控制与决策
Control and Decision

分布式最小二乘估计中隐匿FDI攻击策略的设计

胡明南， 陈博， 俞立

引用本文:
胡明南, 陈博, 俞立. 分布式最小二乘估计中隐匿FDI攻击策略的设计[J]. 控制与决策, 2021, 36(8): 1963−1969.

在线阅读 View online: https://doi.org/10.13195/j.kzyjc.2019.1688

分布式最小二乘估计中隐匿FDI攻击策略的设计

胡明南, 陈 博†, 俞 立

(1. 浙江工业大学 信息工程学院，杭州 310023；2. 浙江工业大学 网络空间安全研究院，杭州 310023)

摘 要: 虽然分布式坏值检测方法能够消除观测数据中坏值对分布式最小二乘估计性能的影响,但是现有的分布式坏值检测方法中依然存在安全漏洞. 针对一类分布式最小二乘估计算法研究了相应隐匿虚假数据注入(FDI)攻击策略的设计问题,设计依赖于部分节点系统信息的分布式隐匿FDI攻击方法,这一方法不仅使得FDI攻击信号无法被现有分布式坏值检测方法检测到,而且可以以预设的偏移量改变估计结果. 最后,通过IEEE 118-Bus电力系统模型验证所设计FDI攻击方法的隐匿性和有效性.

关键词: 安全估计；分布式隐匿FDI攻击；分布式最小二乘估计；坏值检测；分布式系统

中图分类号: TP273 **文献标志码:** A

DOI: 10.13195/j.kzyjc.2019.1688 **开放科学(资源服务)标识码(OSID):**

引用格式: 胡明南,陈博,俞立. 分布式最小二乘估计中隐匿FDI攻击策略的设计[J]. 控制与决策, 2021, 36(8): 1963-1969.

Hidden FDI attack strategy for distributed least square estimation

HU Ming-nan, CHEN Bo†, YU Li

(1. College of Information Science and Engineering，Zhejiang University of Technology，Hangzhou 310023，China；
2. Research Institute of Cyberspace Security，Zhejiang University of Technology，Hangzhou 310023，China)

Abstract: Although the distributed bad data detection algorithm can eliminate the influence of bad values for the least square estimation performence, there are still security vulnerabilities in the existing distributed bad data detection algorithms. Motivated by the above-mentioned fact, this paper proposes a hidden false data injection (FDI) attack method to degrade the performance of the distributed least square estimator, where the hidden attack method is designed based on system information of some nodes. This method not only makes the FDI attack signal unable to be detected by the existing distributed bad value detection methods, but also changes the estimation result with a preset offset. Finally, the IEEE 118-Bus power system is used to illustrate the hiddenness and effectiveness of the proposed methods.

Keywords: security estimation；distributed hidden FDI attack；distributed least square estimation；bad data detection；distributed systems

0 引 言

最小二乘估计是根据可观测数据估计系统内部状态的方法,它借助观测的冗余度来提高数据精度,剔除随机干扰引起的数据噪声以获得系统运行状态的最优估计. 近年来,随着网络通信技术的发展,现代信息处理系统的规模越来越大,传统的集中式最小二乘估计算法对计算资源、数据存储资源和通信带宽的要求越来越高. 因此,对计算和通信要求更低、鲁棒性更高、数据分散存储在各个分布式节点中的分布式估计算法得到了广泛的关注,不同的分布式估计方法相继被提出[1-9]. 然而在分布式框架下,大量原本封闭的系统逐渐通过网络传输信息,直接增加了系统受到外界攻击的风险. 因此,研究分布式最小二乘估计算法中存在的安全隐患是至关重要的.

分布式系统中每个节点只拥有本地的系统参数信息和本地的观测数据,通过所有节点之间执行协作算法来进行最小二乘估计. 文献[10]指出,现有的分布式估计算法按计算结构可以分为分层分布式(有融合中心)和完全分布式(无融合中心)两种. 分层分布式估计算法通常由两步组成:第1步由每个节点利用本地信息进行最小二乘估计,并将估计结果发往融合中心;第2步由融合中心将接收到的状态进行融合,得到次优的估计结果. 与分层分布式估计算法相比,完全分布式算法拥有鲁棒性高、可收敛至最优解的

收稿日期: 2019-12-02；**修回日期:** 2020-03-14.
基金项目: 国家自然科学基金项目(61673351, 61973277)；浙江省自然科学基金项目(LR20F030004).
†通讯作者. E-mail: bchen@zjut.edu.cn.

优点. 在完全分布式的框架下,每个节点与其相邻节点通信,所有节点通过执行协作算法得到全局最优估计值. 完全分布式最小二乘估计算法通常基于拉格朗日松弛的方法[1],通过求解分解迭代优化问题,使各节点的估计结果达成共识[1-4]. 但是,基于拉格朗日松弛的分布式最小二乘估计算法收敛速度十分缓慢,难以满足大规模系统的实时性要求. 为此,文献[6]通过直接求解单区域优化问题设计估计器,但是估计性能不等价于集中式方法. 进一步,文献[7-9]利用求解线性方程的方法设计等价于集中式估计性能的分布式最小二乘估计器. 文献[9]利用Richardson方法[11]并结合预处理矩阵降低系统矩阵的病态系数,极大地加快了算法的收敛速率.

另一方面,在实际系统的应用过程中,观测信息所出现的不良数据会导致估计值大幅偏离真实值. 在此情况下,系统中的不良数据通常采用坏值检测方法进行检测,其中分布式坏值检测方法最早由文献[12]提出,并指出残差只会在分布式系统某一区域内扩散,而不会对其他区域产生影响. 基于这一结论,文献[13]设计了一种降阶的分布式坏值检测方法,通过相邻子系统之间进行数据交换实现完全分布式坏值检测,这种方法使子系统的划分更加灵活. 文献[14-15]分别引入伪量测和数据修补方法,在剔除坏值的同时保证分布式最小二乘估计算法稳定运行. 注意到,与传统的集中式坏值检测方法相比,分布式检测方法可以针对每个子系统都设置更低的坏值检测阈值,从而提高坏值检测算法的识别精度[16].

虽然坏值检测算法经过了多年的发展并应用于诸多领域,但是大多数坏值检测算法均基于如下假设:当观测值中出现坏值时,残差的幅值会十分显著. 然而,这个假设在特定情况下并不成立,基于这一事实,文献[17]于2009年提出了针对电力系统的虚假数据注入(false data injection, FDI)攻击策略. 文献[18]提出了基于部分节点系统参数信息的FDI攻击策略,降低了攻击实施的难度. 另外,文献[19]指出仅通过对量测数据进行分析,也可以实现FDI攻击. 注意到,针对集中式框架下FDI攻击策略的研究受到了学术界的广泛关注,但是还未发现关于分布式框架下FDI攻击的安全研究. 本文从攻击者的角度出发,提出了分布式隐匿FDI攻击方法,使得在不触发坏值检测的情况下仍可以损害分布式估计性能. 由于这种攻击的不可检测性,当攻击者拥有分布式系统的部分节点的参数信息时,能够任意破坏目标系统的估计值. 虽然大多数系统都有严格的信息保密策略,但是恶意的攻击者依然能够通过黑客技术或系统辨识方法获取系统参数信息. 值得注意的是,工业系统由于长期处于封闭的安全环境,数据的安全保护措施并不完善,因此遭受攻击的可能性很大. 本文针对一类分布式最小二乘估计算法,结合分布式系统中现有的坏值检测方法,设计了依赖于部分节点系统信息的分布式隐匿FDI攻击方法. 最后,通过在IEEE 118-Bus电力系统上的仿真结果,验证了所设计FDI攻击方法的隐匿性和有效性.

1 问题描述

1.1 分布式最小二乘估计

针对如下线性模型进行估计:

$$\boldsymbol{z}=\boldsymbol{H}\boldsymbol{x}+\boldsymbol{e}. \tag{1}$$

其中:$\boldsymbol{z}\in\boldsymbol{C}^m$为系统的观测向量,$\boldsymbol{x}\in\boldsymbol{C}^n$为系统的状态向量,$\boldsymbol{H}\in\boldsymbol{C}^{m\times n}$为系统参数矩阵,$\boldsymbol{e}\in\boldsymbol{C}^m$为观测向量中的噪声. 目标是在已知系统观测$\boldsymbol{z}$和系统参数矩阵$\boldsymbol{H}$的情况下,估计系统的状态向量$\boldsymbol{x}$.

本文利用加权最小二乘法[20]来估计状态$\boldsymbol{x}$,即通过优化如下目标:

$$J(x)=[\boldsymbol{z}-\boldsymbol{H}\boldsymbol{x}]^*\boldsymbol{R}^{-1}[\boldsymbol{z}-\boldsymbol{H}\boldsymbol{x}] \tag{2}$$

达到最小值,得到加权最小二乘估计器

$$\hat{\boldsymbol{x}}=(\boldsymbol{H}^*\boldsymbol{R}^{-1}\boldsymbol{H})^{-1}\boldsymbol{H}^*\boldsymbol{R}^{-1}\boldsymbol{z}=\boldsymbol{\Psi}^{-1}\boldsymbol{\alpha}. \tag{3}$$

其中:$\hat{\boldsymbol{x}}$是对系统真实状态$\boldsymbol{x}$的估计结果;$\boldsymbol{\Psi}=\boldsymbol{H}^*\boldsymbol{R}^{-1}\boldsymbol{H}$,$\boldsymbol{\alpha}=\boldsymbol{H}^*\boldsymbol{R}^{-1}\boldsymbol{z}$;矩阵$\boldsymbol{R}\in\boldsymbol{C}^{m\times m}$是对角权重矩阵,通常取系统观测向量的协方差矩阵;$\boldsymbol{A}^*$表示矩阵$\boldsymbol{A}$的共轭转置. 该估计器的特点是计算简单且计算结果稳定,在$\boldsymbol{H}$列满秩,即$\mathrm{rank}(\boldsymbol{H})=n$时,估计结果是唯一的. 一些其他的估计准则,如极大似然准则,如果观测误差服从均值为零的正态分布,则将导出相同的最优状态估计器(3).

引入文献[9]所提出的分布式估计方法. 在分布式框架下,系统被划分为N个相互关联的节点,每个节点只能获取本地的观测信息. 设节点$i\in N$有d_i个状态和p_i个观测,记每个节点的本地状态向量为$\boldsymbol{x}_i\in\boldsymbol{C}^{d_i}$,本地观测向量为$\boldsymbol{z}_i\in\boldsymbol{C}^{p_i}$,本地观测的误差向量为$\boldsymbol{e}_i\in\boldsymbol{C}^{p_i}$,$\boldsymbol{e}_i$服从独立正态分布$\mathcal{N}(0,\boldsymbol{R}_i)$,其中$\boldsymbol{R}_i\in\boldsymbol{C}^{d_i\times d_i}$. 则分布式系统中各节点的模型可以表示为

$$\boldsymbol{z}_i=\sum_{j=1}^{N}\boldsymbol{H}_{i,j}\boldsymbol{x}_j+\boldsymbol{e}_i. \tag{4}$$

各节点的邻居关系由定义1给出.

定义1 定义集合$l_i=\{j\,|\,\boldsymbol{H}_{j,i}\neq\boldsymbol{0}\}$,$l_i$为本地观测值与节点$i$的状态存在关联的节点;定义集合

$\Theta_i = \{j \mid \boldsymbol{H}_{i,j} \neq \boldsymbol{0}\}$, Θ_i为本地状态与节点i的观测存在关联的节点;定义集合$\aleph_i = l_i \bigcup \Theta_i$,称$\aleph_i$为节点$i$的邻居节点.

引入预处理矩阵$\boldsymbol{\Pi}$,并将式(3)表示成如下线性方程形式:

$$\boldsymbol{\Upsilon}\hat{\boldsymbol{y}} = \tilde{\boldsymbol{\alpha}}. \tag{5}$$

其中:$\boldsymbol{\Upsilon} = \boldsymbol{\Pi}^{1/2}\boldsymbol{\Psi}\boldsymbol{\Pi}^{1/2}$, $\tilde{\boldsymbol{\alpha}} = \boldsymbol{\Pi}^{1/2}\boldsymbol{\alpha}$, $\boldsymbol{\Pi}^{1/2}\hat{\boldsymbol{y}} = \hat{\boldsymbol{x}}$.

利用Richardson方法[11],得到式(5)的渐近迭代估计公式

$$\hat{\boldsymbol{x}}(t+1) = (\boldsymbol{I} - \gamma\boldsymbol{\Pi}\boldsymbol{\Psi})\hat{\boldsymbol{x}}(t) + \gamma\boldsymbol{\Pi}\boldsymbol{\alpha}. \tag{6}$$

式(6)在各节点中的分布式形式可以表示为

$$\hat{\boldsymbol{x}}_i(t+1) = \hat{\boldsymbol{x}}_i(t) - \gamma\boldsymbol{\Pi}_i[\boldsymbol{\Psi}\hat{\boldsymbol{x}}(t)]_i + \gamma\boldsymbol{\Pi}_i\boldsymbol{\alpha}_i, \tag{7}$$

其中γ是松弛系数. 在分布式系统中每个节点只能获取本地的系统参数信息($\boldsymbol{H}_i$矩阵)和本地的观测向量($\boldsymbol{z}_i$向量),因此$\boldsymbol{\alpha}_i$和$\boldsymbol{\Psi}$是各节点的未知参数,需要通过节点之间进行信息交换计算出这两个参数的完整信息. 令$\boldsymbol{\alpha}_i^{(k)} = \boldsymbol{H}_{k,i}^{*}\boldsymbol{R}_k^{-1}\boldsymbol{z}_k$, $\boldsymbol{\Psi}_{i,j}^{(k)} = \boldsymbol{H}_{k,i}^{*}\boldsymbol{R}_k^{-1}\boldsymbol{H}_{k,j}$,可以得到

$$\boldsymbol{\alpha}_i = \sum_{k\in l_i}\boldsymbol{\alpha}_i^{(k)}, \tag{8}$$

$$[\boldsymbol{\Psi}\hat{\boldsymbol{x}}(t)]_i = \sum_{k\in l_i}\sum_{j\in\Theta_k}\boldsymbol{\Psi}_{i,j}^{(k)}\hat{\boldsymbol{x}}_j(t). \tag{9}$$

将式(8)和(9)代入(7),得到分布式最小二乘估计器

$$\hat{\boldsymbol{x}}_i(t+1) = \hat{\boldsymbol{x}}_i(t) - \gamma\,\boldsymbol{\Pi}_i\Big(\sum_{k\in l_i}^{N}\sum_{j\in\Theta_k}\boldsymbol{\Psi}_{i,j}^{(k)}\hat{\boldsymbol{x}}_j(t) - \sum_{k\in l_i}\boldsymbol{\alpha}_i^{(k)}\Big). \tag{10}$$

式(10)中估计器的松弛系数γ和预处理矩阵$\boldsymbol{\Pi}$是两个优化参数,其中$\boldsymbol{\Pi}_i$的近似最优条件为

$$\boldsymbol{\Pi}_i = \Big(\sum_{k\in l_i}\boldsymbol{\Psi}_{i,i}^{(k)}\Big)^{-1}, \tag{11}$$

γ的最优取值为

$$\gamma_{\mathrm{opt}} = \frac{2}{\lambda_{\min}^{(\boldsymbol{\Upsilon})} + \lambda_{\max}^{(\boldsymbol{\Upsilon})}}, \tag{12}$$

$\lambda_{\min}^{(\boldsymbol{\Upsilon})}$和$\lambda_{\max}^{(\boldsymbol{\Upsilon})}$分别为$\boldsymbol{\Upsilon}$的最小和最大特征值.

1.2　分布式坏值检测策略

加权最小二乘算法可以抑制观测向量中存在的高斯噪声,但是一些异常的观测数据,如观测值损坏、观测值被非法篡改时产生的坏值,会使估计结果不准确,因此需要利用坏值检测技术保护最小二乘估计不受异常数据的影响. 常用的检测方法为坏数据检测方法,通过判断观测残差的范数(通常使用二范数)$\|\boldsymbol{z} - \boldsymbol{H}\hat{\boldsymbol{x}}\|$是否超出阈值$\tau$,检测观测向量中是否存在坏值. 当观测数据中的噪声服从正态分布时,$\|\boldsymbol{z} - \boldsymbol{H}\hat{\boldsymbol{x}}\|^2$服从$\chi^2(m-n)$分布[21],因此可以通过置信度为$\beta$的假设检验设定阈值$\tau$,即令$\tau_\beta = \chi_\beta^2(m-n)$,则$\|\boldsymbol{z} - \boldsymbol{H}\hat{\boldsymbol{x}}\|^2 \geqslant \tau_\beta^2$的概率为$\beta$.

在分布式框架下,令节点$k \in N$的阈值τ_{β_k}取值为$\chi_{\beta_k}^2(p_k - q_k)$,若不等式

$$\|\boldsymbol{z}_k - \boldsymbol{H}_k\hat{\boldsymbol{x}}\| \geqslant \tau_{\beta_k} \tag{13}$$

成立,则该节点的观测数据中存在坏值. 由于任意节点的本地自由度$p_k - q_k$必定小于全局自由度$m-n$,根据卡方分布的性质,该节点的坏值检测阈值也必定小于集中式坏值检测阈值,加大了攻击者注入攻击的难度[16].

针对上述的分布式最小二乘估计器与坏值检测方法,本文提出了一种隐匿FDI攻击策略以达到如下目标:通过向部分节点的观测值中注入虚假数据,在保证对分布式坏值检测器隐匿的情况下使分布式最小二乘估计结果以预设的偏移量偏离真实值.

2　分布式隐匿FDI攻击策略

在给出具体的攻击构造方法前,对符号变量给出如下定义:将受攻击后的观测向量记为$\boldsymbol{z}^{\mathrm{FDI}} = \boldsymbol{z} + \boldsymbol{a}$. 其中:$\boldsymbol{a} = (\boldsymbol{a}_1^{\mathrm{T}}, \boldsymbol{a}_2^{\mathrm{T}}, \ldots, \boldsymbol{a}_N^{\mathrm{T}})^{\mathrm{T}}$为攻击者向观测中注入的攻击向量,$\boldsymbol{a}_k \in \boldsymbol{C}^{q_k}$为向第$k$个节点的观测注入的攻击向量;$\boldsymbol{z} = (\boldsymbol{z}_1^{\mathrm{T}}, \boldsymbol{z}_2^{\mathrm{T}}, \ldots, \boldsymbol{z}_N^{\mathrm{T}})^{\mathrm{T}}$为受攻击前的观测向量. 将受到攻击后的估计向量记为$\hat{\boldsymbol{x}}^{\mathrm{FDI}} = \hat{\boldsymbol{x}} + \boldsymbol{c}$. 其中:$\hat{\boldsymbol{x}} = (\hat{\boldsymbol{x}}_1^{\mathrm{T}}, \hat{\boldsymbol{x}}_2^{\mathrm{T}}, \ldots, \hat{\boldsymbol{x}}_N^{\mathrm{T}})^{\mathrm{T}}$为未受攻击时的估计向量;$\boldsymbol{c} = (\boldsymbol{c}_1^{\mathrm{T}}, \boldsymbol{c}_2^{\mathrm{T}}, \ldots, \boldsymbol{c}_N^{\mathrm{T}})^{\mathrm{T}}$为攻击者向估计结果中注入的偏移向量,由攻击者篡改观测数据导致. 设攻击者针对M个节点进行攻击,受攻击节点的下标组成集合$\mathcal{I}$. 对于$i \in \mathcal{I}$,攻击者根据要达到的攻击效果设计针对估计结果的偏移向量$\boldsymbol{c}_i$.

定义2　定义节点中与其他节点不存在关联的状态为该节点的非边界状态,节点i中所有非边界状态的下标构成的集合记为$\mathcal{L}^i$. 对于任意节点$k \neq i$,有$\boldsymbol{H}_k\bar{\boldsymbol{x}}^i = 0$. 其中$\bar{\boldsymbol{x}}^i \in \boldsymbol{C}^n$表示节点$i$的非边界状态向量,对于$\forall j \in \mathcal{L}^i$,有$\bar{\boldsymbol{x}}_j^i \neq 0$.

假设1　攻击者有能力获取节点$i \in \mathcal{I}$的系统参数信息,且能够监控和篡改这些节点中的部分观测值.

注1　虽然大多数系统都有严格的信息保密策略,但是恶意的攻击者依然能够通过黑客技术或系统辨识方法获取系统部分节点的参数信息;另外,攻击者可以通过物理入侵或网络入侵的方式窃听和篡改测量仪表,因此假设1是合理的.

根据假设1,攻击者拥有受攻击节点$k \in \mathcal{I}$的系统参数矩阵$\boldsymbol{H}_k$,并且能够篡改观测值$\boldsymbol{z}_k$. 在此情形下,定理1给出分布式FDI攻击策略.

定理1 若攻击者针对目标节点$k\in\mathcal{I}$设计的攻击向量满足$\boldsymbol{a}_k=\boldsymbol{H}_k\bar{\boldsymbol{c}}$,其中$\bar{\boldsymbol{c}}$是攻击者欲向非边界状态估计值注入的偏移向量,则向观测数据中注入虚假数据$\boldsymbol{z}_k^{\rm FDI}=\boldsymbol{z}_k+\boldsymbol{a}_k$可以欺骗分布式坏值检测器(13).

证明 将攻击者向各节点的观测中注入的攻击向量增广为向全局观测中注入的攻击向量(未受攻击的节点补0),可以表示为如下形式:

$$\bar{\boldsymbol{a}}=(\boldsymbol{0}^{\rm T},\boldsymbol{a}_{\mathcal{I}_1}^{\rm T},\boldsymbol{0}^{\rm T},\dots,\boldsymbol{0}^{\rm T},\boldsymbol{a}_{\mathcal{I}_M}^{\rm T},\boldsymbol{0}^{\rm T})^{\rm T}.$$

若攻击者针对目标节点$i\in\mathcal{I}$的非边界状态设计偏移向量$\bar{\boldsymbol{c}}^i\in\boldsymbol{C}^n$,对于$\forall j\in\mathcal{L}^i$,有$\bar{\boldsymbol{c}}_j^i\neq 0$,则根据定义2,对于任意节点$k\neq i$,有$\boldsymbol{H}_k\bar{\boldsymbol{c}}^i=\boldsymbol{0}$. 攻击者向全局状态估计值中注入的偏移向量可以表示为$\bar{\boldsymbol{c}}=\sum\limits_{i\in\mathcal{I}}\bar{\boldsymbol{c}}^i$,可以导出

$$\boldsymbol{H}\bar{\boldsymbol{c}}=\begin{bmatrix}\boldsymbol{H}_1\sum\limits_{i\in\mathcal{I}}\bar{\boldsymbol{c}}^i\\ \boldsymbol{H}_2\sum\limits_{i\in\mathcal{I}}\bar{\boldsymbol{c}}^i\\ \vdots\\ \boldsymbol{H}_N\sum\limits_{i\in\mathcal{I}}\bar{\boldsymbol{c}}^i\end{bmatrix}=\begin{bmatrix}\sum\limits_{i\in\mathcal{I}}\boldsymbol{H}_1\bar{\boldsymbol{c}}^i\\ \sum\limits_{i\in\mathcal{I}}\boldsymbol{H}_2\bar{\boldsymbol{c}}^i\\ \vdots\\ \sum\limits_{i\in\mathcal{I}}\boldsymbol{H}_N\bar{\boldsymbol{c}}^i\end{bmatrix}=$$

$$\begin{bmatrix}\boldsymbol{0}\\ \sum\limits_{i\in\mathcal{I}}\boldsymbol{H}_{\mathcal{I}_1}\bar{\boldsymbol{c}}^i\\ \boldsymbol{0}\\ \vdots\\ \boldsymbol{0}\\ \sum\limits_{i\in\mathcal{I}}\boldsymbol{H}_{\mathcal{I}_M}\bar{\boldsymbol{c}}^i\\ \boldsymbol{0}\end{bmatrix}=\begin{bmatrix}\boldsymbol{0}\\ \boldsymbol{H}_{\mathcal{I}_1}\bar{\boldsymbol{c}}\\ \boldsymbol{0}\\ \vdots\\ \boldsymbol{0}\\ \boldsymbol{H}_{\mathcal{I}_M}\bar{\boldsymbol{c}}\\ \boldsymbol{0}\end{bmatrix}=\bar{\boldsymbol{a}}.$$

假设第k个节点的坏值检测阈值为τ_{β_k},且当前观测值中没有坏值产生,那么第k个节点的分布式坏值检测器计算的残差为

$$\begin{aligned}&\|\boldsymbol{z}_k^{\rm FDI}-\boldsymbol{H}_k\hat{\boldsymbol{x}}^{\rm FDI}\|=\\&\|\boldsymbol{z}_k+\boldsymbol{a}_k-\boldsymbol{H}_k[(\boldsymbol{H}^*\boldsymbol{R}^{-1}\boldsymbol{H})^{-1}\boldsymbol{H}^*\boldsymbol{R}^{-1}(\boldsymbol{z}+\bar{\boldsymbol{a}})]\|=\\&\|\boldsymbol{z}_k-\boldsymbol{H}_k\hat{\boldsymbol{x}}+\boldsymbol{a}_k-\boldsymbol{H}_k(\boldsymbol{H}^*\boldsymbol{R}^{-1}\boldsymbol{H})^{-1}\boldsymbol{H}^*\boldsymbol{R}^{-1}\boldsymbol{H}\bar{\boldsymbol{c}}\|=\\&\|\boldsymbol{z}_k-\boldsymbol{H}_k\hat{\boldsymbol{x}}+(\boldsymbol{a}_k-\boldsymbol{H}_k\bar{\boldsymbol{c}})\|=\\&\|\boldsymbol{z}_k-\boldsymbol{H}_k\hat{\boldsymbol{x}}\|\leqslant\tau_{\beta_k}.\end{aligned}\tag{14}$$

因此,在注入虚假数据后,各节点的坏值检测器计算出的残差值与未受攻击前相同,该攻击无法被坏值检测机制探测到. 受攻击后估计结果的偏移量为

$$\begin{aligned}\hat{\boldsymbol{x}}^{\rm FDI}-\hat{\boldsymbol{x}}=&(\boldsymbol{H}^*\boldsymbol{R}^{-1}\boldsymbol{H})^{-1}\boldsymbol{H}^*\boldsymbol{R}^{-1}\bar{\boldsymbol{a}}=\\&(\boldsymbol{H}^*\boldsymbol{R}^{-1}\boldsymbol{H})^{-1}\boldsymbol{H}^*\boldsymbol{R}^{-1}\boldsymbol{H}\bar{\boldsymbol{c}}=\bar{\boldsymbol{c}},\end{aligned}$$

因此受攻击后估计结果会以预设的偏移量偏离真实值. □.

注2 根据定理1,攻击者只需要盗取目标节点$k\in\mathcal{I}$的系统参数信息和观测信息,就能构造出隐匿的攻击向量,这降低了攻击者实施攻击的难度.

定理2 本文设计的分布式FDI攻击不影响分布式最小二乘估计器的收敛特性.

证明 设估计算法第t次迭代的估计结果与集中式估计结果$\hat{\boldsymbol{x}}_{\rm est}$之间的差值为$\boldsymbol{e}(t)$,结合式(3)和(6)可得

$$\begin{aligned}&\boldsymbol{e}(t+1)=\hat{\boldsymbol{x}}(t+1)-\hat{\boldsymbol{x}}_{\rm est}=\\&\hat{\boldsymbol{x}}(t)-\gamma\boldsymbol{\Pi}(\boldsymbol{\Psi}\hat{\boldsymbol{x}}(t)-\boldsymbol{\alpha})-\hat{\boldsymbol{x}}_{\rm est}=\\&\boldsymbol{e}(t)-\gamma\boldsymbol{\Pi}\boldsymbol{\Psi}(\hat{\boldsymbol{x}}(t)-\boldsymbol{\Psi}^{-1}\boldsymbol{\alpha})=\\&\boldsymbol{e}(t)-\gamma\boldsymbol{\Pi}\boldsymbol{\Psi}\boldsymbol{e}(t)=\\&(I-\gamma\boldsymbol{\Pi}\boldsymbol{\Psi})\boldsymbol{e}(t)=\\&(I-\gamma\boldsymbol{\Pi}\boldsymbol{\Psi})^{t+1}\boldsymbol{e}(0).\end{aligned}\tag{15}$$

通过选取合适的参数,可令估计误差趋于零. 根据式(11)和(12),γ与预处理矩阵$\boldsymbol{\Pi}$由系统参数矩阵$\boldsymbol{H}$确定,因此分布式最小二乘估计算法的收敛性与观测值$\boldsymbol{z}$无关. □

注3 根据定理2,攻击者实施攻击后分布式最小二乘估计的收敛速率不会发生变化,因此无法通过估计过程中的收敛曲线对攻击进行检测,增加了攻击的隐匿性.

根据以上讨论,分布式隐匿FDI攻击算法归纳如下:

算法1 分布式隐匿FDI攻击.

输入: 各节点的原始观测向量$\boldsymbol{z}_k,k\in\mathcal{I}$;攻击者预设的状态偏移量$\bar{\boldsymbol{c}}$.

1) 节点$k\in\mathcal{I}$,计算观测向量的偏移量$\boldsymbol{a}_k=H_k\bar{\boldsymbol{c}}$.

2) 向节点$k\in\mathcal{I}$的观测向量注入虚假数据$\boldsymbol{z}_k^{\rm FDI}=\boldsymbol{z}_k+\boldsymbol{a}_k$.

3) 节点$j\in\{1,2,\dots,N\}$,计算$\boldsymbol{\alpha}_i^{(j)}$,其中$i\in\Theta_j$,并将计算结果发送给节点$i$.

4) 节点$j\in\{1,2,\dots,N\}$接收到节点$i\in l_j$的数据后,计算$\boldsymbol{\alpha}_j^{\rm FDI}=\sum\limits_{i\in l_j}\boldsymbol{\alpha}_j^{(i)}$.

5) 节点$j\in\{1,2,\dots,N\}$,设初始本地状态估计值$\hat{\boldsymbol{x}}_j(1)=0$.

主循环: $t=1,2,\dots,T$.

6) 节点$j\in\{1,2,\dots,N\}$发送$\hat{\boldsymbol{x}}_j(t)$给节点$k\in\aleph_j$.

7) 节点$k\in\{1,2,\dots,N\}$接收到数据后,计算

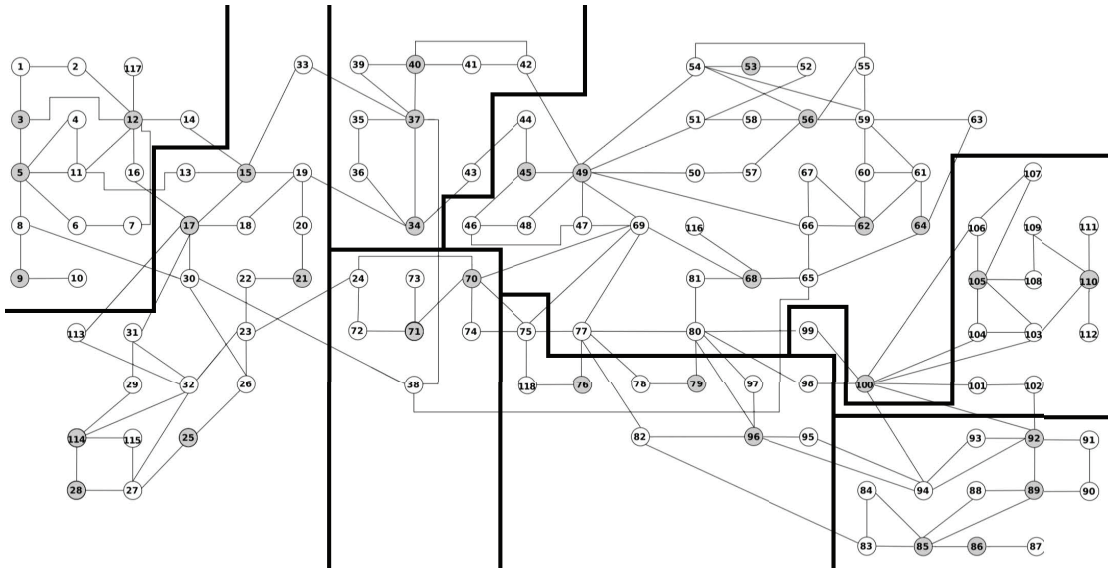

图 1　IEEE 118-bus电网拓扑

$$\check{\boldsymbol{x}}_{i,k}(t)=\sum_{j\in\Theta_k}\boldsymbol{\Psi}_{i,j}^{(k)}\hat{\boldsymbol{x}}_j(t),$$

其中$i\in\Theta_k$. 并将计算结果发送给节点i.

8) 节点$i\in\{1,2,\ldots,N\}$在接收到数据后估计状态

$$\hat{\boldsymbol{x}}_i(t+1)=\hat{\boldsymbol{x}}_i(t)-\gamma\Pi_i\Big(\sum_{k\in l_i}\check{\boldsymbol{x}}_{i,k}(t)-\boldsymbol{\alpha}_i^{\mathrm{FDI}}\Big).$$

3　实　验

本文利用IEEE 118-bus电力系统模型验证所提出算法的有效性. 仿真中,先以攻击者的角度,利用本文提出的分布式虚假数据注入攻击策略生成攻击向量,并将该攻击向量注入系统的测量数据中;其次,利用第2节描述的分布式最小二乘估计方法,结合文献[16]提出的分布式坏值检测方法,以防御者的角度检测系统中是否存在异常数据. 仿真结果显示,本文所构造的隐匿虚假数据攻击成功地欺骗了分布式系统中的坏值检测机制,并且可以使最小二乘估计结果以预设的偏移量偏离真实值.

3.1　仿真系统

利用文献[23]提供的IEEE 118-bus电力系统的真实数据,搭建电力系统的仿真环境[24],如图1所示. 电网的电力总线在图中以圆圈表示,其中使用SCADA (supervisory control and data acquisition)系统作为测量单元的总线以白色圆圈表示,使用PMU (phasor measurement unit)作为测量单元的总线以灰色圆圈表示,圆圈之间的连线表示各总线之间的传输线路. 根据电力总线的地理位置分布,利用聚类方法将电网划分为8个子系统,聚类结果如图1所示,不同子系统之间以黑色粗实线隔开. 仿真中,首先根据定理1利用部分节点的系统信息构造分布式隐匿FDI攻击,将攻击向量注入部分节点的测量值;其次,各相邻子系统之间在每个采样时刻通过分布式最小二乘算法进行数据交换,得到各自的估计结果;最后,使用分布式坏值检测方法,检测各子系统的残差是否超过设定的阈值.

3.2　隐匿FDI攻击仿真结果

本次仿真选取电网中的5个电压状态和5个电压相角状态作为攻击目标,所设计的状态估计偏移量如表1所示. 根据定理1设计针对各个子系统的隐匿攻击. 受攻击前后电网的分布式最小二乘估计如图2～图5所示. 图2和图3分别为未受攻击与受攻击后的电压估计曲线,图4和图5分别为未受攻击与受攻击后电压相角的估计曲线. 可以看出,分布式最小二乘估计的收敛过程在受攻击前后没有发生变化,由此验证了攻击不会改变所提出估计算法的收敛性和收敛速率. 受攻击前后的分布式最小二乘估计误差(状态估计值与真实状态值之间的偏差)分别如图6和图7所示. 未受到攻击时,由测量仪表的测量噪声引起

表 1　预设FDI偏移量

状态类型	状态编号	预设偏移量
电压/V	10	5
	50	4
	82	3.5
	178	3
	232	4.5
电压相角/rad	17	30
	71	30
	75	30
	107	30
	155	30

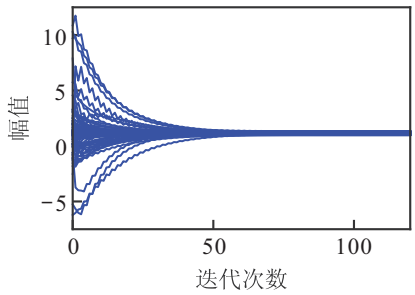

图 2 电压估计值(未受攻击情况)

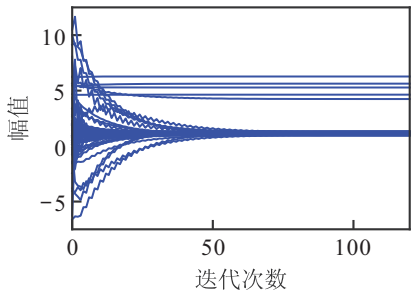

图 3 电压估计值(受攻击情况)

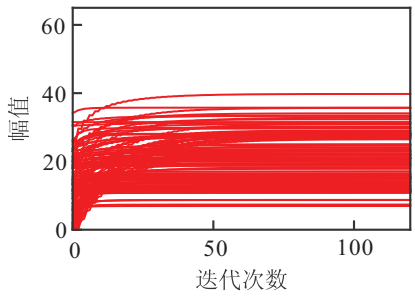

图 4 电压相角估计值 (未受攻击情况)

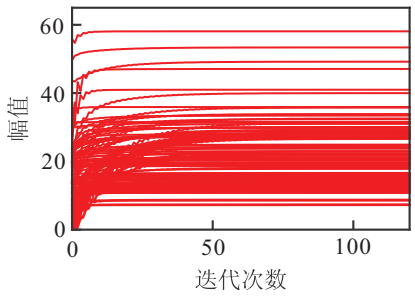

图 5 电压相角估计值 (受攻击情况)

的估计误差很小,如图6所示;在攻击者实施分布式隐匿FDI攻击后,对非攻击目标的状态估计值未产生影响,且受攻击的状态估计值按照表1中预设的偏差值偏离真实状态值,如图7所示.

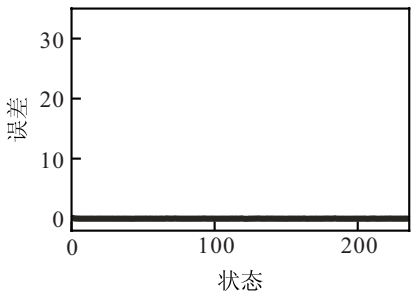

图 6 估计误差 (未受攻击情况)

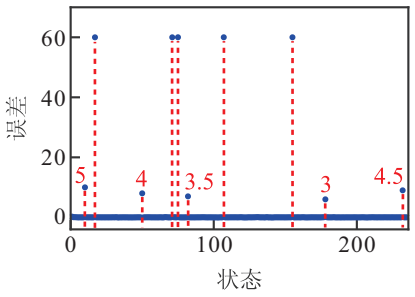

图 7 估计误差 (受攻击情况)

3.3 分布式坏值检测的仿真结果

系统受攻击前后的坏值检测结果分别如图8和图9所示,实线表示各子系统计算得到的测量残差,虚线表示各子系统的坏值检测阈值. 当某个子系统的测量残差超过该子系统的测量阈值时,表示该子系统的测量值中含有坏值. 对比受攻击前后的坏值检测结果,受攻击后每个子系统的测量残差几乎未发生变化,且均未超过坏值检测器的阈值,由此验证了本文所提出的FDI攻击策略可以欺骗分布式坏值检测器.

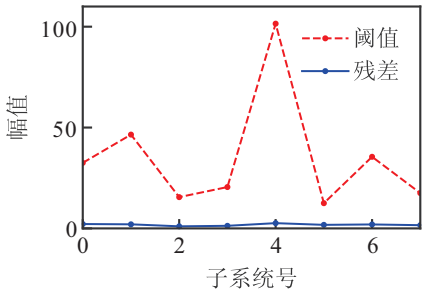

图 8 分布式坏值检测 (未受攻击情况)

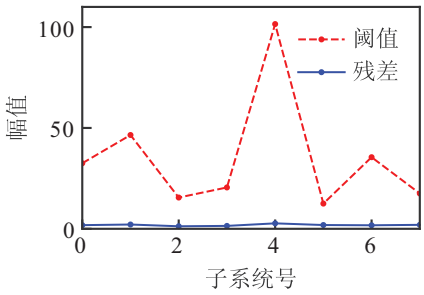

图 9 分布式坏值检测 (受攻击情况)

4 结 论

本文针对一类分布式最小二乘估计算法,利用已有的分布式坏值检测算法中存在的安全漏洞,通过构建合适的状态估计偏移向量,设计了一种只依赖于部分节点系统信息的分布式隐匿FDI攻击策略. 最后,通过IEEE 118-bus电力系统模型验证了这种攻击的有效性. 仿真结果表明现有的分布式坏值检测算法无法检测出本文设计的隐匿FDI攻击信号,且该攻击信号能够在保持隐匿的情况下使分布式最小二乘估计结果以预设的偏移量偏离真实状态值,具有很好的

隐匿性和目标性.

参考文献(References)

[1] Brice C W, Cavin R K. Multiprocessor static state estimation[J]. IEEE Transactions on Power Apparatus and Systems, 1982(2): 302-308.

[2] Kekatos V, Giannakis G B. Distributed robust power system state estimation[J]. IEEE Transactions on Power Systems, 2013, 28(2): 1617-1626.

[3] Cosovic M, Vukobratovic D. Distributed Gauss–Newton method for state estimation using belief propagation[J]. IEEE Transactions on Power Systems, 2019, 34(1): 648-658.

[4] Lin C H, Wu W C, Guo Y. Decentralized robust state estimation of active distribution grids incorporating microgrids based on PMU measurements[J]. IEEE Transactions on Smart Grid, 2020, 11(1): 810-820.

[5] Dong S J, Liu T, Wang W, et al. Identification of discrete-time output error model for industrial processes with time delay subject to load disturbance[J]. Journal of Process Control, 2017, 50: 40-55.

[6] Conejo A J, Sebastin D L T, Canas M. An optimization approach to multiarea state estimation[J]. IEEE Transactions on Power Systems, 2007, 22(1): 213-221.

[7] Xie L, Choi D H, Kar S, et al. Fully distributed state estimation for wide-area monitoring systems[J]. IEEE Transactions on Smart Grid, 2012, 3(3): 1154-1169.

[8] Pasqualetti F, Carli R, Bullo F. Distributed estimation via iterative projections with application to power network monitoring[J]. Automatica, 2012, 48(5): 747-758.

[9] Marelli D E, Fu M Y. Distributed weighted least-squares estimation with fast convergence for large-scale systems[J]. Automatica, 2015, 51: 27-39.

[10] Gómez-Expósito A, de la Villa Jaén A, Gómez-Quiles C, et al. A taxonomy of multi-area state estimation methods[J]. Electric Power Systems Research, 2011, 81(4): 1060-1069.

[11] Bertsekas D P, Tsitsiklis J N. Parallel and distributed computation: Numerical methods[M]. Englewood Cliffs: Prentice Hall, 1989: 134-135.

[12] Clements K A, Krumpholz G R, Davis P W. Power system state estimation residual analysis: An algorithm using network topology[J]. IEEE Transactions on Power Apparatus and Systems, 1981(4): 1779-1787.

[13] Choi D H, Xie L. Fully distributed bad data processing for wide area state estimation[C]. 2011 IEEE International Conference on Smart Grid Communications (SmartGridComm). Brussels: IEEE, 2011: 546-551.

[14] Ho C H, Wu H C, Chan S C, et al. A robust statistical approach to distributed power system state estimation with bad data[J]. IEEE Transactions on Smart Grid, 2019, 11(1): 517-527.

[15] Ren P, Abur A. Avoiding divergence in multi-area state estimation[J]. IEEE Transactions on Power Systems, 2019, 34(4): 3178-3187.

[16] Wang D, Guan X, Liu T, et al. Extended distributed state estimation: A detection method against tolerable false data injection attacks in smart grids[J]. Energies, 2014, 7(3): 1517-1538.

[17] Liu Y, Ning P, Reiter M K. False data injection attacks against state estimation in electric power grids[J]. ACM Transactions on Information and System Security (TISSEC), 2011, 14(1): 1-33.

[18] Rahman M A, Mohsenian-Rad H. False data injection attacks with incomplete information against smart power grids[C]. 2012 IEEE Global Communications Conference (GLOBECOM). Anaheim: IEEE, 2012: 3153-3158.

[19] Yu Z H, Chin W L. Blind false data injection attack using PCA approximation method in smart grid[J]. IEEE Transactions on Smart Grid, 2015, 6(3): 1219-1226.

[20] Meyer C D. Matrix analysis and applied linear algebra[M]. Philadelphia: Society for Industrial and Applied Mathematics, 2000: 429-446.

[21] Wood A J, Wollenberg B. Power generation operation and control[M]. The 2nd Edition. Washington DC: Wiley, 1996: 435-436.

[22] Benzi M. Preconditioning techniques for large linear systems: A survey[J]. Journal of Computational Physics, 2002, 182(2): 418-477.

[23] Christie Rich. 118 bus power flow test case[DB/OL]. [2019-12-02]. http://www.ee.washington.edu/research/pstca/pf118/pg_tca118bus.htm.

[24] Monticelli A. State estimation in electric power systems: A generalized approach[M]. Des Moines: Springer Science & Business Media, 2012: 63-100.

作者简介

胡明南 (1997–), 男, 硕士生, 从事信息物理系统安全估计的研究, E-mail: huminant@gmail.com;

陈博 (1984–), 男, 教授, 博士, 从事信息融合、安全估计与控制等研究, E-mail: bchen@zjut.edu.cn;

俞立 (1961–), 男, 教授, 博士生导师, 从事网络控制、信息融合等研究, E-mail: lyu@zjut.edu.cn.

(责任编辑: 孙艺红)