DOI: 10.3785/j.issn.1008-973X.2019.03.018

感知数据错误化攻击下分布式检测的性能分析

郑晓雁，陈惠芳，谢磊
(浙江大学 信息与电子工程学院，浙江 杭州 310027)

摘 要：针对分布式检测的安全问题，研究分布式检测中感知数据错误化(SDF)攻击及其对检测性能造成的影响. 定义一种概率型翻转攻击模型描述攻击者篡改初始感知数据的恶意行为. 以检测概率、虚警概率和错误概率为性能指标，推导在所提出的攻击模型下分布式系统检测性能的闭合表达式，进而分析 SDF 攻击下分布式检测系统的稳态性能和瞬态性能. 以偏移系数为目标函数，推导使分布式检测失效的盲化条件，即分布式系统无法检测出目标的真实状态时所对应的恶意节点攻击策略. 仿真结果表明：SDF 攻击会恶化系统检测性能，且不同攻击参数会对检测性能造成不同程度的影响.

关键词：分布式检测；安全；感知数据错误化(SDF)攻击；一致性算法；偏移系数
中图分类号：TN 92　　**文献标志码：**A　　**文章编号：**1008−973X(2019)03−0563−08

Performance analysis of distributed detection under sensing data falsification attack

ZHENG Xiao-yan, CHEN Hui-fang, XIE Lei
(*College of Information Science and Electronic Engineering*, *Zhejiang University*, *Hangzhou* 310027, *China*)

Abstract: The performance of distributed detection in the presence of sensing data falsification (SDF) attack was studied for the security problem of distributed detection. A probabilistic inverse SDF attack model was defined to characterize the malicious behavior falsifying initial sensing data. Taking the detection probability, false-alarm probability and error probability as the performance metrics, closed-form performance expressions for consensus-based distributed detection with defined SDF attack model were derived, and the steady-state and transient-state performance of distributed detection under SDF attack was analyzed. Using the deflection coefficient as objective function, the blind condition was deduced, which is the attack strategy of malicious nodes to make distributed detection system invalid. Simulation results show that the system performance degradation is caused by the SDF attack and different attack parameters have effect on the performance of consensus-based distributed detection.
Key words: distributed detection; security; sensing data falsification (SDF) attack; consensus algorithm; deflection coefficient

分布式检测是指网络中的多个传感器节点以协作的方式感知目标信号，然后通过合适的网络协议完成数据交换与信息融合，最后判断目标是否在感兴趣区域内[1-3]. 相较于需要一个融合中心收集各节点的数据进行综合处理的集中式检测，分布式检测能有效减小通信代价，减少计算开

收稿日期：2018−03−17.　　**网址：**www.zjujournals.com/eng/fileup/HTML/201903018.htm
基金项目：国家自然科学基金资助项目(61671410，61471318)；浙江省科技厅公益项目(LGG18F010005，2016C31060)；中央高校基本科研业务费专项资金资助项目(2017FZA5006).
作者简介：郑晓雁(1993-)，女，硕士生，从事分布式信息处理中安全问题研究. orcid.org/0000-0001-9140-1739. E-mail：21631151@zju.edu.cn
通信联系人：陈惠芳，女，教授. orcid.org/0000-0002-1366-1030. E-mail：chenhf@zju.edu.cn

销,具有更好的鲁棒性[4-5]. 分布式检测在无线传感器网络、认知无线电网络、物理信息系统等领域有着广泛应用[6-10].

分布式检测方法提高了检测的精度和效率,但是由于传感器节点部署在开放的环境中,容易被攻击者捕获,成为恶意节点. 攻击可能发生在感知阶段或数据交换融合阶段,可分为感知数据错误化攻击和状态数据错误化攻击 2 种方式[11]. 感知数据错误化(sensing data falsification, SDF)攻击是指攻击者伪造初始测量数据;状态数据错误化攻击是指攻击者在数据融合过程中注入错误数据. 这 2 种攻击方式会降低检测精确度或导致网络状态信息无法收敛,而且攻击者的恶意行为会通过整个网络传播,造成长时间和大范围的影响,严重干扰分布式检测系统正常工作[12]. 因此,分布式检测系统的安全问题显得尤为重要.

目前针对分布式检测系统安全问题的研究多数是从防御者角度考虑,研究防御策略[13-18],较少从攻击者角度入手,分析攻击造成的影响. 本文从攻击者角度出发,研究分布式检测中感知数据错误化攻击问题. 以偏移系数为衡量稳定状态下系统的性能指标,推导使分布式检测无效时的攻击策略. 同时,分析感知数据错误化攻击对系统稳态和瞬态检测性能造成的影响,以及不同攻击参数对系统检测性能的影响程度.

1 系统模型

分布式检测系统是由多个传感器节点和 1 个目标组成,如图 1 所示. 分布式检测通过感知、信息融合和判决 3 个步骤实现目标检测. 节点观测目标获得测量数据后,相邻节点间交换信息并更新状态,不断迭代直到所有节点达成共识,最终各节点根据最终检验统计量与预设门限比较判断目标存在 ($\mathcal{H}_1$) 还是不存在 ($\mathcal{H}_0$).

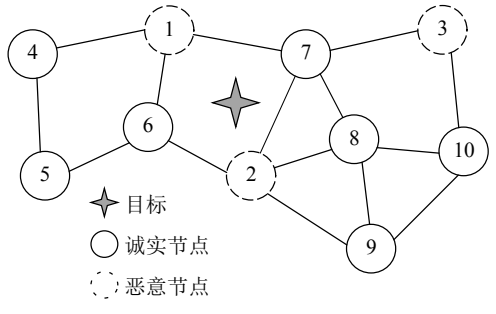

图 1 分布式检测系统示意图

Fig.1 Distributed detection network schematic diagram

1.1 网络模型

假设网络拓扑结构固定,网络部分连通且不存在孤立节点,节点间链路是对称的. 无向图 $\mathcal{G}=(\mathcal{V},\mathcal{E})$ 为网络拓扑结构,其中 $\mathcal{V}$ 为顶点集合,节点总数 $|\mathcal{V}|=N$, N 为节点数目,$\mathcal{E}$ 为边集合,节点 i 的邻居节点集合为 $\mathcal{N}_i$.

为了描述网络拓扑结构,采用度矩阵 $\boldsymbol{D}$ 表示网络中各节点的连通度,$\boldsymbol{D}=\mathrm{diag}\,[d_1,d_2,\cdots,d_N]$,$d_i$ 即节点 i 的邻节点个数. 邻接矩阵 $\boldsymbol{A}$ 表示节点间的邻接关系,若节点 i 与节点 j 相邻,则矩阵元素 $a_{ij}=1$,否则 $a_{ij}=0$. 因此,$d_i=\sum_{j\in\mathcal{V}}a_{ij}$. 拉普拉斯矩阵 $\boldsymbol{L}$ 是度矩阵和邻接矩阵之差,即 $\boldsymbol{L}=\boldsymbol{D}-\boldsymbol{A}$. 由此可得,当 $i=j$ 时,拉普拉斯矩阵元素 $l_{ij}=d_i$;当 $i\neq j$ 时,$l_{ij}=-a_{ij}$.

1.2 感知阶段

在感知阶段,网络中的各传感器节点对目标发出的信号 $s(t)$ 进行独立测量. 考虑到传感器节点的计算能力和供能受限,采用无需先验信息、复杂度低且易于实现的能量检测法. 假设目标与节点间的感知信道是加性高斯白噪声信道,$n_i(t)$ 为 t 时刻节点 i 的噪声量,即 $n_i(t)\sim\mathcal{N}\left(0,\sigma_i^2\right)$,$\sigma_i^2$ 为 $n_i(t)$ 的方差,且噪声与信号相互独立. 在 t 时刻,节点 i 感知到的目标信号为

$$r_i(t)=\begin{cases} n_i(t), & \mathcal{H}_0;\\ h_i s(t)+n_i(t), & \mathcal{H}_1. \end{cases} \tag{1}$$

式中:$1\leqslant i\leqslant N$, h_i 为目标与节点 i 间的感知信道的信道增益. 经过 M 次采样后,感知信号的能量值为 $y_i=\sum_{t=1}^{M}|r_i(t)|^2$. 由于 y_i 是 M 个相互独立并且服从正态分布的随机变量的平方之和,根据中心极限定理,当 M 足够大时,y_i 近似服从正态分布,在 $\mathcal{H}_m$(m=0 或 1)条件下 y_i 的均值和方差分别为

$$\mathrm{E}(y_i|\,\mathcal{H}_m)=\begin{cases} \mu_{0,i}=M\sigma_i^2, & \mathcal{H}_0;\\ \mu_{1,i}=(M+\eta_i)\sigma_i^2, & \mathcal{H}_1. \end{cases} \tag{2}$$

$$\mathrm{Var}\,(y_i|\mathcal{H}_m)=\begin{cases} \sigma_{0,i}^2=2M\sigma_i^4, & \mathcal{H}_0;\\ \sigma_{1,i}^2=2(M+2\eta_i)\sigma_i^4, & \mathcal{H}_1. \end{cases} \tag{3}$$

式中:“E”和“Var”分别表示取均值和取方差,η_i 为节点 i 感知信号的平均信噪比,且

$$\eta_i=|h_i|^2\left(\sum_{t=1}^{M}|s(t)|^2\right)/\sigma_i^2.$$

1.3 感知数据错误化攻击模型

为了降低系统检测性能,诱使系统作出错误判决,攻击者会捕获部分网络中的传感器节点使其成为恶意节点,恶意节点会故意篡改局部数据.

分布式网络中的恶意节点往往会在感知阶段或数据融合阶段发起攻击. 发生在感知阶段的概率型错误数据注入攻击具有较好的隐蔽性. 攻击模型描述如下: 假设恶意节点 i 以概率 p_i 发起攻击. 当恶意节点发起攻击时, 首先进行本地判决, 若本地判决结果为 $\mathcal{H}_0$, 随机生成分布满足 $\mathcal{H}_1\left(\mathcal{N}\left(\mu_{1,i},\sigma_{1,i}^2\right)\right)$ 的数值作为篡改后的局部测量数据; 若本地判决结果为 $\mathcal{H}_1$, 随机生成分布满足 $\mathcal{H}_0\left(\mathcal{N}\left(\mu_{0,i},\sigma_{0,i}^2\right)\right)$ 的数值作为篡改后的局部测量数据. 当不发起攻击时, 维持原始局部测量数据不变[19]. 相较原有攻击方式, 这种攻击方式提高了攻击的隐蔽性、有效性和破坏性. 若节点 i 是正常节点则 $p_i=0$. 恶意节点 i 进行本地判决时, 假设采用 Nayman-Pearson 准则. 恶意节点本地检测时虚警概率可表示为

$$\delta_i=\Pr\{y_i>\gamma_i|\mathcal{H}_0\}=Q\left[\frac{\gamma_i-\mathrm{E}(y_i|\mathcal{H}_0)}{\sqrt{\mathrm{Var}(y_i|\mathcal{H}_0)}}\right]=Q\left[\frac{\gamma_i-M\sigma_i^2}{\sqrt{2M\sigma_i^4}}\right]. \tag{4}$$

式中: $Q(x)$是互补累积分布函数, 且

$$Q(x)=1/\sqrt{2\pi}\cdot\int_x^{\infty}\exp(-t^2/2)\mathrm{d}t.$$

若给定虚警概率值为 $\bar{\delta}_i$, 可得本地检测的门限值 γ_i 和漏警概率 φ_i:

$$\gamma_i=Q^{-1}(\bar{\delta}_i)\sqrt{2M\sigma_i^4}+M\sigma_i^2, \tag{5}$$

$$\varphi_i=\Pr\{y_i\leqslant\gamma_i|\mathcal{H}_1\}=1-Q\left[\frac{\gamma_i-\mathrm{E}(y_i|\mathcal{H}_1)}{\sqrt{\mathrm{Var}(y_i|\mathcal{H}_1)}}\right]=1-Q\left[\frac{Q^{-1}(\bar{\delta}_i)\sqrt{2M\sigma_i^4}-\eta_i\sigma_i^2}{\sqrt{2(M+2\eta_i)\sigma_i^4}}\right]. \tag{6}$$

上述概率型感知数据错误化攻击模型可以表示为

$$\mathcal{H}_0:\tilde{y}_i=\begin{cases}\sim\mathcal{N}(\mu_{1,i},\sigma_{1,i}^2), & \text{w.p. } p_i(1-\delta_i);\\ \sim\mathcal{N}(\mu_{0,i},\sigma_{0,i}^2), & \text{w.p. } p_i\delta_i;\\ y_i, & \text{w.p. } 1-p_i.\end{cases}$$

$$\mathcal{H}_1:\tilde{y}_i=\begin{cases}\sim\mathcal{N}(\mu_{0,i},\sigma_{0,i}^2), & \text{w.p. } p_i(1-\varphi_i);\\ \sim\mathcal{N}(\mu_{1,i},\sigma_{1,i}^2), & \text{w.p. } p_i\varphi_i;\\ y_i, & \text{w.p. } 1-p_i.\end{cases} \tag{7}$$

式中: “w.p.”表示事件发生的概率. 由于 $\tilde{y}_i$ 服从混合高斯分布, 根据混合高斯分布的性质, 篡改后的局部测量数据均值和方差分别为

$$\mathrm{E}(\tilde{y}_i|\mathcal{H}_m)=\begin{cases}\tilde{\mu}_{0,i}=M\sigma_i^2+p_i(1-\delta_i)\eta_i\sigma_i^2, & \mathcal{H}_0;\\ \tilde{\mu}_{1,i}=(M+\eta_i)\sigma_i^2-p_i(1-\varphi_i)\eta_i\sigma_i^2, & \mathcal{H}_1.\end{cases} \tag{8}$$

$$\mathrm{Var}(\tilde{y}_i|\mathcal{H}_m)=\begin{cases}\tilde{\sigma}_{0,i}^2=2M\sigma_i^4+4p_i(1-\delta_i)\eta_i\sigma_i^4+p_i(1-\delta_i)(1-p_i+p_i\delta_i)\eta_i^2\sigma_i^4, & \mathcal{H}_0;\\ \tilde{\sigma}_{1,i}^2=2(M+2\eta_i)\sigma_i^4-4p_i(1-\varphi_i)\eta_i\sigma_i^4+p_i(1-\varphi_i)(1-p_i+p_i\varphi_i)\eta_i^2\sigma_i^4, & \mathcal{H}_1.\end{cases} \tag{9}$$

1.4 融合阶段

在数据融合阶段, 每个节点与相距 1 跳的邻节点交换信息, 并更新节点状态, 信息交换不断发生直到网络中所有节点达成共识, 状态收敛. 在前阶段各节点已经获取检测量信息, 假设节点间能无差错地交换数据, 令节点 i 数据融合前初始状态为 $x_i(0)=\tilde{y}_i$, 采用加权平均的一致性算法, 状态信息 $x_i(k)$更新过程可以表示为

$$x_i(k)=x_i(k-1)+\frac{\varepsilon}{w_i}\sum_{j\in\mathcal{N}_i}\left[x_j(k-1)-x_i(k-1)\right]. \tag{10}$$

式中: ε 为迭代步长, w_i 为节点 i 的权重系数.

上述更新过程可以看作: 各节点的更新结果为前一时刻节点本地状态信息与前一时刻邻居节点状态信息的加权平均, 也可写成矩阵的形式(式(12)). 确保该算法收敛的首要条件是式(10)中各节点的相对加权系数非负(式(12)中的权重矩阵 $\boldsymbol{W}$ 非负). 当 $i=j$ 时, 节点本身的相对加权系数为 $w_{ii}=1-\varepsilon d_i/w_i$; 由 $1-\varepsilon d_i/w_i>0$, 解得 $\varepsilon<w_i/d_i$. 当 $j\in\mathcal{N}_i$ 时, 邻居节点的相对加权系数 $w_{ij}=\varepsilon/w_i$; 由 $\varepsilon/w_i>0$, 解得 $\varepsilon>0$. 其余节点的相对加权系数 $w_{ij}=0$. 迭代步长 ε 适用于所有节点更新过程, 因此需要满足 $0<\varepsilon<\min(w_i/d_i)$, $i\in\{1,2,\cdots,N\}$, 以保证算法收敛性. 该收敛条件具体的物理含义是保证状态更新过程中相对加权系数始终非负, 从而使得更新后节点状态收敛. 当达到共识时, $x_1=x_2=\cdots=x_N=x^*$, 即各节点收敛于 $\boldsymbol{x}=x^*\boldsymbol{1}$.

权重系数的值取决于节点的连通度、感知信道状况等因素. 从节点连通度考虑, 常见的融合准则有平均准则、Metropolis 准则、最大度准则等[20]. 在平均准则中, 权重矩阵元素设置如下: 当 $j\in\mathcal{N}_i\cup\{i\}$ 时, $w_{ij}=(1+d_i)^{-1}$; 否则, $w_{ij}=0$. 在 Metropolis 准则中, 当 $j\in\mathcal{N}_i$ 时, $w_{ij}=\left[1+\max\left(d_i,d_j\right)\right]^{-1}$, 节点自身权重为 $w_{ii}=1-\sum_{j\in\mathcal{N}_i}w_{ij}$, 当 $j\notin\mathcal{N}_i\cup\{i\}$时, $w_{ij}=0$. 在最大度准则中, 当 $j\in\mathcal{N}_i$ 时, $w_{ij}=[1+\max_{i\in\mathcal{V}}(d_i)]^{-1}$, 节点自身权重为 $w_{ii}=1-\sum_{j\in\mathcal{N}_i}w_{ij}$, 当 $j\notin\mathcal{N}_i\cup\{i\}$时, $w_{ij}=0$. 也可以根据网络中各节点感知信道信噪比等信道特征参数设置权重系数. 根据信噪比 η_i、噪声功率 σ_i^2 等信道特征参数分配权重系数, 归一化后的权重系数为

$$w_i = \eta_i \sigma_i^{-2} \Big/ \sum_{i=1}^{N} \left(\eta_i \sigma_i^{-2}\right).$$

1.5 判决阶段

在判决阶段,将各节点达到的一致性收敛结果 x^*作为最终检验统计量,与设定本地门限 λ_i 进行比较,独立判断目标存在与否. 节点 i 的判决过程可以表示为

$$\mathcal{H}_1: \quad x^* > \lambda_i; \ \mathcal{H}_0: \quad x^* \leqslant \lambda_i. \tag{11}$$

如果算法不收敛或不满足收敛条件,可以设定迭代次数或迭代时间上限,当执行次数或时间达到这一上限时,停止执行一致性算法,将当前轮次或时刻节点的状态值作为检验统计量,与本地门限比较,判断目标是否存在.

综上所述,攻击者发起感知数据错误化攻击,致使最终检验统计量有偏差,导致判决失误. 因此,下文将研究感知数据错误化攻击对分布式检测性能的影响,以及采用不同攻击策略与攻击参数对性能的影响程度.

2 性能分析

从攻击者的角度分析感知数据错误化攻击对分布式检测系统稳态性能和瞬态性能的影响,以及采取何种攻击策略可使检测机制无效化. 采用偏移系数、检测概率和虚警概率衡量系统检测性能.

2.1 稳态性能分析

假设网络中有 N_1 个恶意节点,其余为正常节点. 令 x^*表示 $x_i(k)$的稳态值,按照一致性算法进行足够多次迭代运算后,所有传感器节点的检验统计量会达到全局统一的收敛值,网络达到稳定状态. 方便起见,采用矩阵形式表示式(10)的更新过程,即

$$\boldsymbol{x}(k) = \boldsymbol{W}\boldsymbol{x}(k-1), \tag{12}$$

式中:节点状态向量 $\boldsymbol{x}(k) = [x_1(k), x_2(k), \cdots, x_N(k)]^{\mathrm{T}}$,权重矩阵 $\boldsymbol{W} = \boldsymbol{I}_N - \varepsilon \operatorname{diag}(1/w_1, 1/w_2, \cdots, 1/w_N)\boldsymbol{L}$, $\boldsymbol{I}_N$ 是 N 阶单位矩阵.

由权重矩阵 $\boldsymbol{W}$的定义式,可得 $\boldsymbol{W}\boldsymbol{1}=\boldsymbol{1}$, $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{W} = \boldsymbol{w}^{\mathrm{T}}$,其中, $\boldsymbol{1}=[1, 1, \cdots, 1]^{\mathrm{T}}$ 是元素全为 1 的 N 维列向量, $\boldsymbol{w}^{\mathrm{T}} = [w_1, w_2, \cdots, w_N]$ 是各节点权重系数构成的 N 维行向量. 因此, $\boldsymbol{1}$ 和 $\boldsymbol{w}^{\mathrm{T}}$ 分别为 $\boldsymbol{W}$的右特征向量和左特征向量. 根据 Perron-Frobenius 定理, $\lim\limits_{k\to\infty} \boldsymbol{W}^k = \boldsymbol{1}\boldsymbol{w}^{\mathrm{T}}$,最终收敛于各节点初始状态数据的加权平均,即

$$x^* = \lim_{k\to\infty} x_i(k) = \frac{\sum_{i=1}^{N_1}(w_i \tilde{y}_i) + \sum_{i=N_1+1}^{N}(w_i y_i)}{\sum_{i=1}^{N} w_i}. \tag{13}$$

根据正态分布的性质,多个相互独立的正态随机变量的线性组合仍服从正态分布,则最终检验统计量 x^*服从高斯分布,其均值和方差分别为

$$\begin{aligned} \mathrm{E}(x^*|\mathcal{H}_m) = \mu_m = &\sum_{i=1}^{N_1} \left\{ \left(w_i \Big/ \sum_{i=1}^{N} w_i \right) \tilde{\mu}_{m,i} \right\} + \\ &\sum_{i=N_1+1}^{N} \left\{ \left(w_i \Big/ \sum_{i=1}^{N} w_i \right) \mu_{m,i} \right\}, \end{aligned} \tag{14}$$

$$\begin{aligned} \mathrm{Var}(x^*|\mathcal{H}_m) = \sigma_m^2 = &\sum_{i=1}^{N_1} \left\{ \left(w_i \Big/ \sum_{i=1}^{N} w_i \right)^2 \tilde{\sigma}_{m,i}^2 \right\} + \\ &\sum_{i=N_1+1}^{N} \left\{ \left(w_i \Big/ \sum_{i=1}^{N} w_i \right)^2 \sigma_{m,i}^2 \right\}. \end{aligned} \tag{15}$$

式中: $m \in \{0, 1\}$. 各个节点比较收敛值和本地门限的大小完成最终判决. 若先验概率 $P(\mathcal{H}_0)$ 和 $P(\mathcal{H}_1)$ 已知,则节点 i 的虚警概率 $P_{\mathrm{f},i}$、检测概率 $P_{\mathrm{d},i}$ 和错误概率 $P_{\mathrm{e},i}$的表达式为

$$P_{\mathrm{f},i} = \Pr\{x^* > \lambda_i | \mathcal{H}_0\} = Q\left(\frac{\lambda_i - \mathrm{E}(x^*|\mathcal{H}_0)}{\sqrt{\mathrm{Var}(x^*|\mathcal{H}_0)}} \right), \tag{16}$$

$$P_{\mathrm{d},i} = \Pr\{x^* > \lambda_i | \mathcal{H}_1\} = Q\left(\frac{\lambda_i - \mathrm{E}(x^*|\mathcal{H}_1)}{\sqrt{\mathrm{Var}(x^*|\mathcal{H}_1)}} \right), \tag{17}$$

$$P_{\mathrm{e},i} = P_{\mathrm{f},i} P(\mathcal{H}_0) + (1 - P_{\mathrm{d},i}) P(\mathcal{H}_1). \tag{18}$$

其中, λ_i 为节点 i 预设的判决门限值. 利用 Neyman-Pearson 准则,当节点 i 的虚警概率 $\bar{P}_{\mathrm{f},i}$ 给定时,判决门限为

$$\lambda_i = \mathrm{E}(x^*|\mathcal{H}_0) + \sqrt{\mathrm{Var}(x^*|\mathcal{H}_0)} Q^{-1}(\bar{P}_{\mathrm{f},i}). \tag{19}$$

2.2 盲化条件

对于分布式检测系统,也可以采用偏移系数表征达到稳定状态时网络检测性能,偏移系数越大意味着系统检测性能越好,其定义如下:

$$D(x^*) = \frac{[\mathrm{E}(x^*|\mathcal{H}_0) - \mathrm{E}(x^*|\mathcal{H}_1)]^2}{\mathrm{Var}(x^*|\mathcal{H}_0)}. \tag{20}$$

为了恶化网络检测性能达到攻击目标,攻击者会使偏移系数尽可能小. 当偏移系数为 0 时,检测机制被彻底破坏,此时网络检测的结果与随机猜测无异,网络完全盲化,无法根据测量与收集的数据正确判断目标是否存在. 网络盲化时恶意节点占总节点数的最小比例称为盲化比例,记为

α_{blind}. 当 $D(x^*)=0$ 时，即当 $E(x^*|\mathcal{H}_0)=E(x^*|\mathcal{H}_1)$ 时，将式 (8) 和 (9) 代入，网络完全盲化时下式成立：

$$\sum_{i=1}^{N_1}\{w_i[p_i(2-\delta_i-\varphi_i)-1]\}=\sum_{i=N_1+1}^{N}w_i. \tag{21}$$

假设各节点配置相同，$\sigma_i^2=\sigma^2$，$\eta_i=\eta$，$\delta_i=\delta$，$\varphi_i=\varphi$；恶意节点攻击概率相同，$p_i=p$；权重系数相等，$w_i=w$. 设恶意节点的比例为 $\alpha=N_1/N$，可进一步简化式 (21)，得到盲化条件：

$$\alpha p=(2-\delta-\varphi)^{-1}. \tag{22}$$

若 p 是给定的，盲化比例为 $\alpha_{\text{blind}}=1[(2-\delta-\varphi)p]^{-1}$.

2.3 瞬态性能分析

以迭代 k 次后的状态作为检验统计量，分析感知数据错误化攻击下各个传感器节点在迭代到第 k 次时的瞬态检测性能，并以检测概率和虚警概率作为性能指标. 设 k 次迭代后的数据为 $\boldsymbol{x}(k)=\boldsymbol{W}^k\boldsymbol{x}(0)$. 令 w_{ij}^k 为权重矩阵 $\boldsymbol{W}$ 迭代 k 次后的项，设编号为前 N_1 个节点为恶意节点. 恶意节点有 3 种可能状态：不发起攻击，发起攻击且攻击成功，发起攻击但攻击失败. 节点 i 的攻击状态可用 q_i 表示，$q_i=0$ 表示攻击成功，$q_i=1$ 表示攻击失败，$q_i=-1$ 表示没有发起攻击. 方便起见，令 $b_{0,i}=p_i(1-\delta_i)$，$b_{1,i}=p_i\delta_i$，$b_{-1,i}=1-p_i$，$c_{0,i}=p_i(1-\varphi_i)$，$c_{1,i}=p_i\varphi_i$，$c_{-1,i}=1-p_i$，$!|q_i|$表示对 q_i 的绝对值取反. 在概率型感知数据错误化攻击模型下，k 次迭代后节点 i 处的检验统计量在不同假设下的概率密度函数为

$$f(\tilde{x}_i(k)|\mathcal{H}_0)=\sum_{q_{N_1}=\{0,\pm1\}}\cdots\sum_{q_2=\{0,\pm1\}}\sum_{q_1=\{0,\pm1\}}\Bigg\{b_{q_1,1}b_{q_2,2}\cdots b_{q_{N_1},N_1}\times f\left(\frac{\tilde{x}_i(k)-w_{ij}^k\mu_{(!|q_1|),1}-w_{ij}^k\mu_{(!|q_2|),2}-\cdots-w_{ij}^k\mu_{(!|q_{N_1}|),N_1}-\sum_{j=N_1+1}^{N}w_{ij}^k\mu_{0,j}}{\left[\left(w_{ij}^k\sigma_{(!|q_1|),1}\right)^2+\left(w_{ij}^k\sigma_{(!|q_2|),2}\right)^2+\cdots+\left(w_{ij}^k\sigma_{(!|q_{N_1}|),N_1}\right)^2+\sum_{j=N_1+1}^{N}\left(w_{ij}^k\sigma_{0,j}\right)^2\right]^{1/2}}\right)\Bigg\}, \tag{23}$$

$$f(\tilde{x}_i(k)|\mathcal{H}_1)=\sum_{q_{N_1}=\{0,\pm1\}}\cdots\sum_{q_2=\{0,\pm1\}}\sum_{q_1=\{0,\pm1\}}\Bigg\{c_{q_1,1}c_{q_2,2}\cdots c_{q_{N_1},N_1}\times f\left(\frac{\tilde{x}_i(k)-w_{ij}^k\mu_{|q_1|,1}-w_{ij}^k\mu_{|q_2|,2}-\cdots-w_{ij}^k\mu_{|q_{N_1}|,N_1}-\sum_{j=N_1+1}^{N}w_{ij}^k\mu_{1,j}}{\left[\left(w_{ij}^k\sigma_{|q_1|,1}\right)^2+\left(w_{ij}^k\sigma_{|q_2|,2}\right)^2+\cdots+\left(w_{ij}^k\sigma_{|q_{N_1}|,N_1}\right)^2+\sum_{j=N_1+1}^{N}\left(w_{ij}^k\sigma_{1,j}\right)^2\right]^{1/2}}\right)\Bigg\}. \tag{24}$$

根据概率密度函数推导不同节点虚警概率和检测概率的表达式，节点 i 在 k 次迭代后的瞬态虚警概率和检测概率表达式为

$$P_{\text{f},i}(k)=\Pr\{x_i(k)>\lambda_i|\mathcal{H}_0\}=\sum_{q_{N_1}=\{0,\pm1\}}\cdots\sum_{q_2=\{0,\pm1\}}\sum_{q_1=\{0,\pm1\}}\Bigg\{b_{q_1,1}b_{q_2,2},\cdots,b_{q_{N_1},N_1}\times Q\left[\frac{\left(\lambda_i-w_{ij}^k\mu_{(!|q_1|),1}-w_{ij}^k\mu_{(!|q_2|),2}-\cdots-w_{ij}^k\mu_{(!|q_{N_1}|),N_1}-\sum_{j=N_1+1}^{N}w_{ij}^k\mu_{0,j}\right)}{\left(\left(w_{ij}^k\sigma_{(!|q_1|),1}\right)^2+\left(w_{ij}^k\sigma_{(!|q_2|),2}\right)^2+\cdots+\left(w_{ij}^k\sigma_{(!|q_{N_1}|),N_1}\right)^2+\sum_{j=N_1+1}^{N}\left(w_{ij}^k\sigma_{0,j}\right)^2\right)^{1/2}}\right]\Bigg\}, \tag{25}$$

$$P_{\text{d},i}(k)=\Pr\{x_i(k)>\lambda_i|\mathcal{H}_1\}=\sum_{q_{N_1}=\{0,\pm1\}}\cdots\sum_{q_2=\{0,\pm1\}}\sum_{q_1=\{0,\pm1\}}\Bigg\{c_{q_1,1}c_{q_2,2},\cdots,c_{q_{N_1},N_1}\times Q\left[\frac{\lambda_i-w_{ij}^k\mu_{|q_1|,1}-w_{ij}^k\mu_{|q_2|,2}-\cdots-w_{ij}^k\mu_{|q_{N_1}|,N_1}-\sum_{j=N_1+1}^{N}w_{ij}^k\mu_{1,j}}{\left(\left(w_{ij}^k\sigma_{|q_1|,1}\right)^2+\left(w_{ij}^k\sigma_{|q_2|,2}\right)^2+\cdots+\left(w_{ij}^k\sigma_{|q_{N_1}|,N_1}\right)^2+\sum_{j=N_1+1}^{N}\left(w_{ij}^k\sigma_{1,j}\right)^2\right)^{1/2}}\right]\Bigg\}. \tag{26}$$

3 性能仿真与分析

假设网络中共有 10 个节点，存在若干恶意节点，先验概率 $P(\mathcal{H}_0)=P(\mathcal{H}_1)=0.5$，目标发射信号能量为 5，采样 20 次，信道增益相同且均为 1，噪声功率均取 1，假设信道能无差错传输数据，给定本地判决的虚警概率为 0.1.

如图 2 和 3 所示为系统达到稳定状态时，不同恶意节点比例与攻击概率对系统偏移系数 D 和错误概率 P_{e} 的影响. 从图 2 和 3 可以看到，当

网络中没有攻击者时，偏移系数的值最大，全局错误概率最小，系统稳态检测性能最佳. 当攻击概率变大或者恶意节点比例增大时，偏移系数逐渐减小，全局错误概率增大，系统检测性能变差，网络趋于盲化. 当 $\alpha p \approx 0.80$ 时，偏移系数接近于 0，全局错误概率约为 0.5，此时系统盲化，无法根据局部测量与收集数据正确判断目标真实状态，判决结果与随机猜测无异.

如图 4 所示为不同攻击参数下分布式检测系统的受试者工作特性（receiver operating characteristic，ROC）曲线. ROC 曲线以虚警概率 P_f 为横坐标，检测概率 P_d 为纵坐标，曲线下面积越大，表示检测精确度越高. 如图 4 所示，随着网络中恶意节点比例和攻击概率的增大，ROC 曲线不断接近图中的虚线对角线；曲线下方面积减少，网络检测性能逐渐变差.

考虑攻击对瞬态性能的影响，仿真各个节点的检测概率和虚警概率随一致性迭代次数增加时的变化情况. 假设网络中编号 1~3 的节点是恶意节点，攻击概率 p=0.5. 图 5 和 6 比较了有无攻击

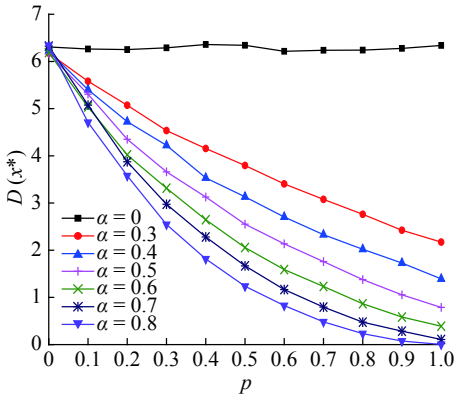

图 2 不同攻击参数对偏移系数的影响

Fig.2 Impact of different attack parameters on deflection coefficient

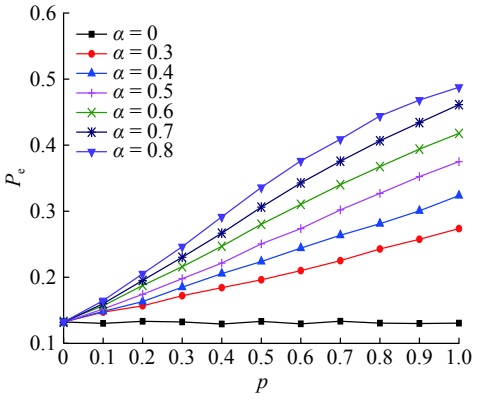

图 3 不同攻击参数对错误概率的影响

Fig.3 Impact of different attack parameters on error probability

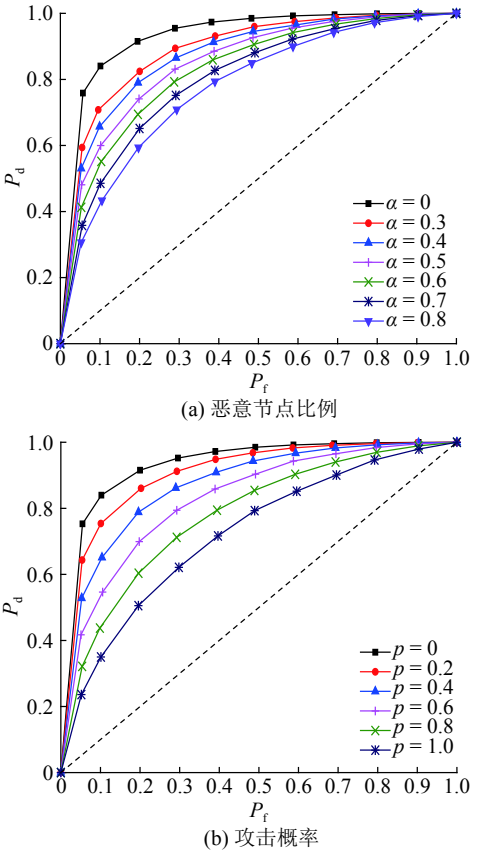

图 4 不同攻击参数下的受试者工作特性（ROC）曲线

Fig.4 Receiver operating characteristic (ROC) curves with different attack parameters

时各节点的检测概率和虚警概率随迭代次数增加时的变化关系. 如图 5 和 6 所示，随着一致性迭代次数增加，各节点的检测概率逐渐增大且最终达到收敛，虚警概率不断减小最终也达到收敛. 当网络没有遭致攻击时，大约迭代 25 次后，各节点的检测概率收敛于 95% 左右，虚警概率收敛速度慢于检测概率，最终收敛于 10%左右；当网络存在攻击时，收敛需要的迭代次数略有增加. 当迭代次数约为 30 次时，各节点检测概率大致收敛于 92%，之后虚警概率大致收敛于 25%. 可见概率型感知数据错误化攻击的确会恶化系统性能，使检测概率减小，虚警概率增大，收敛时间加长，导致系统整体性能变差.

4 结 语

针对没有融合中心的分布式检测，从感知数据错误化攻击影响检测性能的角度展开研究工作. 分布式检测采用一致性算法，各节点将局部测量数据线性融合后获得全局检验统计量，对目

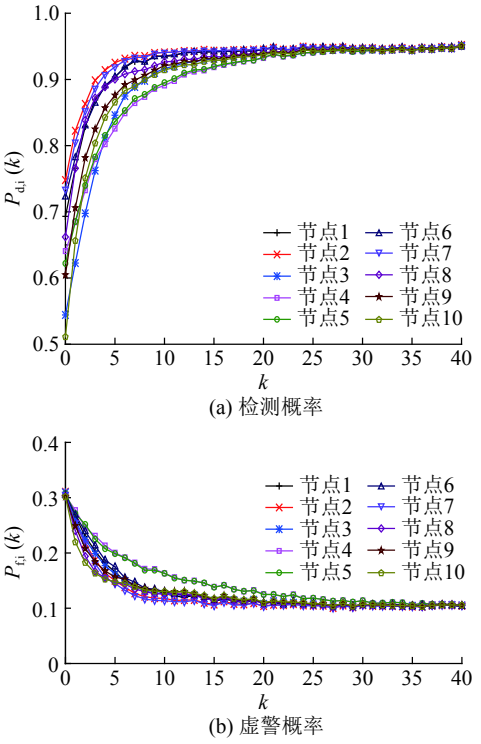

图 5　无攻击时检测概率和虚警概率随迭代次数变化关系

Fig.5　Probability of detection and probability of false alarm as functions of iteration steps without attack

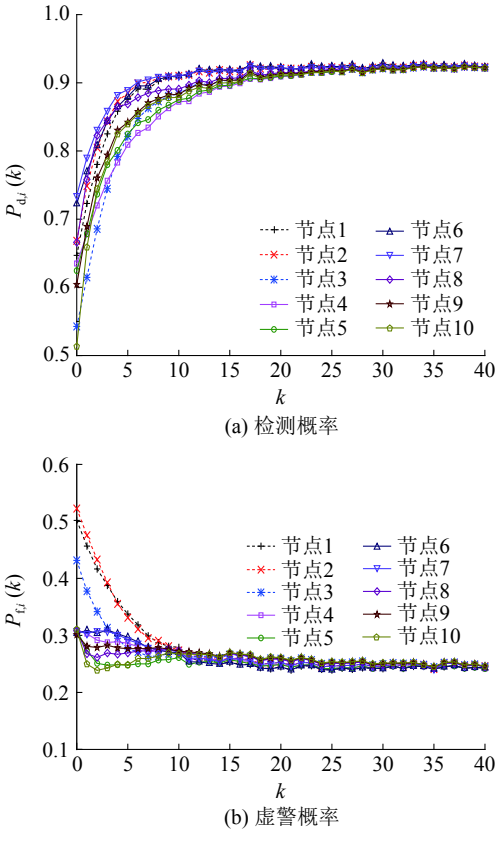

图 6　有攻击时检测概率和虚警概率随迭代次数变化关系

Fig.6　Probability of detection and probability of false alarm as functions of iteration steps with attack

标是否存在作出判决. 本文定义了一种概率型感知数据错误化攻击模型, 以偏移系数、虚警概率、检测概率为分布式检测性能指标, 分析了感知数据错误化攻击对系统检测性能的影响以及不同攻击方式对系统检测性能造成影响的程度. 同时, 利用偏移系数为性能指标, 推导了感知数据错误化攻击使分布式检测失效需要满足的条件. 下一阶段将对状态数据错误化攻击及防御策略展开研究.

参考文献 (References):

[1] KAILKHURA B, HAN Y S, BRAHMA S, et al. Distributed Bayesian detection in the presence of Byzantine Data [J]. **IEEE Transactions on Signal Processing**, 2015, 63(19): 5250–5263.

[2] STANKOVIC S S, ILIC N, STANKOVIC M S, et al. Distributed change detection based on a consensus algorithm [J]. **IEEE Transactions on Signal Processing**, 2011, 59(12): 5686–5697.

[3] 孙小静. 无线传感器网络中分布式检测的相关问题研究 [D]. 南京: 东南大学, 2017.
SUN Xiao-jing. Some issues on distributed detection in wireless sensor network [D]. Nanjing: Southeast University, 2017.

[4] LI Z, YU F R, HUANG M. A distributed consensus-based cooperative spectrum-sensing scheme in cognitive radios [J]. **IEEE Transactions on Vehicular Technology**, 2010, 59(1): 383–393.

[5] ZHANG W, WANG Z, GUO Y, et al. Distributed cooperative spectrum sensing based on weighted average consensus [C] // **IEEE GLOBECOM**. Houston: IEEE, 2011: 1–6.

[6] ALONSO-ROMAN D, BEFERULL-LOZANO B. Adaptive consensus-based distributed detection in WSN with unreliable links [C] // **IEEE ICASSP**. Shanghai: IEEE, 2016: 4438–4442.

[7] GUO J, ROGERS U, LI X, et al. Secrecy constrained distributed detection in sensor networks [J]. **IEEE Transactions on Signal and Information Processing over Networks**, 2018, 4(2): 378–391.

[8] ZHANG W, GUO Y, LIU H, et al. Distributed consensus-based weight design for cooperative spectrum sensing [J]. **IEEE Transactions on Parallel and Distributed Systems**, 2015, 26(1): 54–64.

[9] LIU S, ZHU H, LI S, et al. An adaptive deviation-tolerant secure scheme for distributed cooperative spectrum sensing [C] // **IEEE GLOBECOM**. Anaheim: IEEE, 2012: 603–608.

[10] SADREAZAMI H, MOHAMMADI A, ASIF A, et al. Distributed graph-based statistical approach for intrusion detection in cyber-physical systems [J]. **IEEE Transactions on Signal and Information Processing over Networks**, 2018, 4(1): 137–147.

[11] KAILKHURA B, BRAHMA S, VARSHNEY P K. Data falsification attacks on consensus-based detection systems [J]. **IEEE Transactions on Signal and Information Processing over Networks**, 2017, 3(1): 145–158.

[12] 米士超. 基于无线传感器网络的分布式估计及安全机制设计[D] 上海: 上海交通大学, 2015.
MI Shi-chao. Distributed estimation and security mechanism for wireless sensor networks [D]. Shanghai: Shanghai Jiao Tong University, 2015.

[13] YAN Q, LI M, JIANG T, et al. Vulnerability and protection for distributed consensus-based spectrum sensing in cognitive radio networks [C] // **IEEE INFOCOM**. Orlando: IEEE, 2012: 900–908.

[14] VOSOUGHI A, CAVALLARO J R, MARSHALL A. Trust-aware consensus-inspired distributed cooperative spectrum sensing for cognitive radio ad hoc networks [J]. **IEEE Transactions on Cognitive Communications and Networking**, 2016, 2(1): 24–37.

[15] GENTZ R, WU S X, WAI H T, et al. Data injection attacks in randomized gossiping [J]. **IEEE Transactions on Signal and Information Processing over Networks**, 2016, 2(4): 523–538.

[16] ZHAO C, HE J, CHEN J. Resilient consensus with mobile detectors against malicious attacks [J]. **IEEE Transactions on Signal and Information Processing over Networks**, 2018, 4(1): 60–69.

[17] WANG P, CHEN C, ZHU S, et al. An optimal reputation-based detection against SSDF attacks in industrial cognitive radio network [C] // **IEEE ICCA**. Ohrid: IEEE, 2017: 729–734.

[18] FENG H, LIANG L, LEI H. Distributed outlier detection algorithm based on credibility feedback in wireless sensor networks [J]. **IET Communications**, 2017, 11(8): 1291–1296.

[19] 周明. 认知无线电网络合作频谱感知中的 SSDF 攻击及其防御机制[D]. 杭州: 浙江大学, 2016.
ZHOU Ming. SSDF attack and defense strategies in cooperative spectrum sensing of cognitive radio networks [D]. Hangzhou: Zhejiang University, 2016.

[20] VOSOUGHI A, CAVALLARO J R, MARSHALL A. Robust consensus-based cooperative spectrum sensing under insistent spectrum sensing data falsification attacks [C] // **IEEE GLOBECOM**. San Diego: IEEE, 2015: 1–6.