

无线网络中数据伪造攻击检测方法研究

马小雨

(河南工程学院 计算机学院, 河南 郑州 451191)

摘 要: 针对传统的检测方法一直存在检测误差率大, 能耗多的问题, 提出基于主成分分析方法与平均评分偏离度计算法结合的无线网络中数据伪造攻击检测方法. 采用函数极值法, 确定隶属度, 获取无线网络数据伪造攻击特征. 采用统计学分析法, 确定无线网络数据伪造攻击分类函数, 建立无线网络中数据伪造攻击模型, 采用主成分分析方法对初始无线网络数据伪造攻击的协方差矩阵进行特征值分解, 引入平均评分偏离度计算法, 确定无线网络数据伪造攻击偏离度, 实现无线网络中数据伪造攻击的检测. 实验结果表明, 采用改进检测方法时, 其检测能耗及误检率方面在一定程度上, 相比传统的检测方法要好, 具有一定的优势.

关键词: 无线网络; 数据; 伪造攻击; 检测

中图分类号: TN919

文献标识码: A

文章编号: 1000-7180(2017)02-0105-04

Data Forgery Attack Detection Method in Wireless Network

MA Xiao-yu

(Department of Computer, Henan University of Engineering, Zhengzhou 451191, China)

Abstract: Based on the traditional detection methods have high detection error rate, energy consumption problems, puts forward the data radio network principal component analysis method and the average score combined with deviation calculation method of forgery attack detection method based on. Using function extreme value method to determine the degree of membership, to obtain the characteristics of radio network data forgery attack. Using the discrete wavelet transform method to pre process the data in the radio network. Using statistical analysis method, the determination of the radio network data forgery attack classification function, data forgery attack models for radio network, using principal component analysis method to the covariance matrix of the forgery attack on the initial radio network data eigenvalue decomposition, the average score calculation method is introduced to determine the deviation from the radio network data forgery attack deviation, realize the detection data forgery the radio network attack. Experimental results show that the improved detection method, the detection of energy consumption and false detection rate to a certain extent, compared with the traditional detection method is better, has a certain advantage.

Key words: radio network; data; forgery attack; detection

1 引言

无线网络即分组无线电技术和上联卫星组成网状链路, 其通过改变导体中电流强弱产生无线电波, 利用此无线电波, 通过调制将信息加载于无线电波之上^[1-2]. 当电波通过空间传播到达收信端, 电波引起电磁场变化会在导体产生电流, 通过解调将信息从电流变化中提取出来, 进行信息传递^[3-4]. 近年来随着无线网络的快速发展及广泛应用, 使得无

线电网络数据存在多种特性, 不乏伪造数据的存在, 其对无线网络形成一定威胁, 对数据伪造攻击进行检测, 可在一定程度上增加无线网络数据的使用率, 提高无线网络的安全性^[5-7].

文献[1]提出采用博弈论的无线网络数据伪造攻击检测方法, 该方法结合军用无线网络的应用背景, 引入信号博弈理论对数据伪造攻击进行建模, 建立博弈模型并最终得到了均衡的条件, 实现数据伪造攻击的检测. 最后进行实验结果分析, 验证了

模型的有效性,但此方法存在检测耗时低的问题.文献[8]提出采用文本聚类的数据伪造攻击检测方法,该方法采用迭代聚类过程聚类样本,并根据异常与正常样本分布规律,在聚类进程里选取最优的最大类别作为正常样本类,把其他的作为异样样本类,采用分类误差最小原则,对数据伪造攻击进行检测,但检测运行时间较长.

针对上述问题的产生,提出基于主成分分析方法与平均评分偏离度计算法结合的无线网络中数据伪造共计检测方法.

2 无线网络中数据伪造攻击信息分析

2.1 无线网络中数据伪造攻击特征提取

在对无线网络中数据伪造攻击进行检测时,需先获取数据伪造攻击特征.特征提取是从计算机视觉处理中所发展起来的概念,在计算机进行信息提取时,会对其中每个点,是否具有某个特征进行判断.

假设 $X = \{X_1, X_2, \dots, X_n\}$ 为待检测目标,每个检测目标又有 m 个特征, $X = \{X_{i1}, X_{i2}, \dots, X_{im}\}$, $i = 1, 2, \dots, n$, 则得到原始无线网络中数据矩阵为

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{bmatrix} \quad (1)$$

式中, x_{mn} 表示第 n 个检测对象的第 m 个指标的原始数据.采用函数极值法,确定隶属度,获取无线网络数据伪造攻击特征.将 n 个伪造攻击数据向量 $x_i (i = 1, 2, \dots, n)$ 分为 m 个模糊组,并求出每组的聚类中心,使得非相似性指标的目标函数达到最小,则其目标函数表达式为

$$J = \sum_j^n \sum_i^c (u_{ij})^m X d_{ij}^2 \quad 1 \leq m < \infty \quad (2)$$

式中, m 表示加权指数, u_{ij} 为 x_j 在聚类 i 中的隶属程度, $d_{ij} = \|d_j - d_i\|$ 为第 j 个数据样本与第 j 个聚类中心的欧几里德距离,隶属度计算方法为:

$$u_{ij} = \frac{J}{\sum_{k=1}^c \left(\frac{d_{ij}}{d_{kj}} \right)^{2/(m-1)}} \quad (3)$$

则数据伪造攻击特征表达式为

$$d^2(X, Y) = \sum_{i=1}^r (x_i - y_i)^2 + \lambda \cdot \sum_{i=r+1}^{r+s} u_{ij} \quad (4)$$

式中, λ 为无线网络中数据权值.

2.2 无线网络中数据伪造攻击信息预处理

在获取无线网络中数据伪造攻击特征的基础

上,采用离散小波变换法对无线网络中数据伪造攻击信息进行预处理.根据信息论的定义,计算数据伪造攻击信息变量 Y 的信息熵,计算公式为

$$H(Y) = - \sum_i p(y_i) \log_2 p(y_i) \quad (5)$$

其中, $p(y_i)$ 是算数据伪造攻击信息变量 Y 的每一个组成分量 y_i 的先验概率.再通过无线网络中数据伪造攻击信息的自相似性,获得数据伪造攻击信息的自相关函数.

自相似性是指一个随机过程在各个时间规模上具有相同的统计特性.它是复杂无线电网落的总体与部分或部分与部分间的精细结构或性质所具有的相似性,或从整体中取出局部能够体现整体的基本特征,即在不同放大倍数上的相似性.

对于一个协方差平稳的无线网络数据伪造攻击,设 $T = \{T_t, t = 1, 2, \dots\}$ 为随机到达的数据伪造攻击序列.在网络研究中, T_t 表示在第 t 个单位时间内到达的数据伪造攻击总量. T 过程具有恒定的均值 $\mu = E[T_t]$ 和有限的方差 $\sigma^2 = E[(T_t - \mu)^2]$.其自相关函数记为

$$r(k) = \frac{E[(T_t - \mu)(T_{t+k} - \mu)]}{\sigma^2 H(Y)}, \quad k = 1, 2, 3, \dots \quad (6)$$

得到下面的自相似函数定义,表示为

$$X_j^{<m>} = \frac{(X_{j-m+1} + X_{j-m} + \dots + X_j)}{mr(k)}, \quad j = 1, 2, \dots \quad (7)$$

采用小波系数的能量谱估计自相似参数,做出小波能量谱对尺度 j 做线性拟合.则小波能量函数定义为

$$S_j(q) = \frac{1}{n_j} \sum_{k=0}^{n_j-1} |d_x(j, k)|^q X_j^{<m>} \quad (8)$$

式中, $C(q)$ 和 $\alpha(q)$ 都是关于 q 的函数.

一个无线网络数据伪造攻击序列 $Y = \{Y_j, j = 1, 2, 3, \dots\}$, 选取任意正整数 m 将这个序列分成多个大小为 m 的块,将这些块分别聚合再求平均值,可以得其聚合序列为

$$Y^{<m>}(k) = \frac{1}{m_j} \sum_{j=k-1}^{km} S_j(q) Y(j), \quad k = 1, 2, 3, \dots \quad (9)$$

采用离散小波变换法,对无线网络中数据伪造攻击信息进行预处理.输入无线网络数据伪造攻击值,滤波器可以计算出在时间 T 内的滤波结果,并与预先设置的门限相比较,最终目标为根据下式得出:

$$x(t) = \begin{cases} n(t) \\ hs(t) + n(t) \end{cases} \quad (10)$$

式中, $x(t)$ 为无线网络接收信号, $s(t)$ 是无线网络数据伪造攻击发射信号, $n(t)$ 为附加高斯白噪声, h 为无线网络信道振幅增益。

3 无线网络中数据伪造攻击检测方法的改进

3.1 无线网络中数据伪造攻击模型

在与无线网络中数据伪造攻击信息进行预处理的基础上, 建立无线网络中数据伪造攻击模型。假设无线网络由静态节点构成, 节点数据间的通信为双向。无线网络可信机构 TA 选择一个 t 阶双变量对称多项式如下:

$$f(\alpha, \beta) = \sum_{i,j=0}^{t-1} a_{ij} \alpha^i \beta^j \text{Var} Y^{(m)} \quad (11)$$

式中, a_{ij} 为安全系数, 通过判断目标函数值 $J(\hat{x})$ 是否偏离正常值来确定是否存在数据伪造共计, 根据判断结果, 把状态估计值 $\hat{x}$ 代入到目标函数中, $J(\hat{x})$ 可表示为

$$J(\hat{x}) = r^T R^{-1} r f(\alpha, \beta); r = z - h(\hat{x}) \quad (12)$$

从 $J(\hat{x})$ 的定义可以知道, 在 m 个检测值, 至少有 n 个检测值满足无线网络中数据伪造攻击检测要求, 有 $(m-n)$ 个检测值误差与线性无关, 采用假设检验的方法对 $J(\hat{x})$ 的值进行判断, 判断是否为数据伪造攻击:

$$\begin{cases} H_0: J(\hat{x})P_t < \tau, \text{接受 } H_0 \\ H_1: J(\hat{x})P_t \geq \tau, \text{接受 } H_1 \end{cases} \quad (13)$$

式中, H_0 为无线网络中不存在数据伪造攻击, H_1 为无线网络中存在数据伪造攻击, τ 为控制检测率的值。为了确保建立准确的无线网络中数据伪造攻击模型, 应满足一下约束条件:

$$P_t > D_t + R_t \quad (14)$$

式中, P_t 表示有无线网络数据伪造攻击时, 为目标无线网络提供的防护时间; D_t 表示从攻击者发动攻击开始, 无线网络能够检测出攻击行为所花的时间。 R_t 为检测出攻击行为后, 将无线网络调整回正常状态所需的响应时间。采用统计学分析法, 确定无线网络数据伪造攻击分类函数:

$$f(x) = \text{sign}\left(\sum_{i=1}^L y_i a_i K(\mathbf{x}_i, \mathbf{x}) + b\right) J(\hat{x}) \quad (15)$$

式中, L 为数据伪造攻击训练样本数目, $\mathbf{x}$ 为待分类数据伪造攻击向量, $\mathbf{x}_i, y_i$ 为第 i 个数据伪造攻击训练样本属性向量和类别标识, $K(\mathbf{x}_i, \mathbf{x})$ 为核函数, a_i

和 b 表示模型参数, 则其无线网络中数据伪造攻击模型表达式为

$$\text{TypeDis}_j = \sum_{t=1}^k S_{jt} - \frac{N_j * S_t}{N f(x)} \quad (16)$$

式中, k 为数据伪造攻击类型, S_{jt} 为数据伪造攻击 j 的第 t 次攻击类型, N_j 表示出现第 j 类攻击的总数, S_t 为第 t 个攻击类型中所包含的数据量, N 为所有攻击类型中所含数据的总量。

3.2 改进检测方法的实现

在尽力无线网络数据伪造攻击模型的基础上, 采用主成分分析方法与平均评分偏离度计算法结合, 对数据伪造攻击进行检测。对于无线网络数据伪造攻击, 其攻击的目的为具有攻击依赖漏洞进行攻击, 则存在此攻击漏洞的概率可表示为

$$V(h) = \sum_{i=1}^r \eta_i \text{TypeDis}_j \quad (17)$$

式中, η_i 表示第 i 个攻击漏洞的权重, r, i 均为常数。采用博弈论获取有效抵御无线网络数据伪造攻击的效益函数, 假设 P_i 为全部无线网络数据伪造攻击, 则其效益函数表达式为

$$f_i(p_i, P_i) = R_i f(\gamma_i) V(h) \quad (18)$$

式中, R_i 表示数据伪造攻击速率, γ_i 为数据伪造攻击信息传输的有效率。设置其代价函数为

$$c_i(p_i, P_i) = \lambda \frac{p_i g_i f_i(p_i, P_i)}{\sum_{j=1}^N p_j g_j} \quad (19)$$

通过确定效益函数及效益函数可增加无线网络数据伪造攻击检测处于一个均衡状态, 提高数据伪造攻击检测的效率。由于无线网络数据伪造攻击多为随机选取, 导致攻击点一般极端的分散, 则数据伪造攻击熵大于真实攻击, 假设随机攻击属性变量为 $Q^s \sim p(\text{item} | s)$, 则攻击类为

$$s^{\text{attack}} = \underset{s}{\text{argmax}} H(Q^s) c_i(p_i, P_i) \quad (20)$$

则引入主成分分析方法对初始无线网络数据伪造攻击的协方差矩阵进行特征值分解, 采用协方差矩阵的特征向量组成投影矩阵, 实现无线网络中数据伪造攻击的检测。假设 $\mathbf{X}^{n \times m}$ 表示原始无线网络伪造攻击信息矩阵, 观测数据 $x_i = \{x_{i,1}, x_{i,2}, \dots, x_{i,m}\}$, 则计算协方差矩阵 $\mathbf{C}$:

$$\mathbf{C} = \frac{1}{n-1} \mathbf{X} \cdot \mathbf{X}^T \cdot s^{\text{attack}} \quad (21)$$

式中, n 为常数。引入平均评分偏离度计算法, 确定无线网络数据伪造攻击偏离度, 实现对无线网络数据伪造攻击的检测, 其无线网络数据伪造攻击偏离度表达式为

$$\text{LengthVar}_j = \frac{C|N_i - \bar{r}|}{\sum_{i=0}^N (N_i - \bar{r})^2} \tag{22}$$

无线网络数据伪造攻击检测结果如下：

$$\text{UIRD}_i = \frac{1}{|S|} \sum_{u \in S} \sqrt{(r_{u,i} - \bar{r}_i)^2} * \sum_{u \in S} (r_{u,i} - \bar{r}_u) \tag{23}$$

式中, S 表示对攻击 i 中部分数据伪造攻击信息集合, $r_{u,i}$ 为数据伪造攻击 u 对无线网络的损伤程度, $\bar{r}_u$ 为数据伪造攻击 u 对无线网络的平均攻击时间, $\bar{r}_i$ 为数据伪造攻击 i 在整个无线网络中数据伪造攻击中的属性分量。

4 实验结果分析

为了验证改进检测方法在无线网络数据伪造攻击方面的有效性,在运行时间不定的情况下,进行误检率方面的对比分析,结果如图 1 所示。

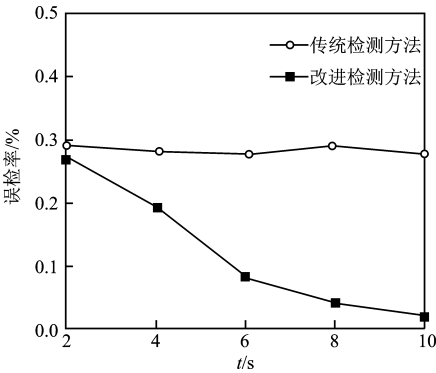


图 1 不同检测方法误检率对比分析图

由图 1 可知,在运行时间不定的情况下,传统的检测方法一直处于稳定的误检率状态,未随着运行时间的增加而增加,也未出现下降的情况,稳定性较好,但误差率较高,约为 0.29;采用改进的检测方法时,其检测误检率约为 0.13,相比传统的检测方法,误检率降低了 0.16,且随着检测运行时间的增加而逐渐下降,具有一定的优势。

为了验证改进检测方法在无线网络数据伪造攻击方面的有效性,在运行时间不定的情况下,进行检测能耗方面的对比分析,结果如图 2 所示。

由图 2 可知,采用文献[1]中的检测方法时,其能耗在 0.2~0.6 s 时出现能量消耗点密集,能耗较大;采用文献[8]中的检测方法时,其能量一开始消耗较少,但在 0.4 s 之后确出现能量消耗点密集,能耗过大的现象;采用改进的检测方法时,其一开始就

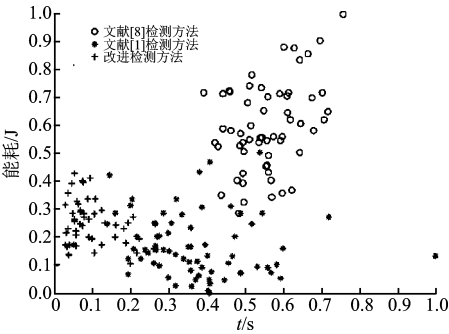


图 2 不同方法下能耗对比分析

出现能耗点密集的现象,但在 0.2 s 之后能量消耗点减少,甚至是无能耗的情况。

5 结束语

针对传统的检测方法一直存在检测误差率大,提出基于主成分分析方法与平均评分偏离度计算法结合的无线网络中数据伪造攻击检测方法。实验结果表明,采用改进检测方法时,其检测能耗及误检率方面在一定程度上,相比传统的检测方法要好,具有一定的优势。

参考文献：

[1] 张树森,张国峰,王腾飞. 基于博弈论的认知无线网络数据伪造对抗研究[J]. 现代电子技术, 2015, 39 (11): 43-45.

[2] 李鑫滨,张孜,韩松,等. 基于防伪码鉴定机制的数据伪造攻击检测方法[J]. 计算机工程, 2016, 42 (8): 112-116.

[3] 王珊珊,郗颖,黄轶岭. 基于特征数据库的认知无线网络模仿主用户攻击解决模型[J]. 电信技术研究, 2014, 55(2): 48-52.

[4] 曹娇华,苏璞睿,应凌云,等. 智能手机点击劫持攻击检测方法研究[J]. 计算机应用与软件, 2014, 31 (11): 312-317.

[5] 杨宏宇,张树茂,江华. 基于多重聚类的网络攻击检测方法研究[J]. 微电子学与计算机, 2015, 32(8): 24-29.

[6] 田玉丹,韦永壮. 认证加密算法 SCREAM 及 iSCREAM 的新伪造攻击[J]. 网络与信息安全学报, 2016, 2(1): 14-18.

[7] 刘俊霞,卜宇,刘智勇. 基于信道分配矩阵的认知无线网络容错频谱分配方案[J]. 电子设计工程, 2016, 24 (14): 77-80.

[8] 杨晓峰,李伟,孙明明,等. 基于文本聚类的网络攻击检测方法[J]. 智能系统学报, 2014, 9(1): 40-46.

作者简介：

马小雨 男, (1978-), 硕士研究生, 讲师. 研究方向为计算机网络与安全. E-mail: maxiaoyu569@163. com.