SPA:新的高效安全协议分析系统

李建欣 李先贤 卓继亮 怀进鹏

(北京航空航天大学计算机学院 北京 100083)

摘 要 研制高效的自动分析系统是密码协议安全性分析的一项关键任务,然而由于密码协议的分析非常复杂,存在大量未解决的问题,使得很多现有分析系统在可靠性和效率方面仍存在许多局限性.该文基于一种新提出的密码协议代数模型和安全性分析技术,设计并实现了一个高效的安全协议安全性自动分析系统(Security Protocol Analyzer,SPA).首先对协议安全目标进行规范,然后从初始状态出发,采用有效的搜索算法进行分析证明,试图发现针对协议的安全漏洞.使用该系统分析了 10 多个密码协议的安全性,发现了一个未见公开的密码协议攻击实例.实验数据显示,该系统与现有分析工具相比,具有较高的分析可靠性和效率,可作为网络系统安全性评测以及密码协议设计的有效辅助工具.

关键词 信息安全;密码协议;形式化分析;搜索算法;攻击序列

中图法分类号 TP309

SPA: A New Efficient System for Security Protocol Analysis

LI Jian-Xin LI Xian-Xian ZHUO Ji-Liang HUAI Jin-Peng

(*School of Computer Science, Beihang University, Beijing* 100083)

Abstract Cryptographic protocols are communication protocols of distributed systems that use cryptography to achieve various goals such as authentication and key distribution in the open network environment. Developing efficient automatic system is a crucial task in the area of security analysis for cryptographic protocols. However, the security analysis for cryptographic protocols is very complex and difficult and a lot of unresolved problems still present in it, which causes some limitations on reliability and efficiency of previously available automatic systems. An efficient automatic analysis system (Security Protocol Analyzer, SPA) for cryptographic protocols is designed and implemented based on the CPA (Cryptographic Protocol Algebra) model and new analysis techniques. In this system, security properties of the cryptographic protocols are specified firstly, and then an effective proof search algorithm starting with the initial state is applied and tries to find all possible attack sequences on them. More than ten cryptographic protocols are analyzed successfully with this system, and moreover, a new attack instance is also found. The analysis results show that SPA has improved the analysis reliability and efficiency compared with other existing tools. It can be used as an efficient tool for the evaluation of the network security and the design of cryptographic protocols.

Keywords information security; cryptographic protocol; formal analysis; search algorithm; attack sequence

收稿日期:2003-01-09;修改稿收到日期:2004-12-30.本课题得到国家自然科学基金(90412011)和国家"八六三"高技术研究发展计划项目基金(2003AA144150)资助.**李建欣**,男,1979 年生,博士研究生,主要研究方向为网络安全、信任协商.E-mail:lijx@act.buaa.edu.cn;myjianxin@sina.com.**李先贤**,男,1969 年生,博士,副教授,主要研究方向为信息安全、计算机科学理论.**卓继亮**,男,1976 年生,硕士研究生,主要研究方向为网络安全.**怀进鹏**,男,1962 年生,博士,教授,博士生导师,主要研究方向为计算机软件与理论、中间件与网格技术、网络安全.

1 引 言

随着互联网技术及应用的飞速发展,如何保护网络传送信息的安全成为一项重要的任务.密码协议(也称为安全协议)是采用密码学技术来达到密钥分配、主体身份认证等目的的通信协议,在很大程度上为网络通信提供了安全性保证.这些密码协议都是经过精心设计的,角色间消息的交互存在着复杂的关系和制约,而且大部分的密码协议是运行在分布式网络环境中.分布式网络具有多主体参与、大规模并发和运行动态性等特点,正是由于这种运行环境的复杂性使得设计出的密码协议难免出现缺陷.大量的事例表明,许多密码协议在使用多年后才发现存在很严重的安全漏洞,例如著名的 Needham-Schroeder 公开密钥协议在公开 17 年后,它存在的中间人攻击漏洞才被 Lowe 发现[1].

由于密码协议的漏洞一般都很隐蔽,使用手工分析非常困难,易于出错,所以研制高效、可靠的自动分析工具成为密码协议安全性分析的一项关键任务.目前国际上推出了许多自动分析工具,然而它们大都仍然存在搜索状态数太多的问题,而且这些工具往往只面向研究人员,一般的技术人员难以使用.本文在文献[2]提出的密码协议形式模型——CPA(Cryptographic Protocol Algebra)模型的基础上,设计并实现了一个高效的安全协议分析系统 SPA(Security Protocol Analyzer),能够有效分析多种类型的密码协议.

2 相关工作

采用形式化方法分析密码协议已有 20 多年的历史,一直属于国际上研究的热点[3].目前,对密码协议进行形式化分析广泛采用模型检测和定理证明的技术.

针对密码协议进行形式化分析的特定模型检测工具包括 Interrogator[4]、Brutus[5]等.其中 Interrogator 是比较早的基于状态空间搜索思想来发现协议安全缺陷的模型检测工具,它使用通信状态机描述协议的参与者,并赋予了攻击者对信息自动处理的能力;Brutus 是采用一阶逻辑描述密码协议的属性,从协议运行的初始状态开始搜索所有合法主体和攻击模型的活动序列来检测是否存在协议的攻击序列,并使用偏序和对称的消减方式来限制状态搜索空间.通用模型检测工具包括 FDR[1]、Murφ[6]等,其中 FDR 是基于进程代数理论 CSP 的自动模型检测工具,牛津大学的 Lowe 首先将其用于密码协议安全性的分析,方法是将协议参与主体视为 CSP 进程,攻击者的进程则描述多种攻击行为,协议的安全性目标被描述为主体的事件序列;Murφ 是由斯坦福大学的 Mitchell 应用于密码协议的分析,它在许多方面的描述类似于 FDR,但它是使用共享变量来描述协议主体事件,使用不变量描述协议的安全属性,同时采用对称消减、可逆规则等技术来提高系统效率.

模型检测工具更有利于通过自动推理生成攻击实例的方法来检测协议的不正确性,但在证明协议正确性时,往往很难做到将无限状态空间映射为一个有限状态空间,从而容易出现无限搜索状态空间的问题.相对于模型检测,定理证明的技术可以处理无限状态空间,因而可以提供协议的正确性证明.采用定理证明技术的 NRL 协议分析器[7]是比较著名的密码协议分析工具,能够分析任意数量参与主体密码协议的安全性,它是以不安全的状态为开始进行回溯搜索,如果不安全状态不能够从初始状态达到,协议就是安全的,否则就是不安全的,同时搜索的过程就是一个攻击的过程.NRL 分析器的缺点是自动化程度较低,往往需要人工干预,相对效率也比较低.

Athena[8,9]是 Dawn Song 和 Sergey Berezin 等结合模型检测和定理证明的方法新开发的密码协议分析工具,它使用扩展的 SSM 模型[8,9]描述协议的运行,采用更为紧致的状态表示方法、变量替换和消减规则来减小状态的搜索空间.Athena 的证明过程类似 NRL 分析器,都是以不安全状态为开始进行状态搜索的,但是它对攻击者的能力做了约束,且变量替换没有类型匹配的限制,在处理复杂密码协议时,容易产生搜索状态空间爆炸问题.

3 密码协议的代数模型

本文采用文献[2]提出的密码协议代数(CPA)模型规范密码协议,并运用了相应的安全性分析理论.为便于理解,这里对 CPA 模型只作简单介绍,关于本节内容的详细论述可参见文献[2],本文未特别说明的符号和术语均与文献[2]一致.

3.1 密码协议代数

原子消息. $PRIM=ID\cup KEY\cup GEN$.其中 ID 表示协议中主体的名称集合;KEY 表示密钥集,

包括非对称密钥中的公私钥和对称密钥；$GEN=\bigcup_{a\in ID}G_a$，G_a对应协议中由标识名为 a 的主体生成的所有消息集.

消息项可如下递归定义：

(i) 若 t 是原子消息，则 t 是消息项；

(ii) 若 t_1，t_2是消息项，则 $t_1\cdot t_2$是消息项；

(iii) 若 t 是消息项，$k\in KEY$，则$\{t\}_k$，$MAC(t,k)$和 $Hash(t)$是消息项.

其中运算符 $\cdot$ 称为联接运算；$\{m\}_k$称为加密运算，满足：

$$\{\{m\}_k\}_{k^{-1}}=m \text{ 和} \{\{m\}_{k^{-1}}\}_k=m,$$

若 $k=k^{-1}$则称为对称密码运算；若 $k\neq k^{-1}$则称为公钥密码运算；

$Hash(m)$称为哈希函数运算，$MAC(m,k)$称为消息认证码函数，满足

$$Hash(m_1)=Hash(m_2)\Leftrightarrow m_1=m_2$$

和 $MAC(m_1,k_1)=MAC(m_2,k_2)\Leftrightarrow m_1=m_2\wedge k_1=k_2$.

消息项集 M 与消息运算构成一个代数系统，称为密码协议代数 CPA.

迁移函数. 设 A 是一个 CPA 代数，σ 是 $M^n\cup\{\Delta\}$到 $M\cup\{\Delta\}$的一个映射，若对任何 $\boldsymbol{b}=(b_1,b_2,\cdots,b_n)\in M^n$，满足 $\sigma(\boldsymbol{b})\in A[b_1,b_2,\cdots,b_n]$，则称 σ 为 A 上的一个迁移函数. 这里引入一个特殊符号“Δ”，表示空的消息项.

变量替换. 一个变量替换 $\phi(t)=t\{t_1/x_1,t_2/x_2,\cdots,t_n/x_n\}$表示使用项 t_i替换变量 x_i在 t 中的所有出现项. 若变量替换满足 $\phi(t)=\phi(t')$，则称 ϕ 是项 t 和 t'的合一.

3.2　密码协议的形式描述

以下叙述中，均假设 P 是一个 N 方密码协议.

基本事件. 基本事件可表示为三元组$\langle\sigma_{ij}(\alpha,l),t,t'\rangle$，其中 $\sigma_{ij}(\alpha,l)$为事件函数，l 为协议的会话轮数，$\alpha=(a_1,a_2,\cdots,a_N)$ 为参与协议会话主体集，下标 i 用于标识该事件主体在 α 中的位置，下标 j 用于标识该事件所在事件序列的序号. t 和 t'分别表示接收到和发送出的消息项，满足映射关系 $t'=\sigma_{ij}(\alpha,l)(t)$.

事件函数偏序关系 $\preceq_P$，满足 $\sigma_{i_1j_1}(\alpha_1,l_1)\preceq_P\sigma_{i_2j_2}(\alpha_2,l_2)$当且仅当 $i_1=i_2\wedge\alpha_1=\alpha_2\wedge(l_1<l_2\vee(l_1=l_2\wedge j_1\leq j_2))$. 设协议 P 的有限序列 $\Gamma=\tau_1\tau_2\cdots\tau_n$，其中 τ_i表示事件函数，若满足：

(1) 对于任何 $1\leq i<j\leq n$，不存在 τ_i，τ_j 使得 $\tau_j\preceq_P\tau_i$；

(2) 当 $\sigma_{ij}(\alpha,l)\in\Gamma$，对于 $1\leq k\leq j$，$\sigma_{ik}(\alpha,l)\in\Gamma$；

则称 Γ 是协议 P 的长度为 n 的一条迹. 当 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)\in M^n$，则称序列 $\Gamma(\boldsymbol{x})$为协议 P 的迹序列.

密码协议的形式定义. 一个 $N(\geq 2)$方密码协议是二元组，其中，

(1) M 是基本消息代数空间；

(2) $P=\langle P_1,P_2,\cdots,P_N\rangle$，其中 P_i是 N 方协议 P 的第 i 个角色，对 $\forall\alpha=(a_1,a_2,\cdots,a_N)\in ID^{(N)}$，$l\in N$，可以导出相应事件函数序列：

$$\Sigma_i(\alpha,l)=\sigma_{i,1}(\alpha,l)\cdots\sigma_{i,s_i}(\alpha,l),$$

其中 $\sigma_{i,j}(\alpha,l)$为协议的原子事件函数. 在这里，迹 $\Sigma_i(\alpha,l)$称为角色 P_i 的正则迹. 当 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)\in M^n$，则称迹序列 $\Sigma_i(\alpha,l)(\boldsymbol{x})$为角色 P_i 的正则迹序列.

正合序列. 设 $\Gamma=\sigma_1\sigma_2\cdots\sigma_n$是迁移函数序列，对于 $a_i\in M\cup\{\Delta\}$，可得到消息迁移序列 $\Gamma(\boldsymbol{a})=\langle\sigma_1,a_1,\sigma_1(a_1)\rangle\langle\sigma_2,a_2,\sigma_2(a_2)\rangle\cdots\langle\sigma_n,a_n,\sigma_n(a_n)\rangle$，$D_0$是攻击者初始知识，若满足 $a_1\in D_0$，$a_i\in D_{i-1}$，$D_i=D_{i-1}[\sigma_i(a_i)]$，$i=1,2,\cdots,n$，则称 $\Gamma(\boldsymbol{a})$是一个正合的 D_0序列.

正合序列描述了攻击者的一个攻击过程.

3.3　密码协议形式分析模型

3.3.1　语法

项包括消息常量$\{a,b,\cdots\}$，迹序列常量$\{\Gamma,\Sigma,\Lambda,\cdots\}$和迹序列变量$\{X,Y,\cdots\}$等. 谓词符号包括$\sqsubseteq$，EXT，$\in$.

(1) 原子命题公式：$\Sigma\sqsubseteq X$，$\mathrm{EXT}(\Sigma,X)$，$(a,\Sigma)\in X$(简记作 $a\in_\Sigma X$).

(2) 若 f_1和 f_2是命题公式，则$\neg f_1$，$f_1\wedge f_2$和 $\forall X.f_1$也是命题公式.

这里 a 是消息常量，Σ 是迹序列常量，X 是迹序列变量，f_1最多含自由变量 X. 可以自然方式引入逻辑符号：$\Rightarrow$，$\vee$和量词$\exists$.

3.3.2　语义

协议 P 的一个模型是五元组 $M_P=(M,\Sigma_P^*,R_P,E_P,\Phi)$，其中 M 是基本消息项集合，Σ_P^* 是迹序列集合，R_P是 Σ_P^* 中的正则迹序列集合，$E_P=\bigcup_{\Sigma\in R_P}Ext(\Sigma)$，其中 $Ext(\Sigma)$表示所有正合的 Σ-序列集合，Φ 是对常量和变量的解释. 语义解释如下：

(i) $\Phi(a)\in M$，$\Phi(\Sigma)\in\Sigma_P^*$，$\Phi(X)\in\Sigma_P^*$ 分别是消息常量、迹序列常量和迹序列变量的解释.

(ii) $M_P\models\Sigma\sqsubseteq X$ 当且仅当$\Phi(\Sigma)\subseteq\Phi(X)$.

(iii) $M_P\models EXT(\Sigma,X)$ 当且仅当 $\Phi(\Sigma)\in R_P$且

$\Phi(X)\in Ext(\Phi(\Sigma))$,即 X 是正合的 Σ-序列.

(iv) $M_P\models a\in_\Sigma X$ 当且仅当 $\Phi(\Sigma)\in R_P$ 且 $\Phi(a)\in_{\Phi(\Sigma)}\Phi(X)$ ($a\in_\Sigma X$ 为 $(a,\Sigma)\in X$ 的简记,表示 $a\in D_0[X]$, 其中 D_0 是关于正则迹 Σ 的攻击者初始知识).

对含逻辑联接词 $\neg$, $\wedge$ 和量词 $\forall$ 的公式语义解释是自然的.

3.3.3 密码协议安全性描述

秘密性安全目标. 协议的秘密安全性是建立共享秘密协议的安全要求,目的是保护某些数据(即会话目标消息)不被攻击者获取,表示为公式

$$\forall X.(EXT(\Sigma,X)\wedge\Sigma\sqsubseteq X)\Rightarrow\neg t\in_\Sigma X$$

(简略形式为 Σ Secret t),

其中 Σ 是正则迹序列,t 是秘密消息项.

对应性安全目标. 认证性是鉴别传输数据来源的真实性或新鲜性. CPA 模型将数据的源认证解释为一种对应性,即当认证方的正则迹序列属于一个正合迹序列,协议预先规定的角色正则迹序列(或部分)也属于这个正合序列. 表示为公式

$$\forall X.(EXT(\Sigma,X)\wedge\Sigma\subseteq X)\Rightarrow\Lambda\subseteq X$$

(简略形式为 Σ Corresp Λ),

其中 Λ 是某个正则迹 $\Sigma_j(\alpha,l)$ 的部分或全部(简称子迹).

密码协议的安全性定义. 设 M_P 是协议 P 的一个模型,G_P 是安全目标集,若 G_P 中的每个公式在模型 M_P 中都为真,则称 P 是一个安全的密码协议.

4 新的安全协议分析系统(SPA)

基于文献[2]中的密码协议代数(CPA)模型,我们设计并实现了一个高效的密码协议自动分析系统(Security Protocol Analyzer, SPA),并应用该系统成功分析了一些密码协议的安全性. 本节首先简要介绍系统的总体结构,然后介绍系统中的核心算法与技术.

4.1 SPA 系统的总体结构

SPA 系统的整体结构框架如图 1 所示,下面简要介绍各个模块的主要功能.

(1) 协议输入及其解析模块

目前的公开密码协议大多使用半自然语言描述,这种描述不太严谨,容易产生二义性. 由于 SPA 系统仅能处理规范好的密码协议,我们采用 EBNF (Extended Backus-Naur Form)定义了一种与 CPA 模型描述语言配套的协议规范描述语言,能够准确描述密码协议组成要素.

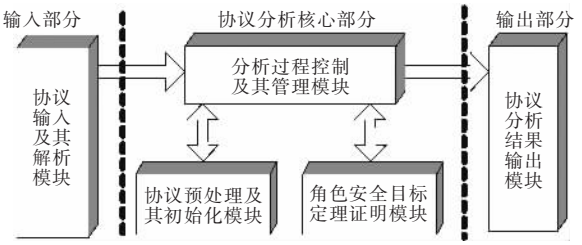

图 1 SPA 系统结构框架

该模块提供了协议规范文本编辑器和协议辅助生成向导两种方式生成协议规范描述文本,以满足不同用户的需求. 通过对生成的协议规范描述文本进行词法和语法解析,如果符合协议描述规范语法,则执行相应的语义动作,生成描述协议的静态数据.

(2)分析过程控制及其管理模块

这部分的主要功能是控制整个协议的分析过程,根据用户指令启动或停止分析进程. 在实际运行过程中,还需要对数据的动态交换、存储进行管理,协调各个具体功能模块,保证以最优效率完成协议的分析.

(3)协议数据预处理及其初始化模块

该模块主要负责协议的预处理和初始化.

首先,由于在协议会话过程中引入攻击者角色,所以需要构造扮演攻击者角色主体参与的会话组,生成各会话主体新的迹序列.

同时,由于密码协议大多运行在开放的网络环境中,网络数据都是公开的,所以需要确定攻击者的初始知识集合,并根据其运算能力作相应的约简操作.

最后,完成协议分析的初始化工作,包括为各个角色的安全目标定理证明分配独立的运行空间,建立对应的初始状态和证明树根节点,加载可能存在的用户预设策略(用户可对分析过程的证明树的深度、宽度以及状态迹序列长度等条件进行约束).

(4)角色安全目标定理证明模块

在 CPA 模型中,密码协议 P 的安全性被规范为安全目标定理集 G_P,这里关键的问题就是通过一种有效的算法来证明 G_P 中的每条定理.

本模块主要应用状态搜索算法 *SearchAlg* 在有限的状态空间中对角色的安全目标定理进行严格的推理证明. 在状态搜索过程中,需要判定各状态的安全性,其中攻击者不可达状态、攻击者可达但不违背安全目标状态都属于安全状态,攻击者可达而且违背安全目标状态属于不安全状态,对不能判定其安全性的状态,应用分裂算法 *Split* 扩张迹序列并生成新的状态集(需要保证新状态的有效性和新鲜性).

(5)协议分析结果输出模块

为了便于分析人员了解协议分析的过程及结果，该模块既可显示协议分析结果的信息摘要(如搜索时间、搜索状态数目、攻击序列数目等)，也可以通过 Tree 结构图直观显示每个角色的证明状态树. 这样，对于存在安全漏洞的协议，协议分析人员就能够直观、方便地了解它的攻击过程.

SPA 系统的各模块间是一种松藕合关系，其总体结构框架具有良好的可扩展性，能够方便插入针对特殊协议的附加处理模块.

4.2　SPA 系统的核心算法与技术

我们使用 Java 语言实现了 SPA 系统，下面分别介绍系统三个组成部分中的一些关键算法及技术.

4.2.1　输入部分

在 SPA 系统的输入部分，使用递归下降子程序解析协议规范描述文本，如果解析通过，将分离出它的三个主体描述类(如图 2 所示)，其中特定迹序列是附带参数的迹序列，用于辅助表示安全目标定理中的正则迹序列 Σ 或其子迹 Λ. 对于角色 P_i 的迹序列 $\Sigma_i(\alpha,l)$ 中的消息项，按照其递归定义形式表示为消息树结构，其中叶节点为原子消息项. 根据这些原子消息项来源和属性(在协议规范描述文本中，各角色中的原子消息属性由对应的参数表指定)的不同，采用唯一标识索引，存入 Hash 表中统一管理，以提高核心算法频繁访问这些信息的效率.

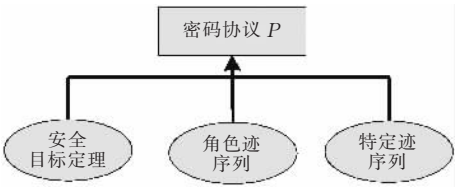

图 2　协议 P 的三个主体描述类

4.2.2　协议分析核心部分

在密码协议分析核心部分，首先需要完成协议的预处理和初始化，然后启用状态搜索算法从初始状态出发对协议描述的各个角色安全目标定理证明，在状态搜索过程中所采用的分裂算法则是 SPA 系统能够快速构造出攻击序列的一个关键部分.

协议的预处理过程.

1. 在确定的会话轮数 l 下，对具有 s 个可信服务器角色(该角色不可由非诚实主体扮演，如证书权威机构 CA 等)的 n 方密码协议 P，生成 $(n-s)!$ 种除可信角色外诚实主体的会话组合，在具有攻击者协同参与情形下，引入 $p=n-s-1$ 个可扮演各个诚实主体角色的攻击者 $(p_1,p_2,\cdots,p_{n-s-1})$，生成 $(n-s)$ 种主体会话组合，最后对 $(n-s)!+(n-s)$ 种组合进行对称约简. 例如对于双方密码协议 $P=\langle a,b\rangle$ 且 $l=1$，会话组合为 $(a,b),(b,a),(a,p)$ 和 (p,b).

2. 对步 1 中的每种主体会话组合，主要应用变量替换等算法，生成各主体新的会话迹序列. 进而求得所有诚实主体的规范序列空间 $Nor_{\Sigma=}\bigcup(Nor\Sigma_i(\alpha,l))$.

3. 设置攻击者本身知识集合 D_0，D_0 包括所有诚实主体 ID 及其公钥，攻击者的密钥 KEY(私钥和共享密钥)及其生成的消息 GEN，也可是协议规范预设的其它公开信息.

4. 构造攻击者可以通过公开网络获取协议会话的知识集合 $D_P=\{t\mid\langle\sigma,s,t\rangle\in Nor_\Sigma\}$.

5. 由步 3 和 4 生成攻击者的初始知识集合 $D=D_0\cup D_p$. $\forall d\in D$，若 $d\in D_0$，则记录 d 的起源 $d\rightarrow D_0$；若 $d\in D_p$，则根据攻击者知识集 D，对 d 不断进行解密、分解等约简操作，直到 D 集合中不再存在可约简的消息子项，对于 d 的不可约简项 $d_1,d_2,\cdots,d_k$，分别记录 d_i 的起源节点 $d_i\rightarrow\langle\sigma_s,t_s,t_s'\rangle$.

协议的初始化过程.

在 SPA 中，采用三元组 (A,B,E) 表示协议的运行迹状态(简称状态)，其中 A 为非空迹序列，B 为安全目标节点集，E 为正合节点集. 针对 CPA 模型描述的两类安全目标，分别采用如下方法来构造它的初始状态：

(1) 对于秘密性安全目标 Σ Secret t，构建一个新节点 $\langle\sigma,t,\Delta\rangle$，相应的初始状态为 $s_0=(\Sigma\bigcup\langle\sigma,t,\Delta\rangle,\varnothing,\varnothing)$.

(2) 对于对应性安全目标 Σ Corresp Λ，相应的初始状态为 $s_0=(\Sigma,\Lambda,\varnothing)$.

建立初始状态后，SPA 采用宽度优先的原则进行状态搜索，主要包括状态搜索算法 $SearchAlg$ 和状态分裂算法 $Split$.

算法 1.　状态搜索算法.

```
SearchAlg(){
//建立初始状态
StateSet={(A0,B0,E0)};
 while(StateSet!=∅){
    //选取一个待分析状态
    (Ai,Bi,Ei)=Select(StateSet);
    //用户预设处理策略处理该状态
    UserPolicy((Ai,Bi,Ei));
    //判断迹序列 Ai 是否包含安全目标节点
    if (Bi⊆ Ai){
      //状态(Ai,Bi,Ei)为安全状态
      StateSet.Remove(Ai,Bi,Ei);
      continue; }
    //判断迹序列 Ai 中节点是否已全部正合
    if ((Bi⊄Ai)&&(Ai⊆Ei))
      //(Ai,Bi,Ei)为不安全状态,
      //正合序列 Ai 描述了协议的一个攻击
```

```
        return FALSE;
      //应用状态分裂算法,得到子状态集 L
      L=Split((A_i,B_i,E_i));
        for each (A_l,B_l,E_l)∈L{
          //状态有效性检查
          if (Validate((A_l,B_l,E_l))
            //将新状态并入集合 StateSet
            StateSet=StateSet ∪{(A_l,B_l,E_l)};
        }
        //状态(A_i,B_i,E_i)分析完毕
        StateSet.Remove(A_i,B_i,E_i);
    }
    return TRUE; //协议无攻击
}
```

对于不能判定安全性的状态 $s=(A,B,E)$，其中 $A=\langle\sigma_1,t_1,t_1'\rangle,\langle\sigma_2,t_2,t_2'\rangle,\cdots,\langle\sigma_n,t_n,t_n'\rangle$. 是一个迹序列,SPA 系统依据迹序列扩张方式,应用状态分裂算法 $Split$(算法 3)求其子状态集,能够有效扩展状态搜索空间,避免大量冗余状态的产生. 在状态分裂算法中首先需要应用算法 $Unify$(算法 2)求解出非正合节点的接收项 t 与攻击者初始知识集合 D 的合一变量替换集合.

变量替换合并运算 Ψ.

令集合 $A=\{a_1,a_2,\cdots,a_m\}$，$B=\{b_1,b_2,\cdots,b_n\}$,其中 $a_i=(\phi,d_{l,a},\cdots,d_{j,a})$,$b_i=(\delta,d_{l,b},\cdots,d_{j,b})$,$\phi,\delta$ 为变量替换,$d_l,\cdots,d_j\in D$. 变量替换合并运算 Ψ 的计算规则如下:

(1)若集合 $A=\{(\lambda,d_{l,a},\cdots,d_{j,a})\}$，$B=\{(\delta,d_{l,b},\cdots,d_{j,b})\}$，则 $C=\Psi(A,B)$定义为

如果变量替换 λ,δ 不矛盾,则

$$C=\{(\lambda\cup\delta,d_{l,a},\cdots,d_{j,a},d_{l,b},\cdots,d_{j,b})\},$$

否则 $C=\varnothing$;

(2)若集合 A 或 B 中的元素不只一项,则 $C=\Psi(A,B)$定义为

$$\{\Psi(a,b)\mid a\in A,b\in B\}.$$

特殊情况,当 $A=\varnothing$,$\Psi(A,B)=B$;当 $B=\varnothing$,$\Psi(A,B)=A$. 算法从略.

算法 2. $Unify$.

输入:消息项 t,攻击者初始知识集合 D

输出:t 的子项组合与 D 中项可以合一的变量替换集合 S_{Unify}

```
Unify(D,t){
  //初始化 S_Unify 为空集
  S_Unify=∅;
  //按照消息项递归定义方式求项 t 的子项组合集
  T={(t_1,t_2,…,t_m)|u(t_1,t_2,…,t_m)=t,u 为代数式}
  for (each t 的子项组合(t_1,t_2,…,t_m) ∈T){
    U=∅;// U 初始为空集
    //由合一算法 MGU 求项 t_i 和 d_i 的变量替换集合
    for (each t_i∈{t_1,t_2,…,t_m}){
      U_i={(λ_i,d_i)|λ_i(t_i) =λ_i(d_i), λ_i是变量替换,为项 d_i 和 t_i 最一般合一, d_i∈D};
      //应用 Ψ 运算规则合并变量替换集合
      U=Ψ(U,U_i);
    }
    //所有完全子项的合一集合求并集
    S_Unify=S_Unify ∪ U;
  }
}
```

算法 3. 状态分裂算法.

输入:待分裂的状态 $s=(A,B,E)$

输出:分裂得到的子状态集合 S'

```
Split(s){
  //从 A 中选取第一个非正合节点
  ⟨τ,t,t'⟩=SelNoExtNode(A);
  //标识⟨τ,t,t'⟩的前驱节点
  ⟨ω,p,p'⟩=PrevNode(A,⟨τ,t,t'⟩);
  //初始化子状态集为空集
  S'=∅;
  //通过算法 2 求项 t 与 D 中元素的可合一变量替换集合
  Unify(D,t)={(ϕ,m_1,…,m_k)|ϕ(t)=u{ϕ(m_1)/x_1,
              …,ϕ(m_k)/x_k},u 是代数式,m_i∈D,
              i=1,2,…,k};
  for(each(ϕ,m_1,m_2,…,m_k)∈Unify(D,t)){
    //依据迹序关系,分为如下几类情形生成新的迹序状态
    if (∀m_i,满足 m_i∈D_0,1≤i≤k){
      //添加一个新状态
      S'=S'∪{(ϕ(A),ϕ(B),ϕ(E)∪
          {ϕ(⟨τ,t,t'⟩)};
      continue; //处理下一个(ϕ,m_1,m_2,…,m_k)组
    }
    //集合 M_R 包含所有不属于 D_0 的 m_i
    M_R={m_1,m_2,…,m_k}-{m_i|m_i∈D_0且
         1≤i≤k};
    对于任一个 m_r∈M_R,根据 m_r→⟨σ_s,t_s,t_s'⟩关系表可查询到其起源节点 n_s=⟨σ_s,t_s,t_s'⟩;
    if (∃m_r∈M_R,其起源节点 n_s的变量替换满足
      ϕ(⟨σ_s,t_s,t_s'⟩)∈ϕ(A)且⟨τ,t,t'⟩≤_P n_s){
      S'=S'∪∅; //不添加新状态
      continue;//处理下一个(ϕ,m_1,m_2,…,m_k)组
    }
    if (∀m_r∈M_R,其起源节点 n_s的变量替换满足
      ϕ(⟨σ_s,t_s,t_s'⟩)∈ϕ(A)且 n_s≤_P⟨τ,t,t'⟩){
      S'=S'∪{(ϕ(A),ϕ(B),ϕ(E)∪
          {ϕ(⟨τ,t,t'⟩)};//添加一个新状态
      continue;//处理下一个(ϕ,m_1,m_2,…,m_k)组
    }
    for(每个 m_r∈M_R){
```

```
        if (m_r起源节点 n_s的变量替换满足
          φ(⟨σ_s,t_s,t'_s⟩)∉φ(A)){
          //正则迹中起源节点前的节点集
        G(m_r)={φ(⟨σ_r1(α,l),t_r1,t'_r1⟩),…,
                φ(⟨σ_rj(α,l),t_rj,t'_rj⟩)};
          G=∪_i G(m_i)−φ(A);
                //去除在 A 中已存在的节点
          for(G 中节点每个满足偏序关系≤_P的有效排
              列 Π_i){
            A'=φ(⟨σ_1,t_1,t'_1⟩)…φ(⟨ω,p,p'⟩)+Π_i+
              φ(⟨τ,t,t'⟩)…φ(⟨σ_n,t_n,t'_n⟩);
            //添加一个新状态
            S'=S'∪{(A',φ(B),φ(E)∪φ(⟨τ,t,t'⟩))};
            }
        }
      }
    }//结束所有 Unify(D,t)中项的遍历
    return (S'); //返回新生成的子状态集合
}
```

值得一提的是，由于 SPA 系统在协议预处理期间限制了攻击者角色数量和会话轮数，依据迹序关系，可匹配的节点将逐渐减少，从而保证了分裂算法收敛. 除此外，算法中的变量替换都遵循类型匹配一致原则等方法都有效减小了状态搜索空间.

与 Athena 系统的几点比较.　卡内基·梅隆大学的 Song 等[8,9]基于 Strand 空间研制的 Athena 系统是一种高效的密码协议自动分析工具. SPA 与 Athena、NRL 协议分析器等分析工具共同之处是均采用构造攻击过程的方法分析协议的安全性；然而，SPA 系统与 Athena 等其它系统相比采用了更有效的新模型和新的算法，主要体现在如下 4 个方面：

(1) CPA 模型中对密码协议的描述方式通常比 Strand 空间更为简洁. 例如，用 CPA 模型描述 Needham-Schroeder 公钥协议[1]仅用 4 个节点(迁移函数)，而在 Strand 空间中需要 6 个角色节点.

(2) Athena 基于的 Strand 空间模型属于图论方法，SPA 基于的 CPA 模型侧重的是代数方法. 在 Strand 空间中通过攻击者 Strand 处理消息运算，难以解决由运算产生的消息集无限性问题，因此，需要人为限制加密深度，并且消息项运算量大；而 SPA 采用具有代数等式和自由生成元(基)的 CPA 代数方法解决消息运算无限生成问题，不需要限制运算方式(如加密深度)，即解决了消息集无限性问题，因而分析结果更可靠.

(3) 协议的攻击序列表示方式不同，SPA 采用算法更直接有效：Athena 算法中 Strand 节点间属于一种消息绑定关系，搜索过程中状态图(bundles)无预期地生长；而 SPA 中迹序列是通过攻击者知识的递增过程表示. 这些区别使基于 CPA 的 SPA 系统在序列生成过程中，收敛更集中，对于一些协议会显著提高其分析效率.

(4) 基于 Strand 的 Athena 系统在状态搜索中主体存在变量类型，从而变量替换中可能重复替换，出现冗余. 基于 CPA 理论的 SPA 中，由于已有关于主体的约简理论，在预处理阶段对不同主体扮演各个角色已做了替换，状态搜索中迹序列中的主体都为常量类型.

4.2.3　输出部分

在 SPA 系统的输出部分，协议分析人员可以获得协议分析的结果信息以及各角色安全目标定理的证明树(结构如图 3 所示)，其中每个节点除记录该状态的自身信息外，还包括分析过程中对该状态属性的判定结果等.

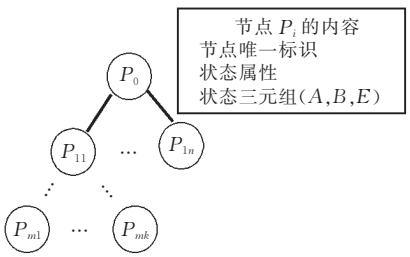

图 3　证明树结构

对于证明树中的节点 P_i，如果其中存储的状态 $s=(A,B,E)$ 属于不安全状态，则迹序列 A 是一个正合序列，描述了协议的具体攻击过程. 一般来说，迹序列 A 中的节点是经过不断扩张生成的，它的来源可能是任意一个主体的迹序列，其中的消息项也都经过了多次的变量替换、等价变换，具体形式可以在下一节的应用实例分析中了解到.

对于 SPA 系统的实现，我们在充分考虑协议分析算法实现效率的同时，提供了友好的人机交互界面和自动处理功能，以方便协议分析人员的使用.

5　SPA 系统的应用

SPA 系统是一个有效的网络系统安全性评测以及密码协议设计辅助工具. 我们借助 SPA 系统分析了一些典型的密码协议，得到了较理想的结果，鉴于篇幅有限，这里仅给出 BAN-Yahalom 协议[10]的一个新攻击过程分析.

5.1　BAN-Yahalom 协议分析

BAN-Yahalom 协议是 Burrows，Abadi 和 Need-

ham 在文献[10]中对 Yahalom 协议的一个修改版本.这个协议的目的是,通信双方通过可信第三方实现相互的身份认证,并建立一个新的会话密钥.它的原始描述如下.

1. $A\rightarrow B:A,N_a$
2. $B\rightarrow S:B,N_b,\{A,N_a\}_{K_{bs}}$
3. $S\rightarrow A:N_b,\{B,K_{ab},N_a\}_{K_{as}},\{A,K_{ab},N_b\}_{K_{bs}}$
4. $A\rightarrow B:\{A,K_{ab},N_b\}_{K_{bs}},\{N_b\}_{K_{ab}}$

其中,A,B 和 S 分别是协议的三个角色 Alice,Bob 及 Server 的简称(在消息中则代表各角色的身份标识),N_a 和 N_b 分别是由 A 和 B 生成的随机数,K_{ab} 是由可信第三方为 A,B 生成的新会话密钥,K_{as} 和 K_{bs} 则分别是 A,B 与 S 之间所共享的长期密钥.

下面给出 BAN-Yahalom 协议在 CPA 模型中的描述:

角色 A 的迹序列 $\Sigma_1(\alpha,l)$:

$\langle\sigma_{11}(\alpha,l),\Delta,(A,N_a)\rangle$

$\langle\sigma_{12}(\alpha,l),(x,\{B,k,N_a\}_{K_{as}},$

$\{A,k,x\}_{K_{bs}}),(\{A,k,x\}_{K_{bs}},\{x\}_k)\rangle,$

这里 x,k 为 Nonce 类型的变量,分别对应于 N_b 和 K_{ab};

角色 B 的迹序列 $\Sigma_2(\alpha,l)$:

$\langle\sigma_{21}(\alpha,l),(A,y),(B,N_b,\{A,y\}_{K_{bs}})\rangle$

$\langle\sigma_{22}(\alpha,l),(\{A,k',N_b\}_{K_{bs}},\{N_b\}_{k'}),\Delta\rangle,$

其中,变量 y,k' 分别对应于 N_a 和 K_{ab};

角色 S 的迹序列 $\Sigma_3(\alpha,l)$:

$\langle\sigma_{31}(\alpha,l),(B,z,\{A,w\}_{K_{bs}}),$

$(z,\{B,K_{ab},w\}_{K_{as}},\{A,K_{ab},z\}_{K_{bs}})\rangle,$

其中,变量 z 和 w 分别对应于 N_b 和 N_a.

协议的对应性安全目标定理:

$\Sigma_1(\alpha,1)Corresp\langle\sigma_{21}(\alpha,1),(A,N_a),(B,N_b,\{A,N_a\}_{K_{bs}})\rangle$

$\Sigma_2(\alpha,1)Corresp\Sigma_1(\alpha,1)\{N_b/x,K_{ab}/k\}.$

SPA 系统接受的输入是类似上面这种描述的协议规范描述语言(语义相同,但形式上稍有区别).在 PC 机(CPU 为 PIII450MHz,内存 128MB)上运行 SPA 系统,在不到 1s 时间内就得到分析的结果:角色 A 的安全目标成立,但角色 B 的安全目标则不成立.系统给出了 5 个不安全状态,从这些不安全状态中,得到一个目前公开文献未见的一个攻击,在 SPA 中的正合序列描述如下:

$\langle\sigma_{11}(\alpha,1),\Delta,(A,N_a)\rangle\rightarrow$

$\langle\sigma_{21}(\alpha,1),(A,N_a),(B,N_b,\{A,N_a\}_{K_{bs}})\rangle\rightarrow$

$\langle\sigma_{21}(\beta,1),(B,N_a),(B,N'_a,\{B,N_b\}_{K_{as}})\rangle\rightarrow$

$\langle\sigma_{31}(\beta,1),(A,N_a,\{B,N_b\}_{K_{as}}),(N_a,\{A,K'_{ab},N_b\}_{K_{bs}},$

$\{B,K'_{ab},N_a\}_{K_{as}})\rangle\rightarrow$

$\langle\sigma_{12}(\alpha,1),(N_b,\{B,K'_{ab},N_a\}_{K_{as}},\{A,K'_{ab},N_b\}_{K_{bs}},$

$(\{A,K'_{ab},N_b\}_{K_{bs}},\{N_b\}_{K'_{ab}})\rangle\rightarrow$

$\langle\sigma_{21}(\alpha,1),(\{A,K'_{ab},N_b\}_{K_{bs}},\{N_b\}_{K'_{ab}},\Delta)\rangle,$

其中 $\alpha=(A,B,S),\beta=(B,A,S)$.

用通常的方法来描述该攻击的过程为

$(\alpha.1)A\rightarrow B:A,N_a;$

$(\alpha.2)B\rightarrow S:B,N_b,\{A,N_a\}_{K_{bs}};$

$(\alpha.3)\cdots;$

$(\beta.1)P(B)\rightarrow A:B,N_b;$

$(\beta.2)A\rightarrow P(S):A,N'_a,\{B,N_b\}_{K_{as}};$

$(\beta.2')P(A)\rightarrow S:A,N_a,\{B,N_b\}_{K_{as}};$

$(\beta.3)S\rightarrow P(B):N_a,\{A,K'_{ab},N_b\}_{K_{bs}},\{B,K'_{ab},N_a\}_{K_{as}};$

$(\alpha.3')P(S)\rightarrow A:N_b,\{B,K'_{ab},N_a\}_{K_{as}},\{A,K'_{ab},N_b\}_{K_{bs}};$

$(\alpha.4)A\rightarrow B:\{A,K'_{ab},N_b\}_{K_{bs}},\{N_b\}_{K'_{ab}}.$

在这个重放攻击过程中,攻击者 P 冒充 B 的身份发起一个并发的协议会话进程,篡改了原协议会话第三步中 S 发给 A 的消息,从而达到了为 A,B 分配一个可能被攻破的旧密钥(K'_{ab})的目的.由于 A,B 间的双向认证并未被破坏,这就欺骗了 A,B,使他们误以为 K'_{ab} 是 S 为本轮会话生成的新会话密钥,从而违反了协议的安全性目标(会话密钥的新鲜性).

文献[11]给出了 BAN-Yahalom 协议的两种攻击,但这些攻击都需要在一定的条件假设下才能成立,其中所列的第一种攻击需要假设"角色 B 在第一步中不检查收到的新鲜值 N_a 的长度"(一般在协议的实现中会规定所用的新鲜值的比特位数),第二种攻击则需要假设"S 在第二步中不对接收到的消息中的 N_a 和 N_b 进行比较"(一般认为在同一次协议会话过程中所用的新鲜值应互不相同,即 $N_a\neq N_b$),然而 SPA 发现的新攻击不需要上述假设条件就可成功对协议攻击,以致这类攻击更容易被攻击者利用.

5.2 分析结果

我们借助 SPA 系统对 Needham-Schroeder 对称密钥[12]及公开密钥协议(以及修正后的版本)[13]、Woo-Lam 的单向认证协议(及其 4 个修改版本)[14]、TMN 协议以及 Kerberos 协议(简化版)[6]等进行了分析,都得到了与以前的研究相似的结果.表 1 给出了部分实验数据,需要说明的是,在分析具体协议时,能够通过配置相应的策略参数使 SPA 系统更快给出分析结果.

表 1　其它协议分析结果

协议	搜索状态数	不安全状态数
BAN-Yahalom 协议	17	5
Needham-Schroeder 对称密钥协议	8	1
Needham-Schroeder 公开密钥协议	8	2
Needham-Schroeder-Lowe 公开密钥协议	29	0
Woo-Lam 单向认证协议(版本 Πf)	2	0
Woo-Lam 单向认证协议(版本 Π1)	2	0
Woo-Lam 单向认证协议(版本 Π2)	4	1
Woo-Lam 单向认证协议(版本 Π3)	4	1
Woo-Lam 单向认证协议(版本 Π)	7	2
TMN 协议	5	4
Kerberos 协议(简化版)	6	1

与其它同类系统的效率比较：从目前的公开文献中，我们获得一些同类系统对部分密码协议的分析结果数据，从表 2 中可以看出 SPA 系统与其它同类系统相比，减小了状态搜索空间，提高了分析效率.

表 2　与同类系统分析效率对比表

协议	Murφ	Brutus	Athena	STA	SPA
Needham-Schroeder 公开密钥协议	1706	1208	36	*	8
Needham-Schroeder-Lowe 公开密钥协议	*	146	19	60	15
TMN 协议	*	3327	*	*	5
Kerberos 协议(简化版)	*	3405	*	*	6

注：表中为各个系统找到攻击或完成证明搜索的状态数，* 表示未获得相关实验数据. Murφ 是斯坦福大学 1992 年研制的模型检测工具，1997 年应用于密码协议分析[6]；Brutus 是卡内基·梅隆大学等 1996 年研制的模型检测工具[5]；Athena 是卡内基·梅隆大学 2001 年研制的密码协议分析工具[8,9]；STA 是意大利佛洛伦萨大学 2001 年研制的密码协议分析工具[15].

6　相关工作比较与结论

本文基于 CPA 模型及其安全性分析技术，设计并实现了一个新的密码协议安全性分析系统(SPA)，SPA 系统与其它同类系统相比，具有以下特点：

(1) 基于新的密码协议代数(CPA)模型规范密码协议，并采用新的高效的分析算法对密码协议的安全目标进行证明.

(2) SPA 系统具有较高的可靠性，能够给出协议的安全性证明或构造出相应攻击序列，特别是发现了一个新的攻击实例.

(3) 从协议的分析结果数据来看，SPA 系统有效减小了状态搜索空间，提高了分析效率.

(4) SPA 系统能够直观显示分析结果，克服了许多协议分析工具给出的分析结果复杂抽象、难于理解的问题.

SPA 系统作为新研制的密码协议分析系统，仍需要不断改进和完善，以适用于分析更广泛类型的密码协议.

参　考　文　献

1 Lowe G.. Breaking and fixing the Needham-Schroeder public-key protocol using FDR. Software-Concepts and Tools, 1996, 17(3): 93～102

2 Huai Jin-Peng, Li Xian-Xian. Algebra model of cryptographic protocols and their security. Science in China(Series E), 2003, 33(12): 1087～1106(in both Chinese and English)
(怀进鹏,李先贤. 密码协议的代数模型及其安全性. 中国科学(E 辑), 2003, 33(12): 1087～1106)

3 Qing Si-Han. Twenty years development of security protocols research. Journal of Software, 2003, 14(10): 1740～1752(in Chinese)
(卿斯汉. 安全协议 20 年研究进展. 软件学报, 2003, 14(10): 1740～1752)

4 Millen J.. The Interrogator model. In: Proceedings of the 1995 IEEE Symposium on Security and Privacy, Oakland, California, USA, 1995, 251～260

5 Clarke E. M., Jha S., Marrero W.. Verifying security protocols with Brutus. ACM Transactions on Software Engineering and Methodology, 2000, 9(4): 443～487

6 Mitchell J. C., Mitchell M., Stern U.. Automated analysis of cryptographic protocols using Murφ. In: Proceedings of the 1997 IEEE Symposium on Security and Privacy, Oakland, California, USA, 1997, 141～153

7 Meadows C.. A model of computation for the NRL protocol analyzer. In: Proceedings of the 1994 Computer Security Foundations Workshop, Franconia, NH, USA, 1994, 84～89

8 Song D.. Athena: A new efficient automatic checker for security protocol analysis. In: Proceedings of the 1999 IEEE Computer Security Foundations Workshop. Los Alamitos: IEEE Computer Society Press, 1999, 192～202

9 Song D., Beresin S., Perrig A.. Athena: A novel approach to efficient automatic security protocol analysis. Journal of Computer Security, 2001, 9(1,2): 47～74

10 Burrows M., Abadi M., Needham R.. A logic of authentication. In: Proceedings of the Royal Society of London A, 1989, 426: 233～271

11 Syverson P.. A taxonomy of replay attacks. In: Proceedings of the Computer Security Foundations Workshop, Franconia NH, USA, 1994, 187～191

12 Needham R., Schroeder M.. Using encryption for authentication in large networks of computers. Communications of the ACM, 1978, 21(12): 993～999

13 Lowe G.. An attack on the Needham-Schroeder public-key authentication protocol. Information Processing Letters, 1995, 56(3): 131～133

14 Woo T., Lam S.. A lesson on authentication protocol design.

Operating Systems Review, 1994, 28(3): 24～37

15 Boreale M., Buscemi M.. Experimenting with STA, a tool for automatic analysis of security protocols. In: Proceedings of the 2002 ACM Symposium on Applied Computing, Madrid, Spain, 2002, 281～285

LI Jian-Xin, born in 1979, Ph. D. candidate. His current research interests include network security, trust negotiation.

LI Xian-Xian, born in 1969, Ph. D., associate professor. His current research interests include information security, computer science theory.

ZHUO Ji-Liang, born in 1976, master. His current research interests include network security.

HUAI Jin-Peng, born in 1962, Ph. D., professor, Ph. D. supervisor. His current research interests include computer software and theory, network middleware and grid computing, network security.

Background

The project, "Technologies and Systems of Network Protocol Analysis", is supported by the National Natural Science Foundation of China under grant No. 90412011, the National High-Tech Research and Development Program of China(863 Program) under grant No. 2003AA144150. This project aims to build a new model for security protocol analysis, propose new efficient analysis algorithm and develop an automatic verification system for security protocol analysis. The security protocol group for the project has built a new algebra called Cryptographic Protocol Algebra (CPA) for security protocol analysis using some mathematical techniques and designed some special protocols such as group key distribution protocol, a fair non-repudiation security protocol. Authors design and implement an automatic analysis system (Security Protocol Analyzer, SPA) based on the CPA (Cryptographic Protocol Algebra) model and new analysis techniques. More than ten cryptographic protocols have been successfully analyzed with this system, and moreover, experiment data have shown that SPA is very efficient.