E-mail 安全协议 PGP*)

宋玉璞[1] 周爱霞[1] 肖 汉[2]
(商丘职业技术学院计算机系 商丘 476000)[1] (郑州师范高等专科学校现代信息技术部 郑州 450044)[2]

摘 要 本文从单钥密码IDEA 算法、双钥密码RSA 算法、单向杂凑算法 MD5 算法等分析了安全电子邮件协议 PGP 的实现原理和实现流程;描述了 PGP 所提供的安全业务;并从 RSA、IDEA、MD5、随机数等安全性方面分别研究了 PGP 的安全性能。

关键词 安全邮件,PGP,单钥密码,双钥密码,单向杂凑算法

E-mail Security Agreement PGP

SONG Yu-Pu[1] ZHOU Ai-Xia[1] XIAO Han[2]
(School of Computer Science,Shangqiu Vocational & Technical College,Shangqiu 476000)[1]
(Department of Modern Information Technique,Zhengzhou Teacher's College,Zhengzhou 450044)[2]

Abstract This paper, from the algorithm of IDEA, RSA, MD5, analyzes the achievable therory and flow of PGP. The secure business which is provided by PGP is also descibed in the essay. It also researchs on the secure capability of PGP according to many aspects of securty, such as IDEA, RSA, MD5, Randseed and so on.

Keywords Secure E-mail, PGP, IDEA, RSA, MD5

PGP—Pretty Good Privacy 是美国 Phil Zimmermann 研究出来的。他创造性地把 RSA 公匙体系的方便和传统加密体系的高速度结合起来,并且在数字签名和密匙认证管理机制上进行巧妙的设计,从而使 PGP 成为流行的公匙加密软件包。PGP 是一个基于 RSA 公匙加密体系的邮件加密软件,可以用于邮件保密,防止非授权者阅读,还能对邮件加上数字签名,从而使收信人可以确信邮件是谁发来的。它让你可以安全地和你从未见过的人通讯,事先并不需要任何保密的渠道来传递密匙。它采用了审慎的密匙管理,一种 RSA 和传统加密的杂合算法;用于数字签名的邮件文摘算法;加密前压缩等,还有一个良好的人机工程设计。实际上 PGP 的功能还不止用来加密邮件,也可以用来加密文件。

1 PGP 实现的原理

PGP 是一种混合密码系统,包含四个密码单元:单钥密码 IDEA;双钥密码 RSA;单向杂凑算法 MD5;一个随机数生成算法。下面逐一介绍各密码系统。

1.1 单钥密码 IDEA

单钥密码 IDEA(International Data Encryption Algorithm)是 1990 年由瑞士联邦技术学院 X.J.Lai 和 Massey 提出的建议标准算法,称为 PES(Proposed Encryption Standard),1992 年强化了抗差分分析能力后改称为 IDEA。输入和输出字长为 64 位,密钥长 128 位,8 轮迭代体制。该算法是近年来提出的分组算法中很成功的一种,安全、运行速度快、实现简单[1]。

1.2 双钥密码 RSA

RSA(Rivest-Shamir-Adleman)算法是 MIT 三位年轻的数学家 1978 年发现了一种用数论构造双钥的方法后而产生的[2],既可用于加密,又可用于数字签字,安全且易于实现,但运行速度慢。该算法是一种基于大数不可能质因数分解这一假设的公匙体系。简单地说,就是找两个很大的质数,一个公开给世界,一个不告诉任何人。一个称为"公匙",另一个叫"私匙"。这两个密匙是互补的,就是说用公匙加密的密文可以用私匙解密,反过来也一样。假设甲要寄信给乙,他们互相知道对方的公匙。甲就用乙的公匙加密邮件寄出,乙收到后就可以用自己的私匙解密出甲的原文。由于没别人知道乙的私匙,因此即使是甲本人也无法解密那封信,这就解决了信件保密的问题。另一方面,由于每个人都知道乙的公匙,他们都可以给乙发信,那么乙就无法确信是不是甲的来信。认证的问题就出现了,一般用数字签名来解决这个问题,甲用自己的私匙将邮件的 128 位的特征值加密,附加在邮件后,再用乙的公匙将整个邮件加密。这样,这份密文被乙收到以后,乙用自己的私匙将邮件解密,得到甲的原文和签名,同时乙也从原文计算出一个 128 位的特征值来和用甲的公匙解密签名所得到的数比较。如果符合,就说明这份邮件确实是甲寄来的,这样两个安全性要求都得到了满足。

1.3 单向杂凑算法 MD5

MD5 算法是 Ron Rivest 1992 年提出的一个单向杂凑函数。所谓杂凑,就是将任意长的数字串 M 映射成一个定长输出数字串 H 的过程。所谓单向,就是任意两个 M 不可能具有相同的 H。由于这两个特征,该函数是实现有效、安全、可靠数字签字和认证的重要工具。PGP 中用到 MD5 算法来获取 128 位消息文摘(message digest),用于对该邮件的签名和认证[3]。

1.4 一个随机数生成算法

随机数的生成是指 PGP 提供两个伪随机数发生器

*)国家高技术研究发展计划"863"项目资助(2001AA132011)。**宋玉璞** 讲师,硕士,研究方向为软件工程;**周爱霞** 讲师,硕士,研究方向为软件工程;**肖 汉** 副教授,硕士,研究方向为软件工程及计算机应用技术。

(PRNG)：一个是 ANSI X9.17 发生器，采用 IDEA 算法，以 CFB 生成；另一个是从用户击键的时间和序列中计算熵值，从而引入随机性。

2 PGP 提供的安全业务

PGP 在安全上的业务有：认证、加密、压缩、同 E-mail 的兼容性、基数-64 变换。

2.1 PGP 的认证性

PGP 还可以只签名而不加密，这适用于公开发表声明时，声明人为了证实自己的身份(在网络上只能如此了)，可以用自己的私匙签名。这样就可以让收件人能确认发信人的身份，也可以防止发信人抵赖自己的声明。认证协议如下：

送信人编制消息 M；用 MD5 产生一个 128 位的杂凑值 H；用送信人的 RSA 密钥对 H 签字；将 $M||H$ 经压缩 Z 后送出；收端对收到的数据进行 Z^{-1} 变换，并以发送人公钥解出 H；用接收的 M 计算杂凑值得 H，与 H 值进行比较签字，如图 1 所示。其中 KR_a：用户 A 的秘钥，KU_a：用户 A 的公钥，Z：压缩，ER：RSA 加密，DR：RSA 解密，Z^{-1}：解压缩。

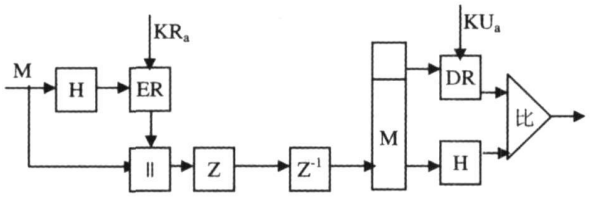

图 1 PGA 认证流程图

2.2 PGP 的机密性

发送者产生消息和 128bit 会话密钥。以密钥对压缩的 M 按 IDEA 体制加密。以接受者公钥按 RSA 体制对密钥加密，接于 M 之后。接受者以 RSA 密钥解密，获得会话密钥。接收者以会话密钥按 IDEA 体制解密，并解压缩，获得原文 M。PGP 的加密流程图如图 2 所示。其中，IE：IDEA 加密，DI_a：IDEA 解密，K：压缩。

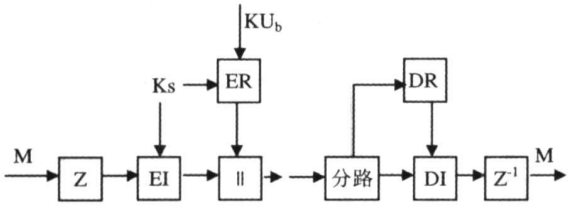

图 2 PGA 加解密流程图

IDEA/RSA 组合方式使消息 M 加密时间大大缩短，且这种一次性会话密钥方式特别适合于存储转递 E-mail 业务，避免了执行交换密钥的握手协议，提高了安全性。RSA 的字长有三种选择：①临时性：384bit，经过努力可以破解；②商用：512bit，可以由专业组织破译。③军用：1024 位，一般相信不可破译。PGP 可同时提供机密性与认证性。它采用先签字后加密的方式，这样的好处是：存储对消息明文的签字较为方便。第三者证实时，无须知道通信者所用的 IDEA 的会话密钥 K。

2.3 PGP 的压缩

压缩可以节省通信时间和存储空间，而且在加密前进行压缩可以降低明文的冗余度，增强机密效果。PGP 中采用 ZIP 算法进行压缩。

2.4 PGP 同 E-mail 的兼容性

PGP 在上述三种业务下，输出消息均有加密数据，故 PGP 输出中的一部分或全体为任意的 8bit 串。而许多 E-mail 系统只允许使用 ASCII 文本。为此 PGP 采用基数－64 变换，将 8bit 字符串变为可以打印的 ASCII 字符串。

2.5 基数-64 变换

基数－64 变换是将任意二元输入变换成可以打印的字符输出。它具有以下特点：不限于某一特定的字符集编码；字符集有 65 个可以打印的字符，每个字符表示 6bit 输入，一个字符作为填充；字符集中不含控制字符，故可通行于 E-mail 系统；不用'－'连字符，此符号在 RFC832 中有特定意义，应避免使用[4]。

2.6 PGP 消息分段和重组

E-mail 对消息长度都有限制。当消息大于长度限度时，PGP 将对其自动分段。分段是在所有处理之后进行的，故会话密钥和签字只在第一段开始部分出现。在接收端，PGP 将各段自动重组为原来的消息[5]。

3 PGP 的安全性

PGP 是一种混合密码系统，其安全就是它的四个加密部分：单钥密码 IDEA；双钥密码 RSA；单向杂凑算法 MD5；随机数生成算法的安全性。

3.1 IDEA 的安全性

在 PGP 中采用 IDEA 的 64-bits CFB 模式。IDEA 比同时代的算法，像 FEAL，REDOC-II，LOKI，Snefru 和 Khafre 都要坚固。而且最近的证据表明，即使是在 DES 上取得巨大成功的 Biham 和 Shamir 的微分密码分析法，对 IDEA 也无能为力。Biham 和 Shamir 曾对 IDEA 的弱点作过专门分析，但他们没有成功。对 IDEA 的攻击方法就“直接攻击”或者说是“密匙穷举”一种了[6]。

3.2 RSA 的安全性

我们知道 RSA 的保密性基于一个数学假设：对一个很大的合数进行质因数分解是不可能的。下面是几种因数分解的算法。试探除法：最老也是最笨的方法；二次筛法(QS)：对 $10\wedge110$ 以内的数是最快的算法；MPQS：QS 的改进版本；分区筛法(NFS)：目前对大于 $10\wedge110$ 的数是最快的算法，曾被用来成功地分解过第九费马数[7]。下面是几种针对 RSA 有效的攻击方法，它们实际上对 PGP 没有效力，因为它们攻击的是加密协议环节上的漏洞，而不是 RSA 本身。从这些例子可以看到 PGP 是如何在实现上堵住这些漏洞的。

• 选择密文攻击：攻击者截取到密文后，将密文和一随机信息合并，传给邮件发送者，请他签名，在签名过程中也就是对原来密文的解密过程。传回给攻击者后，即可破译密文。

• 过小的加密指数 e：看起来，e 是一个小数，并不降低 RSA 的安全性。从计算速度考虑，e 越小越好。可是，当明文也是一个很小的数时就会出现问题。

• RSA 的计时攻击法：这是一种另辟蹊径的方法。至于 PGP，根本不用担心计时攻击，因为 PGP 采用了中国余数理论的方法加速了运算，同时使耗时与操作数无关。

3.3 MD5 的安全性问题

MD5 是一种在 PGP 中被用来单向变换用户口令和对信息签名的单向散列算法。一种单向散列的强度体现在它能把任意的输入随机化到什么程度，并且能产生唯一的输出。对单向散列的直接攻击可以分为普通直接攻击和“生日”攻击。

3.4 随机数的安全性问题

PGP 使用两个伪随机数发生器(PRNG):一个是 ANSI X9.17 发生器,另一个是从用户击键的时间和序列中计算熵值从而引入随机性。

• 用户击键引入随机性:这是真正的随机数,只是尽量使击键无规则就行。

• ANSI X9.17 PRNG:使用 IDEA 而不是 3DES 来产生随机数种子。X9.17 需要 randseed·bin 中的 24 bytes 的随机数,PGP 把其他 384 bytes 用来存放其他信息。Randseed·bin 文件最初是利用用户击键信息产生的,每次加密前后都会引入新的随机数,而且随机数种子本身也是加密存放的。

• X9.17 用 MD5 进行预洗:所谓"洗"就是指像洗牌一样把数据打乱。加密前叫预洗,加密后为下一次加密的准备,叫后洗。PGP 的日常随机数产生器 X19.7 是用明文的 MD5 值来预洗的,它基于攻击者不知道明文这样一个假设。

• randseed·bin 的后洗操作:后洗操作被认为是更安全的。更多的随机字节被用来重新初始化 randseed·bin 文件,它们被用当前的随机临时 PGP 密匙来加密。同样,如果攻击者知道这个密匙,他就不用攻击 randseed·bin 文件。相反,他更关心 randseed·bin 文件当前的状态,因为可能从中获得下次加密的部分信息。因此,对 randseed·bin 文件的保护和公匙环及私匙环文件同样重要。

参 考 文 献

1 王育民,刘建伟,等.通信网的安全—理论和技术[M].西安电子科技大学出版社,2002

2 郑丽娟,刘莉,等.邮件加密软件 PGP 的安全技术研究与应用[J].河北省科学院学报,2005(4)

3 刘雅丽.PGP 保护电子邮件的研究[J].孝感学院学报,2005(3)

4 杨宗德,等.基于 PGP 的安全电子邮件系统设计与实现[J].信息安全与通信保密,2005(9)

5 魏洪波,周建国,等.安全电子邮件协议[J].现代电信科技,2002(3)

6 陈勇.安全电子邮件系统的设计与分析[J].舰船电子工程,2006(4)

7 郑丽娟,郑丽伟,等.邮件加密算法 PGP 的改进[J].河北大学学报(自然科学版),2004(3)

(上接第 30 页)

时间时平均时延随缓冲区长度的变化曲线,可见分形开始时间越小,网络的性能越差。

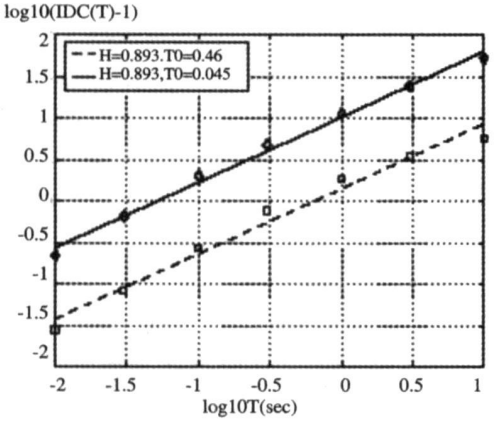

图 5 网络流量的 IDC 图

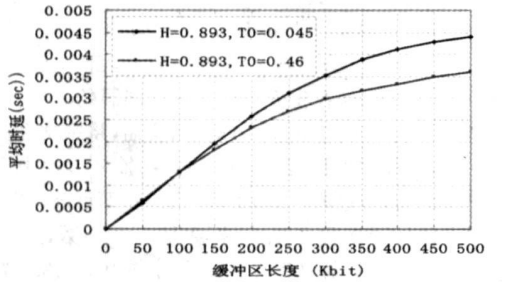

图 6 相同 H 不同 T_0 时的时延

结束语 目前的自相似业务建模和性能评价方法大都片面强调流量过程统计特性的某些方面的影响,而忽略其它因素。而事实上,虽然 Hurst 系数对网络性能有重要影响,但单靠 Hurst 系数并不能全面反映网络的性能,同时 Hurst 系数本身对网络性能的影响还受其它因素(如流量的方差、缓冲区长度、利用率等)的制约。同样,在自相似业务识别与参数估计方面,目前的研究基本上全部集中在 Hurst 系数的估计上,而不考虑其它参数。

本文深入研究影响网络性能的自相似流量关键特性,通过 MATLAB 和 OPNET 相结合的仿真方法研究 Hurst 系数和方差系数对网络性能的影响,结果发现 Hurst 系数和方差系数对网络性能均有重要的影响。然后基于长相关的定义,分析方差对网络性能影响的原因,研究 c_γ 与方差之间的关系及其计算方法。最后给出了基于 IDC 的复合分形更新过程参数的估计算法,分析了分形开始时间对网络性能的影响。

参 考 文 献

1 Leland W E, Willinger W, Taqqu M S, et al. On the self-similar nature of Ethernet traffic (extended version) [J]. IEEE/ACM Trans. Networking, 1994, 2(1): 1～15

2 Park K, Willinger W. Self-similar Network Traffic and Performance Evaluation[M]. New York: Wiley Interscience, 2000

3 Norros I. On the use of fractional brownian motion in the theory of connectionless networks [J]. IEEE J Select Areas Commun, 1995, 13(6): 953～962

4 Norros I. A storage model with self-similar input [J]. Queuing Syst, 1994, 16: 387～396

5 Brichet F, Roberts J, Simonian A, et al. Heavy traffic analysis of a storage model with long-range dependent on/off sources [J]. Queuing System. their Applications, 1996, 23: 197～215

6 Yoshihara T, Kasahara S, Takahashi Y. Practical Time-scale Fitting of Self-similar Traffic with markov-Modulated Poisson Process [J]. Telecomm. Systems, 2001, 17(1-2): 185～211

7 Erramilli A, Narayan O, Neidhardt A, et al. Performance impacts of multi-scaling in wide m a tcp/ip traffic [J]. In: INFOCOM 2000, Tel Aviv, Israel, 2000, 1: 352～359

8 Ryu B K, Lowen S B. Point Process Approaches to the Modeling and Analysis of Self-similar Traffic — Part I: Model Construction [A]. In: Proc IEEE INFOCOM' 96 [C]. San Francisco, March 1996. 1468～1475

9 Beran J. Statistics for Long-memory Processes. London, U K: Chapman & Hall, 1994

10 Watagodakumbura C, Jennings A, Shenoy N. Absolute effects of aggregation of self-similar traffic on quality of service parameters [J]. In: First International Symposium on Control, Communications and Signal Processing, 2004. 511～514

11 Lowen S B, Teich M C. Fractal renewal processes generate 1/f noise [J]. Physics Review E, 1993, 47: 992～1001

12 Lawrence Berkeley National Laboratory. BC-Ethernet traces of LAN and WAN traffic [DB/OL]. http://ita.ee.lbl.gov/html/contrib/BC.htm1, 2003-06