椭圆曲线密码与 SHA 算法在语音加密中的应用①

孙立宏

(阜新高等专科学校 师范部，阜新 123000)

摘 要：VoIP 目前已经发展成为一种专门的语音通信技术，其应用范围越来越广。但是研究发现，除服务质量等问题外，安全问题是企业在做 VoIP 决策时重点考虑的内容。在现实应用中，安全状况不能达到企业的应用标准是企业暂时不想部署 VoIP 的原因。深入研究安全散列算法，论述了其实现的过程，并对其进行了 VC 的实现。使用椭圆曲线密码与 SHA 结合，加密语音数据，实现语音的实时加密。

关键词：椭圆曲线密码；SHA-1；VoIP；散列函数；G.729 协议

Application of ECC and SHA to Voice Encryption

SUN Li-Hong

(Department of Teaching, Fuxin College, FuXin 123000, China)

Abstract: VoIP has developed into a specialized voice communication technology currently, whose application is wider and wider. But the research shows that the security is the main content considered by the enterprise when VoIP is decided besides the service quality issue. In the real application, the reason that the enterprise doesn't want to deploy VoIP is that the security situation can't reach the application standard. This article studies the secure hash algorithm, discusses its implementation process and carries out the implementation of VC. Real-time voice encryption is completed by using the combination between elliptic curve cryptography and SHA, and the encryption between voice and data.

Keywords: ECC; VoIP; Hash function; G.729

基于 IP 网络的语音传输(VoIP)技术目前已经发展成为一种专门的语音通信技术，其应用范围越来越广。VoIP 的一个优势是，用户无需在互联网接入服务费用之外无需再支付其它费用，就像用户发送电子邮件无需再支付其它费用那样。在企业和消费者领域在日益普及。但是研究发现，除服务质量等问题外，安全问题是企业在做 VoIP 决策时重点考虑的内容。在现实中，许多客户表示，安全状况不能达到企业的应用标准是他们暂时不想部署 VoIP 的原因。日前，面临的安全议题主要有 4 个：拒绝服务攻击、非法接入、话费诈欺或窃听等威胁[1-3]。

VoIP 语音服务的传输与安全机制与传统电话服务截然不同。由于协议本身并没有防范攻击的能力，因而未加密的语音数据流量在传输时极易被截取或侦听。尽管目前来看，数据包侦听在安全事件中所占的比例并不高，但由于这种侦听方式技术难度不大，因此，在 VoIP 逐渐成为语音服务的主流之后，语音数据包侦听可能会成为 VoIP 的一个主要安全威胁。根据。安全联盟的调查，目前部分黑客还掌握了如何攻击特定目标及特定网络内语音流量的技术，使。面临安全威胁进一步加大。模拟话机存在并线窃听的问题，当企业用户使用了数字话机之后，由于都是厂家私有的协议，很难通过简单的手段来侦听。但 VoIP 环境下，这个问题又被提了出来。一个典型的 IP 呼叫需要信令和媒体流两个建立的步骤，RTP/RTCP 是在基于包的网络上传输等时话音信息、的协议。由于协议本身是开放的，即使是一小段的媒体流都可以被重放出来而不需要前后信息的关联。如果有人在数据网络上通过

① 收稿时间:2010-08-23;收到修改稿时间:2010-10-10

Sniffer 的方式记录所有信息、并通过软件加以重放，会引起员工对话音通信的信任危机。

1 SHA-1算法

1.1 HA-1

在信息安全技术中，经常需要验证消息的完整性，散列(Hash)函数提供了这一服务，它对不同长度的输入消息，产生固定长度的输出。这个固定长度的输出称为原输入消息的“散列”或“消息摘要”(Message digest)。一个安全的哈希函数 H 必须具有以下属性：

1) H 能够应用到大小不一的数据上。

2) H 能够生成大小固定的输出。

3) 对于任意给定的 x，H(x)的计算相对简单。

4) 对于任意给定的代码 h，要发现满足 H(x)=h 的 x 在计算上是不可行的。

5) 对于任意给定的块 x，要发现满足 H(y)=H(x)而 y=x 在计算上是不可行的。

SHA(Secure Hash Algorithm)算法由 NIST 开发，并在 1993 年作为联邦信息处理标准公布。在 1995 年公布了其改进版本 SHA-1。SHA 与 MD5 的设计原理类似，同样也按留二位数据块为单位来处理输入，但它产生 160 位的消息摘要，具有比 MD5 更强安全性的 SHA-l 是一种数据加密算法，该算法的思想是接收一段明文，然后以一种不可逆的方式将它转换成一段（通常更小）密文，也可以简单的理解为取一串输入码（称为预映射或信息），并把它们转化为长度较短、位数固定的输出序列即散列值（也称为信息摘要或信息认证代码）的过程[2-4,7]。

单向散列函数的安全性在于其产生散列值的操作过程具有较强的单向性。如果在输入序列中嵌入密码，那么任何人在不知道密码的情况下都不能产生正确的散列值，从而保证了其安全性。SHA 将输人流按照每块 512 位(64 个字节)进行分块，并产生 20 个字节的被称为信息认证代码或信息摘要的输出。

1.2 SHA-1 计算过程

该算法输入报文的最大长度不超过 264 位，产生的输出是一个 l60 位的报文摘要。输入是按 512 位的分组进行处理的。SHA-1 是不可逆的、防冲突，并具有良好的雪崩效应。如图 1 所示。

通过散列算法可实现数字签名实现，数字签名的原理是将要传送的明文通过一种函数运算(Hash)转换成报文摘要(不同的明文对应不同的报文摘要)，报文摘要加密后与明文一起传送给接受方，接受方将接收的明文产生新的报文摘要与发送方的发来报文摘要解密比较，比较结果一致表示明文未被改动，如果不一致表示明文已被篡改。

步骤 1：添加填充位(一个 l 和若干个 0)。在消息的最后添加适当的填充位使得数据位的长度满足 length=448 mod 512。

步骤 2：添加长度，一个 64 位块，表示原始消息长度，64 位无符号整数。

步骤 3：初始化 MD 缓冲区。一个 160 位 MD 缓冲区用以保存中间和最终 Hash 函数的结果。它可以表示为 5 个 32 位的寄存器（A，B，C，D，E）。
初始化为：

A=67452301

B=EFCDAB89

C=98BADCFE

D=10325476

E=C3D2E1F0

前四个与 MD5 相同，但存储为大数在前的形式。

步骤 4：以 512 位数据块为单位处理消息。四轮，每轮 20 步。四个基本逻辑函数：f1，f2，f3，f4

步骤 5：输出。全部 L 个 512 位数据块处理完毕后，输出 160 位消息摘要[5-7]。

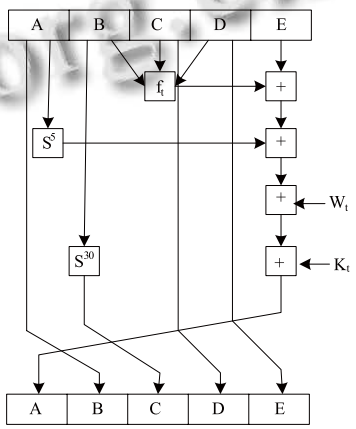

图 1 基本 SHA 操作(单步)

2 语音的实时加密

2.1 G729a 简介

G.729 协议是由 ITU-T 的第 15 研究小组提出的，并在 1996 年 3 月通过的 8Kbps 的语音编码协议。

G.729 协议使用的算法是共轭结构的算法码本激励线性预测(CS-ACELP)，它基于 CELP 编码模型。由于 G.729 编解码器具有很高的语音质量和很低的延时，被广泛地应用在数据通信的各个领域，如 IP phone 和 H.323 网上多媒体通信系统等。

2.2 语音实时加密的原理和方式

下面描述一个用椭圆曲线进行加密通信的协议：

① 用户 A 选定一条椭圆曲线 Ep(a,b)，并取椭圆曲线上一点，作为基点 P。

② 用户 A 选择一个私有密钥 k，并生成公开密钥 Q=kP。

③ 用户 A 将 Ep(a,b)和点 Q，P 传给用户 B。

④ 用户 B 接到信息后，将待传输的明文编码到 Ep(a,b)上一点 M，并产生一个随机整数 r (r<n)。

⑤ 用户 B 计算点 C1=M+rQ: E0=rP。

⑥ 用户 B 将 C1、E0 传给用户 A。

⑦ 用户 A 接到信息后，计算 C1-kE0，结果就是点 M。因为　C1-kE0=M+rQ-k(rP)=M+rQ-r(kP)=M

再对点 M 进行解码就可以得到明文。

在这个加密通信协议中，如果有一个偷窥者 H，他只能看到 Ep(a,b)、Q、P、C1、E0，而通过 Q、P 求 k 或通过 E0、P 求 r 都是相对困难的。因此，H 无法得到 A、B 间传送的明文信息。

然而在加密大文件过程中，需要将明文分解成 N 段，并将每一段明文都编码到 Ep(a,b)上对应点 Mi，协议的步骤⑤⑥⑦要相应的重复 N 次计算，因此极大的影响了加密的速度同时成倍的增加了数据的传输量。如果仅使用一个 r，即：

C1=M1⊕Rq

C2=M2⊕rQ

C3=M3⊕rQ

……

很明显，对于偷窥者 H、假设他得到足够的明文信息：M1,M2,M3……时，C1,C2,C3……之间将有规律可循，这个协议容易受到攻击。

如果引入随机数 D，则加密过程为：

D0=随机数

C1=M1⊕rQ⊕D1, D1=f(D0)

C2=M2⊕rQ⊕D2, D2=f(D1)

C3=M3⊕rQ⊕D3, D3=f(D2)

……

函数 d=f(M)满足关系：

1）给定 M 生成唯一的 d

2）d 的生成空间足够大

这样当 M1，M2，M3……为特定的文件时，C1，C2，C3……之间也无规律可循的，在这里，函数 f() 使用散列算法，它通过把一个单向数学函数应用于数据，将任意长度的一块数据转换为一个定长的、不可逆转的数据。这段数据通常叫做消息摘要。消息摘要代表了原始数据的特征，当原始数据发生改变时，重新生成的消息摘要也会随之变化，即使原始数据的变化非常小，也可以引起消息摘要的很大变化。好的单向散列函数必须具有以下特性：

① 计算的单向性：给定 M 和 f，求 d=f(M)容易，但反过来给定 d 和 f，求 M=f-1(d)在计算上是不可行的

② 弱碰撞自由：给定 M，要寻找另一信息 M’，满足 f(M’)=f(M)在计算上不可行。

③ 强碰撞自由：要寻找不同的信息 M 和 M’，满足 f(M’)=f(M)在计算上不可行。

这里主要是应用 SHA-1 算法的性质 1)，用来生成干扰数据。

因此上面的加密通信的过程从第④步起变为：

④ 用户 B 接到信息后，将待传输的明文编码到 Ep(a,b)上一系列点 Mi，并产生一个随机整数 r(r<n)。

⑤ 用户 B 计算

E0=rP

C1=(M1⊕D1)⊕rQ, D1=SHA-1(rQ)

C2=(M2⊕D2)⊕rQ, D2=SHA-1(D1)

C3=(M3⊕D3)⊕rQ, D3=SHA-1(D2)

……

⑥ 用户 B 将 E0、C1、C2、C3，……传给用户 A。

⑦ 用户 A 接到信息后，计算

C1=C1⊕D1⊕kE0, D1=SHA-1(kE0)

C2=C2⊕D2⊕kE0, D2=SHA-1(D1)

C3=C3⊕D3⊕kE0, D3=SHA-1(D2)

……

对 M1，M2，M3……进行解码，即可得到明文[7]。

3 语音实时加密的程序实现

首先经过 DirectSound 进行声音采集，然后将采集来的数据进行 G729a 压缩算法进行压缩，以减小传输的数据量，双方用于生成的 SHA-1 数据的原始文件数

据为由 ECC 生成的 D0=d*k*p；干扰数据为由 D0 经过不断的 SHA-1 运算生成的数据。然后与声音数据进行合成后经 UDP 协议将数据发送出去。在接收端接到数据后，要先根据事先约定的 ECC 参数计算出 D0 值，然后根据数据的信息头进行相应的 SHA-1 计算，然后将数据进行分解得到压缩后的声音数据，然后进行 G729a 运算还原出声音，通过 DirectSound 将声音播放出来。加密流程如图 2 所示。

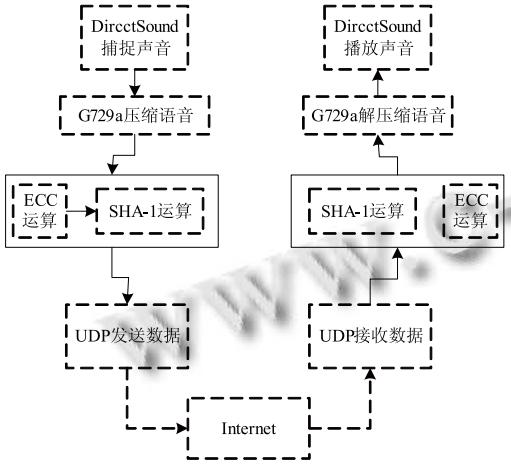

图 2 加密流程

4 结论

本文致力 VoIP 的应用及安全方面的研究，保证通话安全性。深入研究安全散列算法，论述了其实现的过程，并对其进行了 VC 的实现。使用椭圆曲线密码与 SHA 结合，加密语音数据，实现语音的实时加密。

参考文献

1 杨波.现代密码学.北京:清华大学出版社,2007.
2 斯托林斯.孟庆树,王丽娜,傅建明译.密码编码学与网络安全——原理与实践.北京:电子工业出版社,2006.
3 斯皮尔曼,叶阮健,曹英,张长富.经典密码学与现代密码学.北京:清华大学出版社,2005.
4 Stinson DR.密码学原理与实践.北京:电子工业出版社,2003.
5 孙淑玲.应用密码学.北京:清华大学出版社,2004.
6 卢开澄.计算机密码学.北京:清华大学出版社,2003.
7 Sun LH. Study of Authentication Based on Smart Card and Fingerprint Dynamic Password ICICCI 2010.

（上接第 213 页）

报,2009,18(3):249－254.
3 姜明辉,谢行恒,等.个人信用评估的 Logistic-RBF 组合模型.哈尔滨工业大学学报,2007,39(7):1128－1130.
4 刘军丽,陈翔.基于决策树的个人住房贷款信用风险评估模型.计算机工程,2006,32(13):263－265.
5 肖文兵,费奇.基于支持向量机的个人信用评估模型及最优参数选择研究.系统工程理论与实践,2006,26(10):73-79.
6 Ni ZW, Li FG, Yang SL, Liu X, Zhang WL, Luo Q. Attributes reduction based on GA-CFS method. LNCS, 2007(4505):868－875.
7 Narendra PM, Fukunaga K. A branch and bound algorithm for feature subset selection. IEEE Trans. on Computers, 1977, 26(9):917－922.
8 Zhou R, Hansen E. Breadth-First heuristic search. Artificial Intelligence, 2006,170(4-5):385－408.
9 Gheorghies O, Luchian H, Gheorghies A. A study of adaptation and random search in genetic algorithms. Proc. of the 2006 IEEE Congress on Evolutionary Computation (CEC). 2006:2103－2110.
10 Almuallim H, Dietterich TG. Learning boolean concepts in the presence of many irrelevant features. Artificial Intelligence. 1994,69(1-2):279－305.
11 Hall MA. Correlation-based Feature Selection for Discrete and Numeric Class Machine Learning. Proc. of the 17th International Conference on Machine Learning (ICML), 2000:359－366.