{l}$ 个数据包集(包括第一个数据包) 都相同.

如果是第一种情形, 可以利用 A 来构造一个算法 A_1, 它能够在选择消息攻击下伪造消息签名体制(Gen, Sig, Ver) 的一个签名, 从而与(Gen, Sig, Ver) 是选择消息攻击不可伪造的假设矛盾. 算法 A_1 以 A 为子程序, 当 A 询问流签名问答机 Msig 一个流 M_i 的签名时, 由 A_1 模拟问答机 Msig 做出回答. A_1 执行流签名算法 MALC 第一步第二步, 计算认证信息. 因为是对(Gen, Sig, Ver) 的选择消息攻击, 所以 A_1 可以询问消息签名问答机 Sig, 这样在 MALC 签名的第三步时, 询问问答机 Sig 得到签名, 然后把签名流 M_i^s 返回给 A. 最后 A 输出一个伪造的签名流, 它的第一个数据包 $p_0 = (h_1 \mid \ldots \mid h_n \mid l \mid n, s)$ 与 M_i^s, $i = 1, 2, \ldots, \zeta$ 中的第一个数据包都不相同, 因此, 该数据包对问答机 Sig 来说也是一个新的消息, 所以 A_1 可以把 $p_0 = (h_1 \mid \ldots \mid h_n \mid l \mid n, s)$ 输出作为伪造的消息签名.

对于第二种情形, 可以利用 A 来构造一个算法 A_2, 它能够找到哈希函数的一对碰撞值. 构造方法同前, 当 A 输出伪造的签名流 M^s, 假设与问答机 Msig 输出的某个签名流的 $\hat{M}^s$ 相比有 $p_0 = \hat{p}_0$, $p_{1j} \neq \hat{p}_{1j}$, A_2 就输出 p_{1j} $\hat{p}_{1j}$, 他们的哈希值相同, 从而找到一对碰撞.

对于第三种情形, 可以构造一个算法 A_3, 它能够找到哈希函数向量部分分量的一对碰撞值. 当 A 输出伪造的签名流 M^s, 假设与问答机 Msig 输出的某个签名流 $\hat{M}^s$ 的相比有 $p_0 = \hat{p}_0$, $P_i^s \subseteq \hat{P}_i^s$, $i = 1, 2, \ldots, v-1$, 并且

$(p_{ij}, a_{ij}) \neq (\hat{p}_{ij}, \hat{a}_{ij})$. 输出$(p_{ij}, a_{ij}) \neq (\hat{p}_{ij}, \hat{a}_{ij})$为一对部分碰撞. 证毕.

6 性能的分析和比较

密码方案的性能通常可用计算复杂性和通讯复杂性来度量. 为了方便对各种组播协议的比较, 可以用签名次数和哈希计算量来表示计算复杂性, 用附加传输的信息长度来代表通讯复杂性. 因为本方案使用了输出长度不同的两种哈希函数, 对于同一个消息计算单比特输出的 n 维哈希函数簇的哈希值可以用输出长度为n 的单一哈希函数代替, 因此哈希计算量可采用哈希次数和每次哈希值长度的乘积表示. 另外一类比较各种方案性能的指标是发送方和接收者需要缓存的数据包数目, 我们用发送延迟和接收延迟表示. 在目前已经提出的各种可以支持不可靠通讯环境、抗主动攻击的组播认证方案中, 典型的方案有: 1)对所有数据包分别签名的方案; 2)Merkle 树方案[3]; 3) 纠错码方案[7]; 4)擦除容忍方案[9]. 其中第一个方案(称为逐包签名)对每一个包都签名, 计算复杂性太高, 并不实用, 因此仅仅作为比较的参照. 下面从通讯复杂性、计算复杂性和发送接收延迟等方面对这些典型方案和本文提出的 MALC 方案进行比较. 由于这些方案中接收者和发送者的计算复杂性基本相同, 因此仅考虑发送者的计算量. 假设一次组播的数据包数为 m(如果组播消息数据库包总数大于 m, 则需要将组播消息分解成多个 m 大小的序列依次组播), 签名的长度为 k_1, 哈希值长度 k_2, 丢包率 γ, 纠错码方案特定参数伪造包率 β(如果发送 m 的数据包, 攻击者最多可伪造$(\beta - 1)m$ 个数据包), 以及本方案的特定参数 n. 性能比较见表 1.

表 1 各种组播认证方案性能比较

	逐包签名	Merkle 树	纠错码方案	擦除容忍	MALC
签名次数	m	1	1	$\log m$	1
哈希量	mk_2	$2mk_2$	$mk_2 m$	$\log m$	mm
通讯量	mk_1	$m(k_1 + k_2 \log m)$	$\beta k_2 m/(1-\gamma)^2$	$\log m$	mm
发送延迟	1	m	m	m	m
接收延迟	m	m	n	n	n

对于一个组播消息, 本方案仅签名一次, 与 Merkle 树和纠错码方案一样, 而逐包签名的计算效率最低. 在哈希函数计算量和通讯量方面, 由于本方案采用了不同参数, 因此不能直接和其他方案比较. 根据 MALC 的参数取值的原则, $(1-\gamma)n$ 与其他方案的哈希值长度 k_2 相近. 因此, 本文的方案在哈希函数计算量上和其他方案相近, 而通讯复杂性上比 Merkle 树方案有较大提高, 并且略优于纠错码方案. 在接收延迟方面, Merkle 树和纠错码方案需要接收到全部数据包才能验证, 而本文的方案接收者只需缓存少量数据包后就可以对后继的数据包进行验证. 表中给出的擦除容忍方案的性能是理论上可达到的最好性能, 实际性能随着丢包率的增大还会有明显的降低. 因此总体来看, 与其他抗主动攻击的组播认证方案相比, 本文提出的方案在计算复杂性、通讯复杂性和接收延迟等方面都具有很好的性能.

7 结论

与传统签名或者消息认证不同, 在接收的数据存在部分丢失或者窜改的情况下对剩余数据的有效是一个全新的理论问题; 而不可靠信道中的组播认证是急需解决的实际问题. 因此, 在不可靠信道中组播认证是一个在理论和实践两方面都具有较大意义的研究问题. 本文基于抗部分碰撞哈希函数的概念, 提出了一种新的组播认证方案. 基于本文提出的存在数据包丢失情况下抗主动攻击的组播认证安全的形式定义, 证明了该方案的安全性. 与其他同类方案相比, 本文提出的方案在安全性、计算性能和通讯性能都具有较大的优势. 我们进一步的研究方向是改进哈希图结构以及参数以提高协议抵抗突发性大量丢包的能力.

参考文献:

[1] Canneti R. Garay. J., Itkis G., et al. Multicast security: a taxonomy and some efficient construction[A]. In Proceedings of the 6th ACM Computer and Communications Security Conference[C]. Singapore: ACM Press, 1999. 93~ 100.

[2] Gennaro R., Rohatgi P. How to sign digital streams[A]. Advances in Cryptology, CRYPTO' 97[C], Berlin, Springer-Verlag, 1997, 180~ 197.

[3] C. K. Wong and S. S. Lam. Digital signatures for flows and multicasts[A]. In Proceedings of the 1998 International Conference on Network Protocols (ICNP' 98)[C], pages 198~ 209, Austin, Texas, Oct. 1998.

[4] Golle P., Modadugu N. Authenticating streamed data in the presence of random packet loss[A]. ISOC Network and Distributed System Security Symposium[C], 2001, 1322.

[5] A. Perrig, R. Canetti, J. Tygar, and D. Song. Efficient authentication and signing of multicast stream over lossy channels[A]. In IEEE Symposium on Security and Privacy[C], pages 56~ 73, 2000.

[6] 李先贤, 怀进鹏, 高效的动态组播群通信认证签字方案, 软件学报, 2001, v10, 1486~ 1494. Li X. X., Huai J. P., Efficient authentication signature schemes for dynamic multicast groups[J]. Journal of Software, 2001, v10, 14861494. In Chinese.

[7] A. Lysyanskaya, R. Tamassia and N. Triandopoulos, Multicast Authentication in Fully Adversarial Networks[A], In Proceed-

ings of IEEE Symposium on Security and Privacy[C] p. 241-255, Oakland, May 2004.

[8] 何永忠, 冯登国, 一个组播源认证方案的安全分析, 计算机工程, 2006, 10He Y.Z., Feng D.G., Security Analysis on A Multicast Source Authentication Scheme[J], Computer Engineering, 2006, 10.

[9] Desmedt Y., and Jakimoski G., Non-degrading Erasure-Tolerant Information Authentication with an Application to Multicast Stream Authentication over Lossy Channels[A], Topics in Cryptology[C], CT-RSA 2007, LNCS 4377, 2006

[10] Goldwasser, Micali, Rivest. A digital signature scheme secure against adaptive chosen-message attacks[J]. SIAM Journal of computing, 1988.4, 17(2): 281-308.

[11] Hoeffding W. Probability inequalities for sums of bounds random variables [J]. J Amer Statist Assoc, 1963; 50: 13-30

[12] Bellare M., Rogaway P., Random oracles are practical: a paradigm for Designing efficient protocols[A]. In 1st Conference on Computer and Communications Security[C], ACM, 1992. 62-73.

作者简介:

何永忠　男, 1969 年出生于重庆, 博士, 北京交通大学计算机学院讲师, 研究方向: 信息安全.

E-mail: yzhhe@bjtu.edu.cn

(上接第 1278 页)

系统频率同步误差分析[J]. 信号处理, 2007, 23(6): 927-931.

Zhang Yongsheng, Wang Min, Liang Diannong, Dong zhen, Huang Haifeng. Analysis of frequency synchronization error in spaceborne parasitic interferometric SAR system[J]. Signal Processing, 2007, 23(6): 927-931. (in Chinese)

[16] 张永胜, 梁甸农, 孙造宇, 董臻. 时间同步误差对星载寄生式 InSAR 系统相位误差的影响分析[J]. 宇航学报, 2007, 28(2): 151-155.

Zhang Yongsheng, Liang Diannong, Sun Zaoyu, Dong Zhen. Effect of time synchronization error on interferometric phase of spaceborne parasitic InSAR system[J]. Journal of Astronautics, 2007, 28(2): 151-155. (in Chinese)

[17] 袁孝康, 星载合成孔径雷达导论[M]. 国防工业出版社, 北京, 2003.

Yuan Xiaokang, Introduce to the Spaceborne Synthetic Aperture Radar[M], Beijing: National Defense Industry Press, 2003. (in Chinese)

[18] 路兴强, 余安喜, 王敏, 梁甸农. 小卫星分布式雷达仿真系统的集群技术实现[J]. 现代雷达, 2006(8): 1-3.

Lu Xingqiang, Yu Anxi, Wang min, Liang Diannong. A simulation system for distributed spaceborne radar based on PC cluster[J]. Modern Radar, 2006(8): 1-3. (in Chinese)

[19] 李志林, 朱庆. 数字高程模型[M]. 武汉大学出版社, 武汉, 2003.

[20] 陈杰, 周荫清, 李春生, 星载 SAR 自然地面场景仿真方法研究[J]. 电子学报, 2001, 29(9): 1202-1205.

Chen Jie, Zhou Yinqing, Li Chunsheng. Spaceborne synthetic aperture radar image simulation of natural ground scene[J]. Acta electronica Sinica, 2001, 29(9): 1202-1205. (in Chinese)

[21] 陈斌, 一种离散化的最小曲率插值方法[J]. 煤田地质与勘探, 2000, 28(1): 49-54.