

基于聚类分流算法的分布式蜜罐系统设计

柏青^{1*}, 苏阳^{1,2}

(1. 武警工程大学 电子技术系, 西安 710086; 2. 武警工程大学 网络与信息安全研究所, 西安 710086)

(* 通信作者电子邮箱 460545085@qq.com)

摘要:针对现有的网络安全防御系统主动性不足,对未知类型网络数据的判断速度慢、准确性不高的缺陷,设计了一种应用聚类算法对未知类型数据进行聚类分流的分布式蜜罐系统。在聚类过程中,采用一种改进的聚类中心选择算法,对未知类型网络数据进行模糊聚类,将聚类失败的数据分流到蜜罐中进行特征学习,从而尽早地发现新的攻击类型,减轻蜜罐的监控和记录压力,降低蜜罐被攻破的概率,有利于防御时采用更为有效的防御策略。此系统应用在某政府部门的专网中,实验结果验证了在无明显增加系统计算量的情况下,该聚类算法比平均值聚类算法有更高的聚类成功率。

关键词:蜜罐;聚类算法;初始类中心;数据分流;专网

中图分类号: TP393.08 **文献标志码:** A

Design of distributed honeypot system based on clustering and data shunting algorithm

BAI Qing^{1*}, SU Yang^{1,2}

(1. Department of Electronic Technology, Engineering University of Chinese Armed Police Force, Xi'an Shaanxi 710086, China;

2. Institute of Network and Information Security, Engineering University of Chinese Armed Police Force, Xi'an Shaanxi 710086, China)

Abstract: Concerning the lack of activity, the low speed and accuracy of recognizing attacks of the current network security defense system, this paper proposed a distributed honeypot system. During the process of clustering, an improved clustering center selection algorithm was used to cluster the data of the network in a fuzzy way, so as to divide the unclassified data into the honeypot to learn their features. Then a new type of attack can be detected as soon as possible. This design can not only lighten the supervising and recording pressure of honeypots, lower the broken rate of the honeypot, but also help us adopt more effective defense strategy. This system can be used in the private networks of some government. The clustering algorithm used in this paper has a higher rate of success than the average clustering algorithm without increasing the amount of computations of the system obviously.

Key words: honeypot; clustering algorithm; initial clustering center; data shunting; private network

0 引言

伴随着计算机和网络技术的不断发展,人们在享受其带来的高品质生活的同时,也遭受着各种安全问题的侵扰,网络攻击一刻也没有停过。传统的网络安全防御手段如防火墙、入侵检测系统(Intrusion Detection System, IDS)、身份认证等虽已得到广泛应用,但是面对日新月异的攻击方法,它们也只能被动“接招”。蜜罐技术的提出和应用,大大提高了网络安全系统的主动防御能力和对未知类型攻击的响应速度。

蜜罐(Honeypot)从本质上来说是一种资源,其本身价值是引诱攻击者对网络进行攻击。蜜罐能够实时监视攻击者的所有活动,诱惑或者欺骗入侵者,转移攻击目标使他们优先攻击蜜罐系统,并从捕获到的数据中学习入侵者使用的工具、方式和方法,从而赢得研究入侵对策的时间。

蜜罐按照交互能力主要分为低交互型和高交互型,前者的价值在于检测,为网络提供直接的安全保护;而后者的价值在于对恶意软件及黑客攻击的捕获和分析。蜜罐系统主要包括了网络诱骗、数据控制、数据捕获、数据报警、数据分析和日志远程存储等功能模块。由于蜜罐系统一般要结合防火墙、

路由器和IDS来进行数据控制,因此,整个蜜罐系统的数据包括防火墙日志、IDS网络数据包和日志、蜜罐本身日志等。数据报警通常借助一些监视工具(例如Swatch)来实现,通过字符串模式匹配去判断,然后实现E-mail等方式报警^[1]。

1 研究现状

蜜罐技术作为一种新型的网络安全技术,已经得到国外很多研究机构和公司的重视。目前该领域较大型的研究项目有:致力于部署分布式蜜罐的“分布式蜜罐项目组”(Distributed Honeypot Project);研究蜜罐网络诱骗技术的“蜜网项目组”(Honeynet Project);而“蜜网研究联盟”(Honeynet Research Alliance)在此领域有较大影响。“分布式蜜罐项目组”主要研究将蜜罐散布在网络的正常系统和资源中,利用闲置的服务器端口进行欺骗,将欺骗分布到更广范围的IP地址和端口空间中,从而增大欺骗在网络中的百分比。“蜜网项目组”致力于提高人们的网络安全意识,同时提供必要的网络安全知识及该组织开发的开源工具软件。“蜜网研究联盟”主要是通过使用蜜罐网络这样一个真实的环境来研究入侵者使用的工具、策略和动机,其所有的研究成果都是开放

收稿日期:2012-10-26;修回日期:2012-12-10。

基金项目:国家自然科学基金资助项目(61103231);陕西省自然科学基金基础研究计划项目(2012JM8014)。

作者简介:柏青(1988-),男,陕西宝鸡人,硕士研究生,主要研究方向:网络与信息安全;苏阳(1975-),男,河南三门峡人,教授,博士,主要研究方向:网络与信息安全、密码学。

的,向整个安全研究领域公布。目前蜜罐网络研究的重点被放在数据捕获和数据分析上,以提高蜜罐网络的易用性,即增加工具软件的界面友好性和操作的简便性等,并在此基础上提出了第三代蜜罐网络结构模型。

我国在蜜罐技术研究方面起步较晚,2001 年才对其立项研究。2004 年 9 月北京大学计算机研究所正式成立的蜜罐网络项目研究组,开展“狩猎女神”项目研究,该研究组于 2005 年 2 月 22 日正式成为“蜜网研究联盟”的一员。这是我国第一个正式研究蜜罐技术的组织,标志着我国蜜罐技术与世界接轨^[2]。

虽然如此,蜜罐技术在我国的应用现状仍不理想。它自身存在着几个局限性。首先是视野狭窄:蜜罐只能看见针对它自身的攻击行为进行记录和响应,不能对同一网络中其他部分的攻击作任何数据采集和分析。其次是蜜罐可能将新的风险引入到它所在的环境中:蜜罐系统一般设置较为简单,自身防护能力弱。一旦被黑客攻破,就可能被当作攻击和渗透其他系统的跳板。

本文采用一种改进的聚类算法,对在基于分布式蜜罐技术的系统中的经过防火墙和 IDS 检测后仍未被识别出类型的数据进行聚类分析,对聚类成功的数据通过其分簇情况来判断其所属类型,并采取相应措施。而对于聚类失败的数据,则重定向到蜜罐中,对其进行监视、特征提取。这样可以拓宽蜜罐监视的视野,接触到更多未知种类的攻击。而通过聚类算法,又可以对未知类型数据即疑似攻击数据做分析,将可以判断出类型的数据从疑似攻击数据中分离出来,以此减少进入蜜罐的数据量,降低蜜罐被攻破的概率,进而提高基于分布式蜜罐系统的网络的主动防御能力。

2 基于聚类分流技术的分布式蜜罐系统设计

2.1 聚类分流技术

聚类是将物理或者抽象对象的集合分成由类似的对象组成的多个类的过程。由聚类所生成的簇是一组数据对象的集合,这些对象与同一簇中的对象彼此相似,与其他簇中的对象相异^[3]。

本系统通过聚类算法对已捕获的未知种类的网络数据进行分簇;之后利用分流技术,对不同类型的数据分别进行相应的处理。

2.2 系统设计

网络中常见的攻击数据类型有四大类,分别为 PROBE(端口监视或扫描)、DoS(拒绝服务攻击)、U2R(未授权的本地超级用户特权访问)、R2L(来自远程主机的未授权访问),再加上 NORMAL(正常)数据,共有五类,所以本系统将数据聚类为五类^[4]。由于已捕获的数据经本系统采用的数据预处理方法整理后在格式上与 KDD CUP99 数据集相似,所以系统中各初始类中心的选定是通过 KDD CUP99 数据集一子集中已知种类的数据,结合初始类中心选择算法来实现的。在聚类过程中,利用这些得到的初始类中心对系统收集到的网络数据进行聚类。聚类成功的数据判定其种类,采取相应的防御措施,并将聚类成功的数据回送到选定的 KDD CUP99 子集中,对各初始类中心作动态调整。聚类失败的数据被重定向到蜜罐,对其进行深层次监控和研究。进行聚类分流,可以降低蜜罐被攻破的概率,减轻其监控和记录压力,同时也能缩短发现和掌握未知攻击方法的时间,从而及时调整防御策略。系统逻辑流程如图 1。

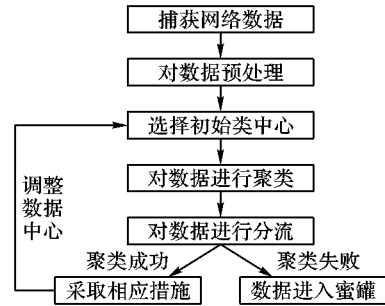


图 1 系统逻辑流程

2.2.1 数据捕获和预处理

蜜罐的主要目的之一就是获得有关攻击和攻击者的所有信息,捕捉入侵者从扫描、探测、攻击攻陷蜜罐主机,到最后离开蜜罐的每一步动作,因此,蜜罐必须有强大的信息捕获功能,在不被入侵者发现的情况下,捕获尽可能多的信息,包括输入/输出数据包和击键特征等^[5]。本系统采用四层数据捕获方式,分别由防火墙、IDS、路由器和蜜罐完成。IDS、防火墙和路由器捕获的数据为网络数据包,格式为:

(日期,时间,源 IP,源端口,目的 IP,目的端口,协议,头部信息,附加信息)

另外,可以通过事件查看器来查看审计日志中蜜罐捕获的数据,其格式为:

(日期,时间,类别,类型,事件标识符,来源,用户,主机名描述)

由于两种数据的格式不一样,需要先对其格式作调整:

1) 在网络层和传输层对网络数据包进行协议分析,分析 IP、TCP 和 UDP 协议头,并记录其特征,这样便可从网络数据包中找出源 IP 地址、源端口、目的 IP 地址、目的端口等信息,在规则提取时引用。对于蜜罐主机记录的攻击行为,可用作网络数据包属性的扩展,添加到网络包的属性中,比如文件生成操作的数目、错误登录次数等。

2) 将扩展的网络数据包整理成统一的格式,将各属性用逗号分隔。统一格式后的网络数据包的各属性如表 1 所示。

表 1 统一格式后的网络数据包属性情况

名称	类型	名称	类型
duration	连续	is_guest_login	离散
protocol_type	离散	count	连续
service	离散	srv_count	连续
flag	离散	serror_rate	连续
src_bytes	连续	srv_serror_rate	连续
dst_bytes	连续	rerror_rate	连续
land	离散	srv_rerror_rate	连续
wrong_fragment	连续	same_srv_rate	连续
urgent	连续	diff_srv_rate	连续
hot	连续	srv_diff_host_rate	连续
num_failed_logins	连续	dst_host_count	连续
logged_in	离散	dst_host_same_srv_rate	连续
num_compromised	连续	dst_host_diff_srv_rate	连续
root_shell	连续	dst_host_same_src_port_rate	连续
su_attempted	连续	dst_host_srv_diff_host_rate	连续
num_root	连续	dst_host_serror_rate	连续
num_file_creations	连续	dst_host_srv_serror_rate	连续
num_shells	连续	dst_host_rerror_rate	连续
num_access_files	连续	dst_host_srv_rerror_rate	连续
num_outbound_cmds	连续	dst_host_srv_rerror_rate	连续
is_host_login	离散		

3)整理好的数据包属性分为离散型和连续型,在聚类前应分别对两种类型的属性进行预处理。

①对于离散型属性对其统一进行0-1编码,使其变为连续型属性变量。

②连续型属性可通过中心化和标准化对其进行规范。

中心化 一般都是在观测值上减去相应变量的平均值。以属性 a 的 n 个数据实施中心化为例:

$$x'_{ia} = x_{ia} - \bar{x}_a; i = 1, 2, \dots, n$$

其中: x_{1a}, \dots, x_{na} 均为属性 a 的特征值; \bar{x}_a 是属性 a 的平均值,

$$\bar{x}_a = \frac{1}{n} \sum_{i=1}^n x_{ia}$$

标准化 在中心化的基础上,使各变量在相同范围内变化,用标准差标准化方法:

$$s_a = \sqrt{\frac{1}{n-1} \sum_{i=1}^n x'^2_{ia}}$$

标准化的特征属性值 $z_{ia} = (x_{ia} - \bar{x}_a) / s_a$ 。

这样便形成了具有标准属性的数据集,每条数据有41个特征属性。此外,要使用一台安全性高,不提供任何服务的系统来作为远程日志服务器,存储捕获到的网络数据。传输已捕获数据要使用专用线路,在传输前要对数据进行加密。

2.2.2 数据控制

数据控制是为防止蜜罐被攻破进而成为黑客攻击其他设备的跳板所必须采取的限制措施。如果黑客突破蜜罐,他们最需要的就是网络连接,以便从网络上下载攻击工具包,打开IRC(Internet Relay Chat)连接等。为了不引起黑客的怀疑,系统必须允许他们做大部分的“合法”事情,如蜜罐对各种服务的请求;但又要对攻击其他系统的“需求”一概禁止,如发起拒绝服务攻击、对外部扫描以及使用漏洞攻击他人的行为。本系统采用两种数据控制方法:第一种是通过设置防火墙外出连接次数来进行数据控制;第二种是路由控制,利用路由器的包过滤能力,限制连接,这时可以用没有网络地址的网关来代替路由器,更加隐蔽。

2.2.3 数据分析

从已捕获的数据中分析出黑客的攻击活动、所用方法和意图,对提高系统防御能力有着至关重要的作用。但是网络数据量极大,如果单纯依靠模式匹配等方法对每条数据进行甄别,再区分攻击数据和正常数据势必会降低网络工作效率^[6]。本系统首先通过防火墙和IDS过滤掉已知攻击,对剩余的未知种类的数据利用聚类算法,重新进行分类,之后进行分流,从而达到对数据的分析判断。

1)数据初始类中心选择算法。

聚类算法中最关键的部分是初始类中心的选择。在很多聚类算法中,为了降低计算量,初始类中心的选取是通过对数据各属性求平均值来确定的。这种方法容易淹没孤立点数据对初始类中心选取的影响,或者由于孤立点数据与初始类中心距离较远,而将孤立点作为初始类中心,进而造成聚类结果不准确^[7-8]。本文采用了一种改进的初始类中心选择算法。

先来定义四个相关概念。

① $CDT = (U, A, V, f)$ 是一个分类数据表, U 是对象的非空有限集合,称为论域。 A 是属性的非空子集。 $V = \bigcup_{a \in A} V_a$, V_a 是属性 a 的值域,并且是有限的。 $U \times A \rightarrow V$ 是一个信息函数,使得对于 $a \in A$ 和 $x \in U$,有 $f(x, a) \in V_a$ 。

② $d(x, y) = \sum_{a \in A} E_a(x, y)$ 称为对象 x 和 y 之间关于属性

集 A 的相异度,其中

$$E_a(x, y) = \begin{cases} 1, & f(x, a) \neq f(y, a) \\ 0, & f(x, a) = f(y, a) \end{cases}$$

③ $Dens_a(x) = |\{y \in U | f(x, a) = f(y, a)\}| / |U|$, 它是对象 x 关于属性 a 在论域 U 的密度。

④ $Dens(x)$ 是 x 关于属性 A 在论域 U 上的平均密度,且 $Dens(x) = \sum_{a \in A} Dens_a(x) / |A|$ 。由公式可看出, $Dens(x)$ 越大,则 x 周围聚集的对象越多, x 成为类中心的可能性越大,因此可以选择密度最大的对象作为第一个类中心。对于其余类中心的选择,不仅要考虑对象之间的距离,而且要考虑对象的平均密度。如果只考虑对象和已知类中心之间的距离,孤立点有可能被作为类中心。类似地,如果只考虑对象的密度,许多类中心可能位于一个类中心的周围。基于以上分析,下面对一种改进的初始类中心选择算法进行描述:

步骤1 输入 $CDT = (U, A, V, f)$ 和 k , 其中 k 是期望的类别数,令 $Centers$ 为空集。

步骤2 对于每一个 $x_i \in U$, 计算 $Dens(x_i)$, 令 $Centers = Centers \cup \{x_{i_1}\}$ 。其中 x_{i_1} 满足 $Dens(x_{i_1}) = \max_{i=1}^{|U|} \{Dens(x_i)\}$, 第一个类中心被选择。

步骤3 计算第二个类中心, 令 $Centers = Centers \cup \{x_{i_2}\}$ 。其中 x_{i_2} 满足 $d(x_{i_2}, x_m) \times Dens(x_{i_2}) = \max_{i=1}^{|U|} \{d(x_i, x_m) \times Dens(x_i) | x_m \in Centers\}$, 转至步骤4。

步骤4 如果 $|Centers| < k$, 则转至步骤5; 否则转至步骤6。

步骤5 对于任意一个 $x_i \in U$, $Centers = Centers \cup \{x_{i_3}\}$, 其中 x_{i_3} 满足 $d(x_{i_3}, x_m) \times Dens(x_{i_3}) = \max_{x_m \in Centers} \{ \min_{x_i \in U} \{d(x_i, x_m) \times Dens(x_i)\} | x_i \in U \}$, 则转至步骤4。

步骤6 结束, 输出 $Centers$ 。

2)数据聚类算法。

通过上面的算法选择出各初始类中心, 利用 k -modes 算法将一个 $n = |U|$ 个对象聚集成 k 类的目标是找到 W 和 C , $W = [w_{li}]$ 是一个 $k \times n$ 的矩阵, $C = [c_1, c_2, \dots, c_k]$, c_l 是第 l 个类中心。目的是使目标函数 $F(W, C)$ 最小。

$$F(W, C) = \sum_{l=1}^k \sum_{i=1}^n w_{li} d(c_l, x_i), \text{ 其满足最小化的条件是}$$

$$\textcircled{1} w_{li} \in \{0, 1\}, 1 \leq l \leq k, 1 \leq i \leq n; \sum_{l=1}^k w_{li} = 1, 1 \leq i \leq n;$$

$$\textcircled{2} 0 < \sum_{i=1}^n w_{li} < n; 1 \leq l \leq k。$$

聚类过程如下:

步骤1 选择初始点 $C^{(1)} \in \mathbf{R}^{|A|k}$, 计算 $W^{(1)}$, 使 $F(W^{(1)}, C^{(1)})$ 最小。

步骤2 设置标记 W 和 C 变化的变量 t , 并令 $t = 1$, 计算 $C^{(t+1)}$, 使得 $F(W^{(t)}, C^{(t+1)})$ 最小, 如果 $F(W^{(t)}, C^{(t+1)}) = F(W^{(t)}, C^{(t)})$, 停止; 否则转至步骤3。

步骤3 计算 $W^{(t+1)}$, 使得 $F(W^{(t+1)}, C^{(t+1)})$ 最小。如果 $F(W^{(t+1)}, C^{(t+1)}) = F(W^{(t)}, C^{(t+1)})$, 则停止; 否则执行 $t = t + 1$ 并转至步骤2。

步骤4 如果在执行了两次 $t = t + 1$ 后, 等式 $F(W^{(t)}, C^{(t+1)}) = F(W^{(t)}, C^{(t)})$ 仍无法成立, 则停止此次聚类, 并判定此数据无法聚类, 与其他数据分流后经重定向送入蜜罐。

3 分布式蜜罐系统在某政府某专网中的应用

3.1 政府专网介绍

随着计算机技术和网络技术的不断推广普及,政府单位也利用其提高自身工作效率,实现“无纸化”办公,节约资源。但是互联网上频繁的网络攻击对政府职能的发挥产生了巨大的负面影响。因此,自 1998 年,全国政府系统开始组建自己的专网,称为“全国政府系统办公业务资源网”(简称“政府专网”)。该网是政府内部的办公业务网络,与互联网物理隔离,满足政府系统内部办公、管理、协调、监督和决策的需要,并于 1999 年投入使用。虽然有了专网,避免了大量来自互联网的 attack,但是专网内部攻击依然存在。据统计,网络攻击中有 60% 是来自内部人员的攻击。加之一些部门的专网仍处于建设初期,网络防御能力薄弱,而网络攻击手段却不断地快速更新。因此需要将分布式蜜罐技术与传统的防火墙、IDS 相结合,对系统检测后无法判别其类型的网络数据,通过聚类算法快速进行分簇聚类,聚类成功的数据则判断其所属类型,聚类失败的数据则通过重定向技术将其送入蜜罐,进行更深入的监控、学习和特征提取,以达到学习掌握新攻击方法的目的。

3.2 基于分布式蜜罐系统的政府专网设计

由于政府单位职能业务不同,各部门专网结构也有所不同。下面以政府某部门的行政专网为例,将分布式蜜罐系统应用其中并进行研究。在此专网中部署蜜罐,构成一定规模的低交互度的分布式蜜罐网络,以此降低主机受扫描入侵的命中率,同时在蜜罐上放置蜂蜜信标来诱捕来自网络内部的非法访问。要实现攻击诱骗,需要有 IP 空间欺骗、网络流量仿真、系统动态配置、系统漏洞、网络服务、虚拟操作系统和端口重定向技术支持。专网具体结构如图 2。

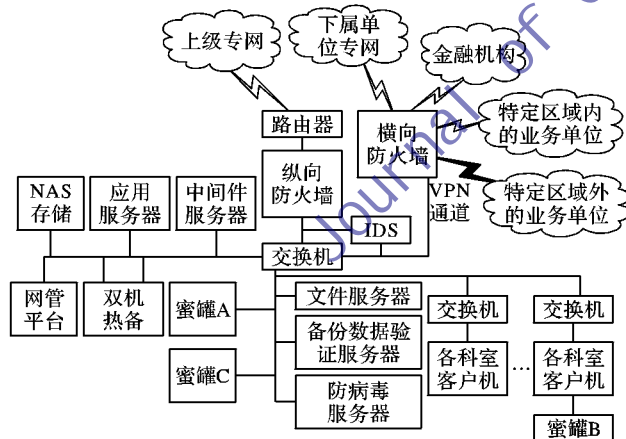


图2 模拟专网结构

系统中接入三个蜜罐,分别配置在服务器群和客户机之间,用以吸引绕过防火墙、IDS 的攻击和监督、记录经过聚类分流的未知类型数据,学习提取其特征。IDS 用于识别已知和未知类型的数据,并对未知类型数据进行记录。聚类分流器布置在 IDS 上,对未知类型数据进行第二次识别,之后根据聚类结果进行分流。

4 实验结果

为检验该系统所采用的聚类算法对未知类型数据的聚类效果,本文进行了一组对比实验。实验分为两部分,第一部分用 KDD CUP99 数据集中一子集,分别通过求平均值选择初始

类中心算法和本文采用的初始类中心选择算法确定初始类中心。第二部分进行了一组对比实验,用得到的两组初始类中心,分别对网络在一定时间内捕获的 492 条未知类型数据进行聚类分析。

为使选定的初始类中心更具有普适性,本文选取了一个包含 4000 条数据的子集,各类数据量见表 2。

表2 所选子集中各类型数据情况

数据类型	数据量
NORMAL	850
PROBE	300
DoS	2600
U2R	125
R2L	275

利用不同的初始类中心选择算法得到的聚类中心,对系统捕获的数据进行聚类,结果见表 3。

表3 平均值法和本文方法确定初始类中心得到的聚类结果

数据类型	平均值法		本文方法	
	数据量(条)	聚类结果	数据量(条)	聚类结果
NORMAL	5	成功	3	成功
PROBE	116	成功	125	成功
DoS	268	成功	257	成功
U2R	9	成功	16	成功
R2L	20	成功	40	成功
UNKNOWN	74	失败	51 失败	

聚类成功率 = $1 - \text{聚类失败数据量} / \text{聚类总数据量}$ 。由于是对数据进行第二次识别,聚类成功率普遍不高。平均值法的聚类成功率为 84.96%,改进的初始类中心选择算法聚类成功率为 89.63%。从表中可以看出,平均值法虽然在 NORMAL 和 DoS 两类数据的聚类上效果较好,但聚类算法属于模糊判别理论,并不能完全确定已经聚类成功的数据的真实类型,并且平均值法是通过求各条数据属性的平均值来确定初始类中心。该中心对数据的共性特征体现较明显,但个性特征显示不足,会造成将一些特征接近正常数据的攻击数据误判为正常数据,进而导致表面提高了聚类成功率,实际降低了聚类准确率情况的发生。实验结果显示平均值法的最终聚类成功率较低,对不常见的攻击如 R2L 等,识别能力不足。本文采用的初始类中心选择算法在聚类成功率和识别不常见攻击方面效率较高。针对其 NORMAL 和 DoS 数据聚类效果欠佳,可以通过增大确定初始类中心时选取的子集中相应类型数据量来改进。另外,采用平均值法的聚类时间为 5.6598 s,而采用本文改进的初始类中心选择算法的聚类时间为 5.7359 s。在没有明显增加系统计算量的情况下,第二种算法可以有效提高聚类成功率。

5 结语

本文将分布式蜜罐系统与防火墙、IDS 相结合,通过一种改进的初始类中心选择算法的聚类方式,实现了对未知种类数据的识别及分流功能。将其应用到政府某专网中提高了专网识别和响应攻击的整体速度,缩短了安全策略的调整时间;同时降低了蜜罐被攻破的概率,增强了专网的主动防御能力。在下一步的研究中,应该重点改进现有初始类中心选择算法,使其具有更多的初始类中心,进而提高聚类精确度和成功率。

(下转第 1084 页)

的空间占用比 RC5 多 4198 字节、比 RC6 多 3596 字节。其主要原因是该算法需要存储 2 个的 S 盒,以及 2 个置换矩阵。

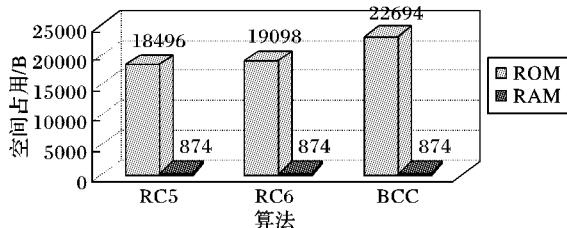


图8 算法占用空间比较

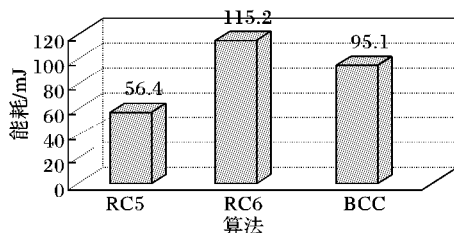


图9 算法在 WSN 中的能耗仿真

由图9 仿真结果,可看出 RC5 算法的能量消耗相比 RC6、BCC 算法具有明显优势。BCC 算法比 RC6 算法能耗稍低。

对于 WSN 在考虑加密算法时,需要从安全性、空间占用、耗能等进行考虑,安全性越高,算法越复杂,空间占用、能量消耗相应会增加。

5 结语

在实际应用中需要对各种参数进行综合考虑,寻找平衡点。BCC 算法正是基于这种观点,利用空间消耗来换取安全性的提高和能量消耗的降低。通过对 BCC 算法与 RC5、RC6 算法进行了对比分析,结果显示 BCC 算法除了占用空间多,其他的安全指标和性能指标具有一定的综合优势。随着存储器件容量不断提高,体积不断减小,价格不断下降,BCC 算法将会在 WSN 中得到较好的应用。

参考文献:

- [1] 裴庆祺,沈玉龙,马建峰.无线传感器网络安全技术综述[J].通信学报,2007,28(8):113-122.
- [2] LAW Y W, DOUMEN J, HARTEL P. Survey and benchmark of block ciphers for wireless sensor network [J]. ACM Transactions on Sensor Network, 2006, 2(1):65-93.
- [3] GUIMARAES G, SOUTO E, SADOK D, *et al.* Evaluation of security mechanisms in wireless sensor network[C]// Proceeding of the 2005 Systems Communications. Washington, DC: IEEE Computer Society, 2005: 428-433.
- [4] 吴文玲,冯登国,张文涛.分组密码的设计与分析[M].北京:清华大学出版社,2009:64-84.
- [5] 廖晓峰,肖迪,陈勇,等.混沌密码学原理及其应用[M].北京:科学出版社,2009:55.
- [6] CHEN Z, ZHANG Z W, JIANG N. A session key generator based on chaotic sequence[C]// International Conference on Computer Science and Software Engineering. Washington, DC: IEEE Computer Society, 2008:635-637.
- [7] PENG J, JIN S Z, LIN H L, *et al.* A block cipher based on a hybrid of chaotic system and Feistel network[C]// 2009 Fifth International Conference on Natural Computation. Piscataway, NJ: IEEE Press, 2009:17.
- [8] 邓绍江,黄桂超,陈志建.基于混沌映射的自适应图像加密算法[J].计算机应用,2011,31(6):1502-1504.
- [9] 肖迪,赵秋乐.一种基于 Logistic 混沌序列的图像置乱算法的安全分析[J].计算机应用,2010,30(7):1815-1817.
- [10] LIANG W, XU J B, TANG M D, *et al.* A new embedded encryption algorithm for wireless sensor networks[C]// IFITA'09: Proceedings of the 2009 International Forum on Information Technology and Applications. Washington, DC: IEEE Computer Society, 2009:119-122.
- [11] ULUAGAC A S, BEYAH R A, LI YINGSHU, *et al.* WEBEK: virtual energy-based encryption and keying for wireless sensor networks[J]. IEEE Transactions on Mobile Computing, 2010, 9(7):994-1007.
- [12] MASUDA N, JAKIMOSKI G, AIHARA K, *et al.* Chaotic block ciphers: from theory to practical algorithms[J]. IEEE Transactions on Circuits and Systems I: Regular Papers, 2006,53(6):1341-1352.
- [13] 陈帅.网络混沌加密理论及其关键技术研究[D].重庆:重庆大学,2006.
- [14] TANG G P, LIAO X F, CHEN Y. A novel method for designing S-boxes based on chaotic maps[J]. Chaos, Solitons & Fractals, 2005, 23(2):413-419.
- [15] 杨吉云.混沌密码在无线传感器网络安全中的应用研究[D].重庆:重庆大学,2007.

(上接第1080页)

参考文献:

- [1] 田俊峰,刘永立.一种新的蜜网模型——BRHNS[J].计算机工程与应用,2007,43(7):139-143.
- [2] 尚治宇.蜜罐系统的研究与实现[D].北京:北京交通大学,2010.
- [3] 曹付元.面向分类数据的聚类算法研究[D].太原:山西大学,2010.
- [4] KDD99. KDD99 cupdataset[EB/OL]. [2005-11-02]. <http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html>.
- [5] Intrusion detection working group. Intrusion detection message exchange format data model and EXtensible Markup Language (XML) document type definition[R]. Internet-Draft,2003:21-26.
- [6] 朱明.数据挖掘[M].合肥:中国科技大学出版社,2002.
- [7] 孙吉贵,刘杰,赵连宇.聚类算法研究[J].软件学报,2008,19(1):45-61.
- [8] RALAMBONDRAIN H. A conceptual version of the k-means algorithm[J]. Pattern Recognition Letters, 1995,16(11):1147-1157.
- [9] 肖宇,于剑.基于近邻传播算法的半监督聚类[J].软件学报,2008,19(11):2803-2813.
- [10] JIANG R, LIU Q, LIU Q, *et al.* A proposal for the morphological classification and nomenclature of neurons [J]. Neural Regeneration Research, 2011,6(25):1925-1930.
- [11] XU R, WUNSCH D. Clustering[M]. Hoboken, New Jersey, USA: John Wiley & Sons, 2009.
- [12] 杨博,刘大有,LIU J M,金弟,等.复杂网络聚类方法[J].软件学报,2009,20(1):54-66.
- [13] ZHANG S H, WANG R S, ZHANG X S. Identification of overlapping community structure in complex networks using fuzzy c-means clustering[J]. Physica A, 2007,374(1):483-490.
- [14] 石立宪.网络陷阱的研究与实现[D].大连:大连海事大学,2006.