模糊入侵识别引擎的研究与设计*)

徐 慧 戚 涌 张 宏 刘凤玉

(南京理工大学计算机科学与技术系 南京 210094)

摘 要 模糊入侵识别引擎是一个用模糊理论来针对计算机网络的恶意活动的网络入侵检测系统。本文将模糊理论中知识的模糊表示、特征的模糊匹配及模糊推理用于入侵检测中,提出了一种新的模糊入侵识别引擎(IFIRE)。以具有模糊属性的特征元素为最小成份,组成模糊特征因子、模糊特征表达式及模糊特征树来描述具有模糊特性入侵活动特征的知识体系;通过特征因子的相似度计算进行特征的模糊匹配;最后,用基于产生式规则的模糊推理进行检测决策。该方法能有效地降低误报率及漏报率。

关键词 模糊,入侵检测,特征,匹配,推理

Design of Fussy Intrusion Recognition Engine

XU Hui QI Yong ZHANG Hong LIU Feng-Yu

(Department of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing 210094)

Abstract The Fuzzy Intrusion Recognition Engine (FIRE) is a network intrusion detection system that uses fuzzy systems to assess malicious activity against computer networks. This paper originally explores some fuzzy theories, including fuzzy knowledge expression, fuzzy match and fuzzy inference, geares towards intrusion detection. The model of FIRE based on above fuzzy theories is built and discussed in detail. The intrusion features are presented by fuzzy elements, fuzzy factors, fuzzy expression and fuzzy trees. Fuzzy match is carried out by calculate of resemblance. Finally decisions are made by fuzzy inference. These methods can effectively improve the false negative rate and false positive rate of IDS.

Keywords Fuzzy, Intrusion detection, Features match, Inference

1 引言

当前应用在入侵检测引擎设计中的两种主要技术,一种是基于统计分析的入侵检测技术,另一种是基于规则/特征的检测方法。统计分析常用于异常检测中,即通过得到一组测量值与正常值的偏离来预测用户行为的变化,然后做出决策判断的检测技术。但由于:a)入侵而非异常;b)非入侵且是异常;c)非入侵且非异常;d)入侵且异常等复杂情况,使得检测结果误报率(false negative rate)高。基于规则/特征的检测方法,常用于误用检测中,即假设攻击能够被精确用某种方式的编码表示,根据已知的入侵模式来检测入侵。总体上说,这种方法对未知入侵无能为力,因而漏报率(false positive rate)高。新的方法不断被引入,如数据挖掘、神经网络、遗传算法、免疫机理等[1,3]。当人们注意到入侵检测中的模糊性后,模糊概念被引入,出现了多种模糊入侵识别引擎(Fuzzy Intrusion Recognition Engine,FIRE),如模糊数据挖掘、模糊神经网络、模糊遗传算法等在入侵检测中应用的研究[1~3]。目前,这些技术大都处于研究探讨阶段。在商业化的软件产品中,应用最广泛的就是特征分析技术。因此,根据入侵检测的"模糊特性",把"模糊"引入特征分析中,是IDS研究中最具备实用价值的研究。

本文提出了一种新的基于推理的模糊入侵识别引擎(Inference_based Fuzzy Intrusion Recognition Engine, IFIRE)的设计,把模糊概念引入判据的模糊表示(3.1节)、特征的模糊匹配(3.2节)、基于模糊树的模糊推理(3.3节)中,并根据入侵检测的特点,提出了具体的算法,就模糊理论在入侵检测中应用方面具有一定的独创性。与John E. Dickerson设计的FIRE相比[3],由于充分考虑了特征之间的关联与因果关系,将不仅仅用于基于数据包的异常检测,而且在日志分析、系统行为序列分析等方面都可应用,并在降低误报率和漏报率方面有明显改善。

2 入侵检测中模糊性及系统模型

2.1 入侵检测中的模糊性

任何试图危害资源的完整性、保密性和可用性的活动集合,被称为入侵。随着计算机及网络应用的日益广泛深入,入侵活动越来越多样性、频繁性,加之这种活动本身的特性,使得入侵检测在多个方面呈现出模糊性,可概括为以下三个方面:(1)特征的模糊性:入侵检测中包括许多量的特性,使用模糊集的表示,更有助于特征的创建。例关于一定时间内某主机被建立的连接数,与异常之间就是一个模糊隶属关系。(2)特征匹配的模糊性:为了反入侵,攻击者经常变化手段、策略等,即使同一类入侵会以多种形式出现,这要求检测时不能求得精确匹配,更应注重检测到的特征与先验知识的相近程度。(3)判据的模糊性[4]:一方面,系统自身定义的保密性和完整性标准不足以作为入侵的评判标准,例如违反系统访问控制的活动不一定是入侵,它可能是合法用户的误操作或非法用户出于好奇的偶尔操作;另一方面,系统正常超越和异常超越的界限比较模糊,例资源消耗与遭受Dos攻击之间存在一种模糊关系。

2.2 基于模糊匹配与推理的FIRE设计

整个FIRE主要由以下几个部分组成:数据采集组件、分

*)项目资助:国防科工委应用基础基金(NO. J1300D004)、江苏省自然科学基金(NO. BK2001055)、江苏省南通市青年学术带头人带课题进修计划(NO. Z3008)。徐 慧 博士研究生,研究方向:人工智能与信息安全;戚 涌 博士研究生,研究方向:网络故障诊断与信息安全;张 宏 博士生导师,研究方向:软件工程;刘凤玉 博士生导师,研究方向:信息安全与多媒体技术。

析组件、知识获取组件、模糊运算组件、报警组件、用户界面及模糊知识库等，如图 1 所示。

知识获取组件：用于机器学习时对知识的获取，知识来源于专家关于入侵检测的知识的模糊表示。知识的输入可以采用多种形式。元素的形式、知识因子的形式或规则形式，分析组件根据不同的形式进行处理，形成模糊推理树，存取模糊知识库。

数据采集组件：主要收集并整理与安全相关的信息，提供给模糊证据分析器，从中提取特征的数据，数据形式为元素。收集对象包括系统日志、网络数据流、目标系统各进程的关键运行状态等。它也可以是一个组件组，由若干个组件组成，如 TCP 连接组件，用于收集 TCP 报连接情况，类似有 IP 报连接组件等。

模糊匹配组件：依据模糊知识库对入侵模糊证据的定义，把特征数据库中的数据与之进行模糊匹配（即计算相近程度），以决定此特征因子是否作为模糊树的一部分，参予后续的推理。

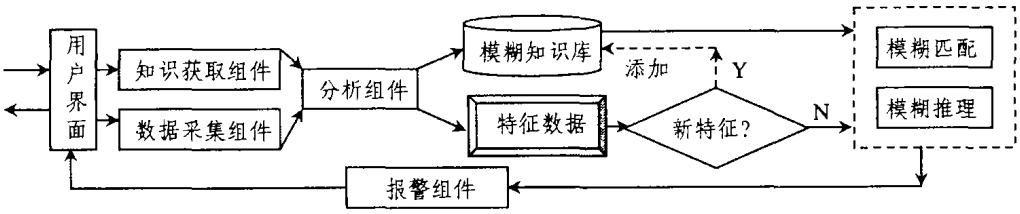

图 1　IFIRE 逻辑结构图

模糊推理组件：选择知识库中相关的模糊产生式规则，根据上述已获得的推理前提，进行模糊推理，本文用前提的可信度、结论的支持度、规则可信度及规则触发阈值反映推理的模糊性。

报警组件：是针对推理结论采取的动作，以某种方式告知检测员检测结论及辅助说明，例如响铃提示、屏幕提示、结束进程等。

3　系统设计与实现

本文采取的模糊产生式规则一般形式为：IF $P(\mu_p)$ THEN $R(\mu_r)$ (CF,τ)或 $R(\mu_r)\leftarrow P(\mu_p),(CF,\tau)$。表示：如果前提 P 在 $\mu_p(0\leqslant\mu_p\leqslant 1)$程度上成立，则可推出在 $\mu_r(0\leqslant\mu_p\leqslant 1)$程度上成立的结论 R，规则的可信度为 $CF(0\leqslant CF\leqslant 1)$，规则触发条件为：$\mu_p\geqslant\tau$。如果规则被触发，$R$ 的可信度为：$\mu_r = CF * \mu_r$。下面从判据的模糊表示、模糊匹配及模糊推理三个方面叙述。

3.1　特征的模糊表示

特征是进行检测的依据。本文以以下层次表示它：特征元素→特征因子→特征表达式。

（1）特征元素　是特征的最小构成单位，根据对入侵情况的分析，将特征元素分为数元素、符号元素、真值元素 3 种。

〈数元素〉::〈整数〉|〈实数〉|〈模糊数〉；

〈符号元素〉::＝〈字母〉|〈符号元素〉〈字母〉|〈符号元素〉〈数字〉；

〈真值元素〉::＝True |False |〈语言真值〉

设 A 是某一个入侵特征属性的集合，对每一个 $x\in A$，有一个对应的论域 U（一个集合），$F(U)$表示 U 上的全体模糊集构成的一个集合，$v\in U$，则特征元素的表示形式为：$x\overset{\mu_x}{\theta}v$ 其中，θ 为运算符，可以为：＝、＞、＜、≮、≯、≤、≥等，μ_x 是特征元素值的隶属度，$0\leqslant\mu_x\leqslant 1$。

例 1：TCP 包头中出现标志位示 FIN，表示为：$f\ fin\overset{1}{=}1$

例 2：用户 30min 内，登录次数超过 10 次，可信度 0.95，表示为：$login\ on\overset{0.95}{>}10$

（2）特征因子　由特征元素经过某种运算组成，这些元素组合起来可以反映某一入侵全局或局部的特征，用图 2 表示，$x1,x2,\cdots$为特征元素，f 是特征因子的名字，μ_f：$0<\mu_f\leqslant 1$ 是该因子的可信度。

〈运算〉::＝ 一元运算|二元运算；

〈一元运算〉::＝ NOT|∃ |∀ ；

〈二元运算〉::＝ AND | OR

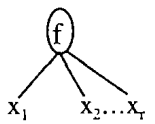

图 2　特征因子图符

（3）特征树及特征表达式　特征库可以形象地用多棵树表示。由能够比较完整判断某一类入侵的特征元素及其相关规则组成一棵树，如图 3 示。图中画出 4 层，具体层数由特征之间的关系及推理来定。图中，0 层为树叶层，其节点为特征元素，与节点相连的线上的 μ，表示该特征元素的值的隶属度；1 层为特征因子层，若干有联系的属性反映出某一特征属性，节点旁边的 μ 表示该特征属性存在的可信度；2 层为若干特征因子组合起来的特征表达式，反映出更详细的、更准确的特征信息；树根为检测结论，即系统是否被入侵或遭到了何种入侵及其可能性。特征因子、特征表达式、树根之间的连线，表示规则，线上的 CF,τ,μ 分别表示规则的可信度、该规则被触发的条件及结论本身的模糊程度。图 3 表示了产生组：

1：$f11(\mu_{11})\leftarrow x1,x2\ x3$
2：$f12(\mu_{12})\leftarrow x4,x5$
3：$f13(\mu_{13})\leftarrow x6$
4：$e21(\mu_{21})\leftarrow f12,f13\ (CF_{21},\tau_{21})$
5：$R1(\mu_{1r})\leftarrow f11(CF_{1r},\tau_{1r})$
6：$R1(\mu_{2r})\leftarrow e21(CF_{2r},\tau_{2r})$

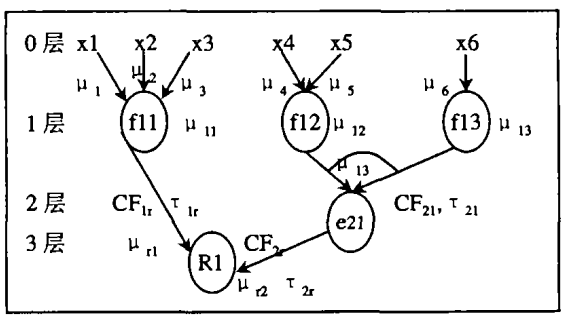

图 3　特征树

3.2　模糊匹配—特征的相似度计算

特征定义永远要在效率和精确度间取得折衷。大多数情况下，简单特征比复杂特征更倾向于误报，因为前者很普遍；

复杂特征比简单特征更倾向于漏报，因为前者太过全面，攻击软件的某个特征会随着时间的推进而变化。知识库中的知识包括特征元素、特征因子、特征表达式及规则一般通过以下途径获取：(1)对已知入侵的分析；(2)现有软件、硬件的漏洞；(3)专家的经验，这些都是基于过去的情况，现在或将来往往很难与之100%匹配，因此，检测中不能追求精确匹配，本文通过计算已测特征元素与知识库中已有特征因子所需特征元素之间的相似程度，确定是否选择该特征因子作为入侵的判据，并根据计算的相似度，对特征因子的可靠程度进行打折，使其尽量反映真实情况。相似度的计算参考文[5]，选择余弦法，具体如下。

设有模糊向量 $\boldsymbol{A}=(a_1,a_2,\cdots,a_n)$，$\boldsymbol{B}=(b_1,b_2,\cdots,b_n)$，则 $\boldsymbol{A},\boldsymbol{B}$ 之间的相似度 r_{AB} 为：

$$r_{AB}=\frac{\sum_{i=1}^{n}a_i*b_i}{\sqrt{(\sum_{i=1}^{n}a_i^2)*(\sqrt{\sum_{i=1}^{n}b_i^2})}}$$

如图3中所示，特征因子 $f11$ 由特征元素 $x1,x2,x3$ 决定，已知特征库中，$x1,x2,x3$ 的值的模糊度分别为 μ_1,μ_2,μ_3，用向量 $\boldsymbol{A}=(\mu_1,\mu_2,\mu_3)$ 表示，设实际测得特征元素 $x1,x3,x2$ 未出现，用向量 $\boldsymbol{B}=(\mu_1,0,,\mu_3)$ 表示，则模糊向量 $\boldsymbol{A},\boldsymbol{B}$ 之间的相似程度为：$r_{AB}=(\mu_1^2+\mu_3^2)/((\mu_1^2+\mu_2^2+\mu_3^2)*(\mu_1^2+\mu_3^2))$。设特征因子 $f11$ 的模糊程度为 μ_{12}，因为特征元素 $x2$ 未出现，则 $\mu'_{12}=r_{AB}*\mu_{12}$。

3.3 模糊推理

推理从1层开始，逐步向树根方向深入。分为两种情况：一是前提有多个单一证据的合取或析取；二是多个独立证据推出同一假设。算法设计应该满足下列要求：随着判据的丰富，结论的可信度增强，反之则减弱。

1)前提有多个单一证据的合取或析取，如图3中 $e21$ $(\mu_{21})\leftarrow f12,f13$ (CF_{21},τ_{21})。设有如下规则：IF $A(\mu_A)$ AND $B(\mu_B)$ THEN $f1(\mu_f)(CF_1,\tau 1)$，规则的前提是两个独立证据 A、B 的合取，则规则的触发条件是当且仅当满足下列不等式：$\mu'=\mu_A\wedge\mu_B\geqslant\tau 1$，即 $\min(\mu_A,\mu_B)\geqslant\tau 1$，结论的可信度为：$CF_1*\mu'$。

设有如下规则：IF $A(\mu_A)$ OR $B(\mu_B)$ THEN $f1(\mu_f)$ $(CF_1,\tau 1)$，规则的前提是两个独立证据 A、B 的析取，则规则的触发条件是当且仅当满足下列不等式：$\mu_A\geqslant\tau 1\vee\mu_B\geqslant\tau 1$，取 $\mu'=\max(\mu_A,\mu_B)$，结论的可信度为：$CF_1*\mu'$。

2)多个独立证据推出同一假设，如图3中 $R1(\mu_{1r})\leftarrow f11$ (CF_{1r},τ_{1r})、$R1(\mu_{2r})\leftarrow e21$ (CF_{2r},τ_{2r})。设有如下规则：IF A (μ_A) THEN $R(\mu_{Ra})$ (CF_A,τ_A)、IF $B(\mu_B)$ THEN $R(\mu_{RB})$ (CF_B,τ_B)，每条规则触发条件分别为：$\mu_A\geqslant\tau_A$、$\mu_B\geqslant\tau_B$，被触发规则的结论的可信度分别为：$CF*\mu$，结论的综合可信度由下式决定：

$$\mu'_R=\begin{cases}\mu'_a+\mu'_B-\mu'_A*\mu'_B & \mu'_A\geqslant 0,\mu'_B\geqslant 0\\ \mu'_A+\mu'_B+\mu'_A*\mu'_B & \mu'_A<0,\mu'_B<0\\ \dfrac{\mu'_A+\mu'_B}{1-Min\{|\mu'_A|,|\mu'_B|\}} & \mu'_A\text{ 与 }\mu'_B\text{ 异号}\end{cases}$$

3.4 应用举例

1)特征信息及分析　Synscan是一个流行的用于扫描和探测系统的工具，它发出的信息包具有多种可分辨的特性。设通过它得到如下一组信息：不同的来源IP地址信息，TCP来源端口21，目标端口21，服务类型0，IP鉴定号码39426(IP identification number)，设置SYN和FIN标志位，不同的序列号集合(sequence numbers set)，不同的确认号码集合(acknowledgment numbers set)，TCP窗口尺寸1028。

上述数据分析如下：(1)只具有SYN和FIN标志集的数据包，这是公认的恶意行为迹象。(2)没有设置ACK标志，但却具有不同确认号码数值的数据包，而正常情况应该是0。(3)源端口与目标端口相同，即端口"反身"(reflexive)。虽然大多数情况下它们并非预期数值，但"反身"端口本身并不违反TCP标准。(4)源端口和目标端口都被设置为21的数据包，通常为与FTP服务器关联。(5)TCP窗口尺寸为1028，IP鉴定号码在所有数据包中为39426。根据IP RFC的定义，这2类数值应在数据包间有所不同，如果持续不变，就表明可疑，因此确定特征元素为：TCP包的SYN、FIN、ACK标志，TCP窗口大小属性、IP鉴定号属性、具有不同确认号码集合的数据报数，分别用f_SYN、f_FIN、f_ACK、TCP_WinSize、IP_in、Packages DAN表示。

2)推理判断　根据上述分析，创建下列模糊推理树，如图4所示。

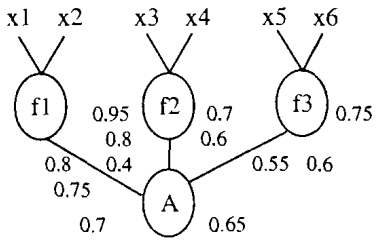

图4　实例特征树

$x1:f\_SYN\overset{1}{=}1$　　$x2:f\_FIN\overset{1}{=}1$

$x3:IP_in\overset{1}{=}$常数　　$x4:TCP_winSize\overset{1}{=}$常数

$x5:f\_ACK\overset{1}{=}0$　　$x6:packages\_DAN\overset{1}{\neq}0$

Rule1(恶意数据报)：$f1(0.95)\leftarrow x1\wedge x2$
Rule2(异常)：$f2(0.7)\leftarrow x3\wedge x4$
Rule3(异常)：$f3(0.75)\leftarrow x5\wedge x6$
Rule4(攻击)：$A(0.7)\leftarrow f1(0.8,0.65)$
Rule5((攻击)：$A(0.5)\leftarrow f1(0.7,0.6)$
Rule6((攻击)：$A(0.65)\leftarrow f1(0.55,0.6)$

(设特征元素与知识库中的特征元素正好相配，根据3.3节的2)，结论的可信度计算如下：

Step1　由Rule4、Rule5、Rule6均符合触发条件，若分别被触发得到 A 的可信度 μ'_1、μ'_2、μ'_3 为：$\mu'_1=0.7*0.8=0.56$；$\mu'_2=0.5*0.7=0.35$；$\mu'_3=0.65*0.55=0.3575$

Step2　由 μ'_1、μ'_2 得 $\mu'_{12}=0.56+0.35-0.56*0.35=0.714$

Step3　由 μ'_3、μ'_{12} 得 $\mu'_A=\mu'_{123}=0.3575+0.714-0.3575*0.714=0.8162$

所以，系统基于通讯的数据报信息的异常，判断出受到攻击的可能性为：81.62%。

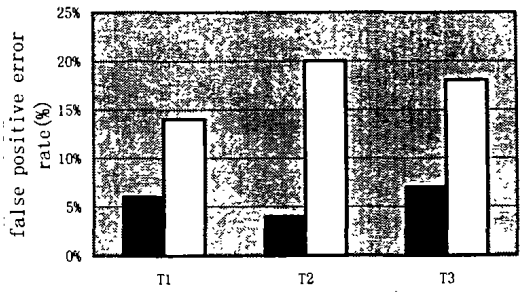

图5　误报率的比较

结论　在IDS中，把特征数据映射到模糊集，可以更好地找出入侵检测依据，从而降低误报率和漏报率[3,4]；如果在

检测过程中，引入模糊匹配及模糊推理，可以进一步降低误报率和漏报率，特别是对原来攻击的变形所形成的新攻击，本文初步实验也证明了这点。通过在局域网上对 Denial of Service、Portscan、Unauthorized Service 等攻击行为的预演与检测，得到图 5、图 6 所示结果。

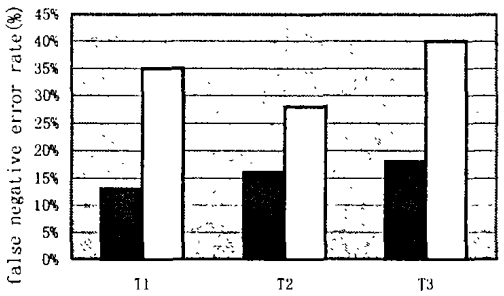

图 6　漏报率的比较

图中 T1、T2、T3 分别代表上述 3 种类型的行为；深颜色是未考虑模糊的情况，即推理过程中采用精确匹配及精确推理；浅颜色为本文提出方法的实验结论。需要说明的是，实验结果与所建立特征库、采用的攻击手段有极大关系，本图表是对若干种手段测试的平均值。

基于特征分析的 IDS 是目前商用 IDS 中使用最多的一种方法，提高它检测的准确率具有很深的现实意义，如何全面地获取特征信息是所有基于特征分析的 IDS 的难点之一，是影响 IDS 性能主要原因。另外，为了解决 IDS 的跨平台、跨语言问题，特征的表述是关键。这些问题是需一步研究的问题，也是本设计推向应用前所必须解决的问题。

参 考 文 献

1　张琨，徐永红，王珩，刘凤玉．基于免疫学的入侵检测系统模型．南京理工大学学报，2002，26(4)：337～340

2　Bridges S M，Vaughn R B，FUZZY DATA MINING AND GENETIC ALGORITHMS APPLIED TO INTRUSION DETECTION. http://www.cs.msstate.edu/～bridges/papers/nissc2000.pdf

3　Dickerson J E，Dickerson J A. Dickerson. 2000. Fuzzy network profiling for intrusion detection [J]. In：Proc. of NAIFIPS 2000. 301～306

4　Luo Jianxiong，Bridges S M，Vaughn R B. Jr. Fuzzy Frequent Episodes for Real-Time. http://www.cs.msstate.edu/～bridges/papers/fuzzieee-2001.pdf

5　何新贵．模糊知识处理的理论与技术[M]．国防工业出版社，1999. 63～65

（上接第 79 页）

免有经验的用户最终发现 *LICENSE* 号的存在而复制、传播，将之放在用户的硬盘上，而不以文件形式存储在 *MP3* 播放器的文件夹中。每次用户听歌时必须经历一次验证过程，而每次验证过程必须在 *LICENSE* 数组中查找，随着用户付费的歌曲增多，*LICENSE* 数组会越来越大，因此给 *LICENSE* 数组设置一个上限，比如：200。然后在 *LICENSE* 类中加一个 *TIME* 域，当 *TIME* 为 0 时在数组中删除这个对象。这样可以有效减少查询的时间，使插件的运行速度比较快，让合法用户感觉不到 *MP3* 疫苗免疫数字细菌的过程。同时由于用户换机器必须重新注册安装插件，可以减少用户的意外损失。类图如图 4 所示。

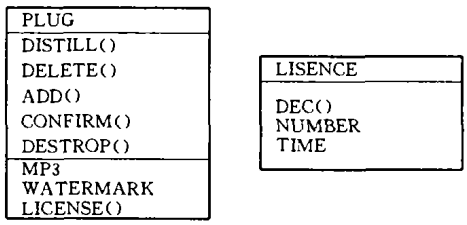

图 4　MP3 音乐插件类图

为了增强版权保护的可行性，可以对插件加密，达到反拷贝、防静态分析、反动态跟踪的基本要求。也可以将插件制作为不可拆卸的硬件——由可编程器件 DSP 来实现。

5　测试结果

小波域水印仿真实验结果如下：

a) 原始音频信号　　b) 嵌入 MP3 疫苗后的音频信号

图 5　测试 S 结果

由上图我们可以看出仿真实验的结果：原始音频信号（如图 a 所示）与嵌入 MP3 疫苗后的音频信号（如图 b 所示）之间的失真很小，能够很好地保护 MP3 的版权。

结论　将 MP3 疫苗合成在 MP3 音乐中，隐藏在计算机的插件中的数字细菌自动判别用户的 MP3 音乐是否是盗版。若是盗版，MP3 音乐中就无 MP3 疫苗，那么音乐就无法正常播放。即使用户是复制的带 MP3 疫苗的 MP3 音乐，那么插件是无法全部复制的，不全的插件也无法播放带 MP3 疫苗的 MP3 音乐，从而达到主动保护版权的目的。MP3 疫苗和插件的双重使用即使失效，MP3 疫苗中还有版权水印可以事后追踪消费者的盗版责任。不过插件的至少几秒钟的运行以及安装插件需要花一定的时间也会使消费者感到不快，但是在这方面的负面影响是很小的。

进一步的研究将包括数字细菌在其它方面的应用。MP3 疫苗可扩展为计算机疫苗用作其它用途，当然还可以产生各种各样的疫苗应用到其它方面。

参 考 文 献

1　Noel S，Szu H. Multimedia authenticity with ICA watermarks，Wavelet Applications VII，SPIE Proc.，Vol. 4056，April 2000. 175～184

2　Yim B，Noel S，Szu H. Detecting Electrocardiogram Abnormalities with Independent Component Analysis. In：15th Annual Intl. Symposium on Aerospace/Defense Sensing，Simulation，and Controls，Orlando，Florida，April 2001

3　Noel S，Szu H，Doppler Frequency Estimation with Wavelets and Neural Networks. presented at 12th Annual International Symposium on Aerospace/Defense Sensing，Simulation，and Controls，Orlando，Florida，April 1998

4　Szu H，Hsu C. Landsat spectral Unmixing à la superresolution of blind matrix inversion by constraint MaxEnt neural nets. In：Proc. SPIE 3078，1997. 147～160

5　Bassia P，Pitas I. Robust audio watermarking in the time domain. In：Proc. EUSIPCO 98，vol. 1，Rodos，Greece，Sept. 1998. 25～28

6　Gruhl D，Lu A，Bender W，Echo hiding. in Information Hiding，Springer Lecture Notes in Computer Science，1996，1174：295～315

7　Neubauer C，Herre J. Digital watermarking and its influence on audio quality In：Proc. 105th Convention，Audio Engineering Society，San Francisco，CA，Sept. 1998

8　Szu H，Progress in Unsupervised Atificial Neural Networks for Image Demixing Applications. IEEE Industrial Electronics SocietyNewsletter，46127，June 1999