面向移动传感网的攻击风险监测系统

倪 铭 张 宏 李千目

(南京理工大学计算机科学与工程学院 南京 210094)

摘 要 提出基于稳定性、安全性、簇大小选择和簇首节点合理选择的方法,从逻辑上对整个网络进行了簇的划分并选出最佳簇首节点,引入基于自回归模型的间隙性异常识别,实现了在簇首节点上执行自动检测网络流量异常和自动报警的功能。实验证明了该方法的有效性。

关键词 攻击风险监测,簇移动传感网,自回归

中图法分类号 TP393 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2015.3.030

Security Risk Monitor System of Mobile Sensor Network

NI Ming ZHANG Hong LI Qian-mu

(College of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China)

Abstract This paper proposed an approach based on stability, security, cluster size selection and rational cluster head node choice, which divides the entire network into clusters logically and selects the best head node. With a intermittent anomaly identification based on the auto-regressive (AR) model, it achieves automatic detection of network traffic anomalies on the head node of the cluster and automatic alarm. Experiments show the effectiveness of the proposed method.

Keywords Monitoring program on risk of attack, Cluster, Mobile sensor network, Auto-regressive

1 引言

无线信道、动态拓扑、合作的路由算法、缺乏集中的监控等特点使得移动传感网的安全性非常脆弱。目前,移动传感网中攻击风险监测技术的研究尚处于起步阶段。文献[1]采用RIPPER分类算法,设计了一个对路由表的异常更新进行静态检测的检测模型;文献[2]在文献[1]的研究基础上引入了代理技术;文献[3]综合运用了安全路由和攻击风险监测技术进行攻击风险监测。上述攻击风险监测技术应用于移动传感网中还存在问题:首先,资源的有限性要求攻击风险监测的部署要轻负载;第二,节点生命周期短并处于不断移动中,知识库学习和更新困难;第三,网络拓扑的改变使得很难判别一个提供了错误消息的节点是Byzantine式还是仅仅失去同步。

本文在现有的分布式协作攻击风险监测方案的基础上,提出一些改进来实现有效和可靠的攻击风险监测。

2 现有攻击风险监测方案分析

为了保障移动传感网的安全,已有许多安全解决方案[4]被提出,主要集中于密钥的分配与认证[5]和路由安全算法两个方面[6,7]。这些措施用于移动传感网之中,能够发挥安全防范作用,但如果节点被截获或泄露密钥,这些技术都会遭受风险。因此,研究攻击风险监测是极有意义的。表1列出了典型的移动网络攻击风险监测方案。

表1 典型的攻击风险监测方案分析

提出人	检测模式	决策方式	工作特点	优点	缺点
NASA	异常检测	协作	每个节点本地监控本地用户活动,若确定检测到入侵行为,可独立响应。若不能确定是否是入侵行为,向相邻节点发送入侵状态询问触发合作检测流程	各Agent合作响应,能适应不同审计日志	对拜占庭节点脆弱,占用资源较多
Oleg Kachirski Ratan Guha	异常检测	独立	在NASA方案的基础上,只在某些节点上驻留监视网络的Agent,并且Agent的数量可按要求进行增减	具有较好的扩展性和灵活性;大多数节点保存资源	MA的安全是最大的难题,协议复杂

到稿日期:2014-08-20 返修日期:2014-10-27 本文受江苏省自然科学重大研究项目:无线传感网的安全感知与全域安全控制(BK2011023),江苏省自然科学基金:无线传感网安全风险的全域感知、评估与控制(BK2011370),国家自然科学基金:信息系统的安全风险跟踪评价与动态预测(6090302)资助。

倪 铭(1989—),男,博士生,主要研究方向为无线网络通信技术、信息安全,E-mail:nq1027@gmail.com;张 宏(1958—),男,博士,教授,博士生导师,主要研究方向为无线通信技术、无线传感网络、网络安全等;李千目(1979—),男,博士后,教授,博士生导师,主要研究方向为无线网络通信技术、信息安全、高性能无线通信等。

（续表）

提出人	检测模式	决策方式	工作特点	优点	缺点
Huang,Lee	异常检测	独立	将攻击风险监测限定在某一些节点来降低SRM资源使用，从而提高MANET SRM的有效性	能动态调整Agent的数量，提高网络资源和CPU的效率	需组织被泄露的节点作为簇头，没提及误警率
Chin Yang Tseng	基于规范的检测	合作	利用分布在网络中的监测节点，合作监视在AODV路由查询过程中被监视节点是否按路由规范进行操作	不需要训练数据，较低的管理费用	基于一种路由协议设计，通用性差
Sun,Wu,Pooch	异常检测	分布协作	用于攻击风险监测的节点间的通信协议和区域建立算法都较为复杂，所以该方法占用较大数量的资源	较低的误警率，能合并多检测技术	复杂结构需要多协议合并，建区计算复杂
Guha,Schwartz	误用检测	分布协作	自动构造异常检测	考虑了MANET的分布式特性，较好的带宽意识	MA的安全性问题，丢包率较高
Huang,Lee,Yu	异常检测	独立	基于丢包和路由表异常变化，自动构造异常检测模型	能自动地构造异常检测模型	较高的计算开销
Albers,Camp	异常检测	独立	基于Agent的可伸缩体系结构，如果SRM需要更多的功能，它只需合并更多的移动Agent，同时它也能减少SRM网络流量	伸缩性强，较小的网络流量	创建和管理MA较为复杂
Rajavar,Shah	异常检测	分布协作	在每一跳都要检测数据包	不同于传统的本地攻击风险监测机制	Packet需在每一跳检测，效率较低
Pttini,Percher	异常检测	分布协作	基于管理信息库(MIB)，执行多级检测攻击	分布性和伸缩性，能检测多级入侵	只适合于MA，存在安全、计算花费和管理问题

分析表1可以看出：(1)根据移动传感网络独有的特点，设计一个具有自组织和移动特性的无线网络SRM(Security Risk Monitor)体系结构是一个重要方向；(2)移动传感网中没有统一的监测点，任何节点收集到的信息都是不完全的，导致审计日志的部分性和局部性的特点，因此，需要一种能有效处理审计日志数据的方法；(3)需要一种能从正常数据流中有效地鉴别出攻击数据流的方法。

3 自适应移动传感网攻击风险监测

本文思想是：逻辑上对整个网络进行簇划分，选择最佳簇首节点。簇首节点执行网络检测，普通节点执行本地检测。在簇首节点上应用基于自回归(AR)模型的间隙性异常识别算法，对所考虑的时间窗建立拟合AR模型，然后用所得残差计算统计量，以该统计量为判断异常的依据；计算得到的统计量是当前时刻的残差与滑动窗口的相关量(σ)之间的比较结果，当前观测值是否异常可以立即判断出来。因此，本方法是一种实时的检测方法，即每个异常点都能被检测到。

阶段一：簇与簇首节点的选择

假设：(a)无线信号传播符合自由空间传播模型[8]，采用文献[9]提出的预测方法，假定该模型中接收到的信号强度主要由接收方和发送方的距离来决定；(b)节点有效发射距离相等，若两个节点处在彼此的发射范围内，称这两个节点为邻居节点，且存在一条链路相连；(c)在每个移动节点上安装GPS，获取每个节点的移动速度、方向和方位。本文仅考虑一跳的情况，因为过多跳数会导致簇首节点监视不到一部分数据包，从而引起决策失误。

设节点 i 的邻居节点数为 n_i，速度为 v_i，位置坐标为 (x_i, y_i)，移动方向为 $\theta_i(0\leqslant\theta_i\leqslant 2\pi)$，信号传播距离为 s，p_i 为簇首节点标识，id_i 为节点的惟一标识，N_{max} 为簇允许的最大普通节点数，T_{min} 为允许的节点 i 与节点 j 的最短连接时间。对簇及簇首节点的选择按照如下步骤执行：

步骤1：任一节点 i 将 n_i、v_i、(x_i, y_i)、id_i 和 θ_i 发送给自己的邻居 j。

步骤2：任一节点 i 依据所获得的数据预测到自己邻居 j 的连接时间，预测公式为：

$$C_{i,j}=\frac{(ab+cd)+\sqrt{(a^2+c^2)s^2-(ad-bc)^2}}{a^2+c^2}$$

其中，$a=v_i\cos\theta_i-v_j\cos\theta_j$；$b=x_i-x_j$；$c=v_i\sin\theta_i-v_j\sin\theta_j$；$d=y_i-y_j$。

步骤3：节点 i 选取最大的且没有被选取过的 n_j，并验证 $n_j\geqslant n_i$ 是否成立；若成立，则执行步骤4；否则不向其他节点投票。

步骤4：节点 i 检查 $C_{i,j}\geqslant T_{min}$ 是否成立；若成立，则向节点 j 投票；否则执行步骤3。

步骤5：若节点 i 收到一个投票，则置 $p_i=1$；同时检查投票数是否大于 N_{max}，如果大于，即从投票的节点中选出与其连接时间较短的邻居节点 j，并向 j 发送拒绝信号，让 j 重新选择簇首节点；如果节点 i 没有收到拒绝信号，则将向其投票的节点 j 作为自己的簇首节点；否则执行步骤3、4，重新选择簇首节点。

阶段二：自适应检测

移动传感网的网络流量在短时间内会产生突发性变化，这在统计上是不平稳的，因此需要对网络流量进行适当处理来达到或近似稳定。由此，采用方差分析的方法(ANOVA)对实际网络流量进行预处理。

(a)数据处理：作为设计和分析的基础，共收集6周的网络流量数据，每周5个工作日，每天24小时，采集数据时间间隔为5min，以周为单位，对这6周的数据取平均值；同时考虑不确定因素的影响，需要删除一些坏值，本文根据格拉布斯准则剔除坏值：

$$\bar{x}=\frac{1}{6}\sum_{i=1}^{6}x_i,\ \mu=\sqrt{\frac{1}{6}\sum_{i=1}^{6}(x_i-\bar{x})^2}$$

其中，$\bar{x}$ 为 $x_1, x_2, \cdots, x_6$ 的平均值，μ 为标准差，若 x_i 满足式 $|x_i|>k\mu$，则 x_i 为坏值，将其删除，其中 k 表示格拉布斯准则系数，置信区间达95%时，$k=2.03$。对 $x_1, x_2, \cdots, x_6$ 中剩余值求平均，由此得到6周内1200个数据的平均值，记为每个实际观测值 M_{ij} ($i=1,2,\cdots,240$；$j=1,2,3,4,5$)，其表示为该流量参数在一周的第 j 个工作日第 i 个时刻的正常行为模

式。若当前观测值与其相比明显偏离，则判其为异常。

记 A 表示一周总平均观测值，A_i' 表示每天第 i 个时刻的平均值与总平均值 A 的偏差，A_j'' 表示每周第 j 天的平均值与总平均值 A 的偏差，如下：

$$A=\frac{\sum_{i=1}^{5}\sum_{i=1}^{240}M_{ij}}{5\times240},A_i'=\frac{\sum_{j=1}^{5}M_{ij}}{5}-A,A_j''=\frac{\sum_{i=1}^{240}M_{ij}}{240}-A$$

同时，$\sum_i A_i'=0,\sum_j A_j''=0$。

将每个实际观测值 M_{ij} 划分为 4 个部分：

$$M_{ij}=A+A_i'+A_j''+y_{ij}$$

可变换为：

$$y_{ij}=M_{ij}-A-A_i'+A_j''$$

经过转换，把原始观测值序列转换成 y_{ij}。通常，在不同工作日以及同一天的不同时刻，网络变动情况是不同的。经过上述处理，可把一周当中不同的工作日以及每天的不同时刻这两个影响因素从原始观测值 M_{ij} 中删除。同样，对于这些节点经常变动的移动网络，还可以用类似的方法剔除不同月份甚至不同年份的影响。为表示方便，把转换后的序列 y_{ij} 按时间先后顺序排列，记成$[y_t]$，$t=1,2,\cdots$；实际运用时，每周更新一次 $A,A_i'(i=1,2,\cdots,240),A_j''(j=1,2,3,4,5)$。

从总体上说，网络流量的观测值序列是不平稳的，但可在局部将其近似统计为平稳的，这个局部记作一个滑动时间窗，窗的大小记为 $N+1$，每次取出 $N+1$ 个数。滑动窗的 $N+1$ 个数可表示为 $y_1,y_2,\cdots,y_N,y_{N+1}$，前 N 个数用于建立自回归(AR)模型，可判断第 $N+1$ 个数是否为异常(即在实际应用时，将时间窗口不断向前滑动一次)。

(b)建立拟合 AR 模型

为建立拟合 AR 模型，必须零均值化前 N 个数，设 $\bar{y}$ 为 $y_1,y_2,\cdots,y_N,y_{N+1}$ 的平均值，即

$$\bar{y}=\frac{1}{N}\sum_{t=1}^{N}y_t,x_t=y_t-\bar{y},t=1,2,\cdots,N,N+1$$

则 $x_1,x_2,\cdots,x_N,x_{N+1}$ 为零均值时间序列。一般情况下，观测值序列$[x_t]$($t=1,2,3\cdots$)是不平稳的，但可以假定滑动时间窗近似地平稳，所以窗口的大小 N 不应该过大。拟合模型时，选择合适的模型阶数 p，自回归模型 AR(p)拟合时间序列时，可采用 Akaike 的 FPE(Final Prediction Error)衡量其准确性，相应的最小 FPE 的 AR 的阶数 p 为最佳模型阶数，但是 N 与 p 必须满足约束条件：$0\leqslant p\leqslant 0.1N$。

用时间序列 $x_1,x_2,\cdots,x_N$ 估计 AR(2)模型参数 ϕ_1 和 ϕ_2 以及 σ_e^2，过程如下：

$$Y=\begin{bmatrix}x_3\\x_4\\\vdots\\x_N\end{bmatrix},X=\begin{bmatrix}x_2&x_1\\x_3&x_2\\\vdots&\vdots\\x_{N-1}&x_{N-2}\end{bmatrix},\Phi=[\phi_1,\phi_2]$$

记 T 表示矩阵转置，系数 ϕ_i 的估计式为：$\Phi=[X^TX]^{-1}X^TY$，白噪声 e_t 的方差 σ_e^2 为 $\bar{\sigma}_e^2=\frac{1}{(N-1)}\sum_3^N(x_t-\bar{\phi}_1x_{t-1}-\bar{\phi}_2x_{t-2})^2$。

$$X^TX=\begin{bmatrix}\sum_2^{N-1}x_t^2&\sum_2^{N-1}x_tx_{t-1}\\\sum_2^{N-1}x_tx_{t-1}&\sum_1^{N-2}x_t^2\end{bmatrix},X^TY=\begin{bmatrix}\sum_3^{N}x_tx_{t-1}\\\sum_3^{N}x_tx_{t-2}\end{bmatrix}$$

得

$$\begin{bmatrix}\bar{\phi}_1\\\bar{\phi}_2\end{bmatrix}=[X^TX]^{-1}X^TY \quad (1)$$

$$\bar{\sigma}_e^2=\frac{1}{(N-2)}\sum_3^N(x_t-\bar{\phi}_1x_{t-1}-\bar{\phi}_2x_{t-2})^2 \quad (2)$$

由此，得到 AR(2)参数。

(c)执行异常检测

利用 AR(2)模型执行异常检测，AR(2)模型为：

$$x_t=\phi_1x_{t-1}+\phi_2x_{t-2}+e_t \quad (3)$$

变换为

$$e_t=x_t-\phi_1x_{t-1}-\phi_2x_{t-2} \quad (4)$$

若 Δ 是一步后移算子，即 $x_{t-1}=\Delta x_t$，则

$$\begin{aligned}e_t&=x_t-\phi_1\Delta x_t-\phi_2\Delta x_{t-1}\\&=x_t-\phi_1\Delta x_t-\phi_2\Delta^2x_t\\&=(1-\phi_1\Delta-\phi_2\Delta^2)x_t\\&=\phi(\Delta)x_t\end{aligned} \quad (5)$$

其中，$\phi(\Delta)=1-\phi_1\Delta-\phi_2\Delta^2$，由此

$$e_t=\phi(\Delta)x_t,t=1,2,\cdots,N \quad (6)$$

令 $\sigma^2=N^{-1}\sum e_t^2$，定义 $e_{N+1}=\phi(\Delta)x_{N+1}$ 及 $\psi=\frac{e_{N+1}}{\sigma}$。

其中，σ^2 表示时间序列中当前时刻向后 N 个相应残差 e_t 平方和的平均值。

ψ 表示当前观测值的残差与 σ 比值，则 ψ 作为检测 x_{N+1} 是否异常的统计量。统计量 ψ 的大小标志着异常点偏离正常的大小。设 Z_1 和 Z_2 为两个大于零的常数：(i)当 $\psi<-Z_1$ 或 $\psi>Z_2$ 时，x_{N+1} 为异常值；(ii)$\psi<-Z_1$ 时，异常值比正常值小，ψ 越小，则异常值偏离正常范围越大；(iii)当 $\psi>Z_2$ 时，异常值比正常值大，ψ 越大，则异常值偏离正常范围越大。

(d)执行异常报警

当实时读取参数时，若某个性能参数被检测出异常，则发出参数级警报。在使用异常检测方法时，预先对设定的常数 Z_1 和 Z_2 进行适当的调整，可以防止出现过多的假警报。为了保证检测率，可根据统计量 ψ 的不同大小，把异常情况分成不同等级：设 Z_{20}、Z_{21}、Z_{22} 满足 $0<Z_{20}<Z_{21}<Z_{22}$，Z_{10}、Z_{11}、Z_{12} 满足 $Z_{12}<Z_{11}<Z_{10}<0$。当 $Z_{20}<\psi\leqslant Z_{21}$ 或 $Z_{11}\leqslant\psi<Z_{10}$ 时，发出较低级警报；当 $Z_{21}<\psi\leqslant Z_{22}$ 或 $Z_{12}\leqslant\psi<Z_{11}$ 时，发出警报；当 $Z_{22}\leqslant\psi$ 或 $\psi\leqslant Z_{12}$ 时，发出高等级警报。

本文的异常检测方法特别适用于间歇性的、短时间内异常的检测，由于实现了实时对网络流量的检测，因此算法的复杂性是必须考虑的因素。从第二步建立拟合 AR 模型开始估算，按照 $N=30$ 估算，零均值化的计算量很小(主要是有加减运算 46 次)，在异常检测阶段，乘除运算约有 112 次，加减运算约有 96 次。经测算，在 Pentium4 1.8GHz、512MB RAM 的机器上检测一次约需 30μs，所以在现有的机器上进行实时检测是可行的。

4 仿真实验

本文的仿真环境是 NS2；节点移动范围：1500m×300m，节点传输距离为 250m。取网络节点数 $N=50$，移动速度分别为 0m/s，2m/s，…，20m/s 时，对算法进行仿真，其结果如图 1 所示。从图 1(a)可看出，簇首节点监测到网络中的报文比率基本不受移动速度的影响，均达到 96%以上，其原因是簇首节点分布均匀，且有一定冗余。从图 1(b)中可看出，簇首节点占网络节点的百分比会随移动速度增加而略有增大，当移动速度达到 20m/s 时，其比率还控制在 40%以下，其原因是随着节点移动速度的增大，节点可能移出原簇而成为新的簇首节点，网络中的簇首节点数目有所增加。

当网络节点数分别取为 40、60、80、100，移动速度为 10m/s 时，进行算法的仿真，结果如图 1(c)所示。可以看出，簇首节点监测到网络中的报文比率基本不受网络规模的影响，均达到 97%以上。

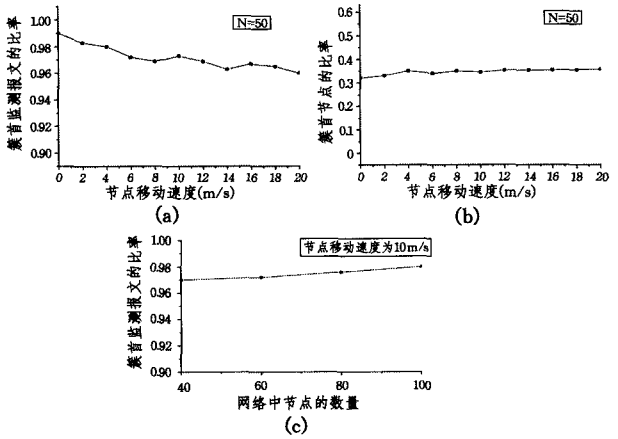

图 1　仿真结果

为了衡量检测效果，定义检测率的概念：在正确检测出真实异常点的情况下(这可以通过调整阈值达到)，它等于序列中正确检测异常点数与发生异常事件总数之比，显然检测率越高，效果越好，它的最大值是 1。图 2 和图 3 绘出了当节点在一维直线移动时，检测率随 IpForwDatagrams 和 IpInReceives 的观测值序列增加时的变化情况，同时也比较了网络节点数目不同时网络的异常检测情况。从图中观察得知，在增加网络节点时，网络流量变化变得越来越陡，即在网络节点数目增多时，网络流量变化会出现跃变现象；当检测率超过 0.5 时，异常点误检率(FAR=误检异常点个数/参与检测点总数)呈下降趋势。

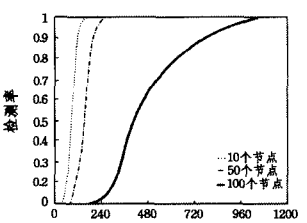

图 2　一维 IpForwDatagrams 观测值序列的检测结果

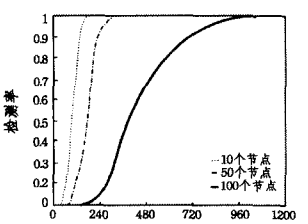

图 3　一维 IpInReceives 观测值序列的检测结果

图 4 和图 5 显示了节点在二维平面移动时有与图 2 和图 3 中相同的变化趋势；同时，当检测率超过 0.5 时，异常点误检率也呈下降趋势，如表 2 所列。

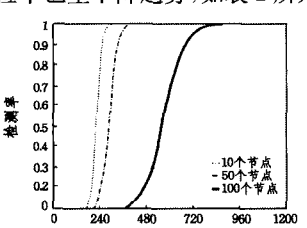

图 4　二维 IpForwDatagrams 观测值序列的检测结果

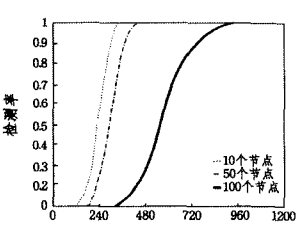

图 5　二维 IpInReceives 观测值序列的检测结果

表 2　两种移动方式的异常点误检率

节点数	10	50	100
一维直线移动	0.2306	0.0783	0.0462
二维直线移动	0.2705	0.0858	0.0493

结束语　本文研究分析了传统攻击风险监测技术，比较了传统攻击风险监测技术的差异。在改进现有的基于分布式协作攻击风险监测方案的基础上，提出了一种自适应移动传感网攻击风险监测方案。

参 考 文 献

[1] Zhang Yong-guang, et al. Intrusion Detection in Wireless Ad hoc Networks[C]// Proc. of the 6th Annual ACM/ IEEE International Conference on Mobile Computing and Networking. Boston: ACM Press, 2000: 275-283

[2] Kachirski O, Guha R. Effective Intrusion Detection Using Multiple Sensors in Wireless Sensor Networks [C]// Proc. of the 36th Hawaii International Conference on System Science. Hawaii, 2003: 102-110

[3] Venkatraman L., Aagrawal D P. Strategies for Enhancing Routing Security in Protocols for Mobile Ad hoc Networks[J]. Journal of Parallel and Distributed Computing, 2003, 63(2): 214-227

[4] Gomez J, Campbell A T. PARO: Supporting dynamic power controlled routing in wireless Ad Hoc networks[J]. ACM/Kluwer J. Wireless Networks, 2003, 9(5): 443-460

[5] Toh C-K. *Maximum battery life routing to support ubiquitous* mobile computing in wireless Ad hoc networks[J]. IEEE Communicating Magazine, 2001(6): 2-11

[6] Rodoplu V, Meng T. Minimum energy mobile wireless networks [J]. IEEE Journal on Selected Areas in Communications, 1999, 17(8): 1333-1344

[7] Wattenhofer T. Distributed topology control for power efficient operation in multihop Ad Hoc networks[C]// Proc. IEEE INFOCM Conf.. 2001: 1388-1397

[8] Rappaport S. Wireless communications: principles and practice [M]// Upper Saddle River. NJ: Prentice Hall, 1995: 72-75

[9] Bae S H, Lee S J, Su W, et al. The design, implementation and performance evaluation of the on demand multicast routing protocols in multihop wireless networks[J]. IEEE Network, 2000, 14(1): 70-77

[10] Li Qian-mu. Multiple QoS Constraints Finding Paths Algorithm in TMN[J]. Journal of Information, 2011, 14(3): 223-231

[11] Li Qian-mu, Yin Jie. A Service Access Security Control Model in Cyberspace[C]// Intelligent Computing and Information Science (ICICIS2011). 2011, 134: 418-426

[12] 王芳，易平，等. 基于规范的移动 Ad hoc 网络分布式入侵检测[J]. 计算机科学，2010，37(10)：118-122