信息与控制　2018 年　第 47 卷　第 1 期：16 ~ 21

DOI：10.13976/j.cnki.xk.2018.0016　　文章编号：1002 - 0411(2018) - 01 - 0016 - 06

网络化 CPS 的优化、监控与安全专刊

基于自适应卡尔曼滤波器的智能电网隐蔽假数据攻击检测

罗小元[1]，朱鸣皋[1]，王新宇[1]，关新平[2]

1. 燕山大学电气工程学院，河北 秦皇岛　066004；　2. 上海交通大学电信学院，上海　200240

基金项目：河北省自然科学基金资助项目(2016203119)

通信作者：罗小元，xyluo@ysu.edu.cn　收稿/录用/修回：2017 - 10 - 12/2018 - 01 - 29/2018 - 02 - 08

摘要

基于自适应卡尔曼滤波器研究了智能电网假数据注入攻击检测问题. 首先通过给出电网的状态空间模型，对两种攻击进行描述：随机攻击和假数据注入攻击. 其次，研究表明常用的卡方检测器可以有效检测随机攻击，但无法检测出隐蔽假数据注入攻击. 由此进一步在分析假数据注入攻击隐蔽性的基础上，考虑实际噪声变化的影响，提出基于 $\boldsymbol{R}$，$\boldsymbol{Q}$ 自适应动态估计的卡尔曼滤波器的方法检测此类攻击，给出其设计过程，并对隐蔽假数据注入攻击提出了检测判断方法. 最后进行 Matlab 仿真实验，验证了提出的隐蔽假数据注入攻击检测方法的有效性.

关键词

智能电网

隐蔽假数据注入攻击

攻击检测

自适应卡尔曼滤波器

中图法分类号：TP273.1

文献标识码：A

Adaptive Kalman Filter-based Detection of Covert False Data Injection Attacks in Smart Grids

LUO Xiaoyuan[1], ZHU Minggao[1], WANG Xinyu[1], GUAN Xinping[2]

1. *School of Electrical Engineering*, *Yanshan University*, *Qinhuangdao* 066004, *China*;

2. *School of Electronic*, *Information and Electrical Engineering*, *Shanghai Jiao Tong University*, *Shanghai* 200240, *China*

Abstract

The detection problem of covert false data injection attack (FDIA) for smart grids by using adaptive Kalman filter (AKF) is herein presented. The state-space model of the smart grid is employed to study the characteristics of the random attack and the FDIA. However, an accurate analysis revealed that the commonly used χ^2-detector could only detect random attack efficiently, but not the FDIA. To overcome this drawback, the AKF is applied to assess FDIA by analyzing the covert characteristics of the FDIA and the change of actual noise. Finally, the simulation experiments verify the effectiveness of the proposed FDIA detection method.

Keywords

smart grid;

covert false data injection attack;

attack detection;

adaptive Kalman filter

0　引言

随着计算机、网络通讯、自动控制等新兴产业的相互交融和促进，信息物理融合网络系统(cyber-physical system, CPS)随之出现，并朝着大规模、智能化的方向发展. 智能电网是一种典型的 CPS 系统，但由于其开放性与智能性的日益增强，导致其安全风险不断增加[1-2]. 最近发生在乌克兰的大规模停电事件和印度整体电网瘫痪事件[3]引起了社会各界对智能电网遭受攻击的关注. 面对各种恶意的网络攻击[4-6]，研究针对智能电网的假数据注入攻击检测具有现实意义.

电力系统状态估计欺诈性数据攻击首次由 Liu 等人提出[7]，该种攻击可以有效避开传统的不良数据检测方式，被普遍认为是针对电力系统状态估计的新型网络攻击. 与传统的物理攻击相比，FDIA(false data injection attack)可以多次发起而不被发现. 如果 FDIA 与物理攻击相互协调使用，由物理攻击引起的线路中断可以被掩盖[8]. 因此，隐蔽假数据注入攻击的检测对于电力系统的安全性具有重要意义. 文[9]对电网指标的不确定性问题进行了建模分析. 为了量化电网的潜在威胁，文[10]引入了两类不同类型的 FDIA，即稀疏攻击和小规模攻击. 在此基础上，文[11]在非线性交流动力模型上描述了线性攻击策略的局限性. 更多关于 FDIA 的资料请参见文[12].

对于虚假数据攻击的检测，文[13]提出了针对 FDIA 的检测方法：一种是确保战略选择所需要的基本测量的安全，另一种是独立验证战略性选择的状态变量的异常. 在

此基础上，文[14]提出 H_∞-范数检验法，该方法通过量测残差的局部异常可以有效地检测随机攻击. 但是，检验设定的阈值大小对检测精度的影响较大，应用中易出现漏检，而且无法检测具有隐蔽性的假数据注入攻击. 文[15]提出基于卡尔曼滤波器构造的欧几里德检测方法可以检测隐蔽假数据注入攻击，但是噪声等不确定因素引起的异常波动会影响系统的检测结果. 文[16]提出状态可达集来检测隐蔽假数据攻击，但是使用的标准卡尔曼滤波器仍旧没有考虑实际噪声变化的影响. 文[17]从处理系统参数变化角度考虑使用交互多模型卡尔曼滤波来对系统状态进行估计，这在一定程度上提高了估计精度，但是仍旧没有考虑噪声变化的影响，这样就会导致观测系统的状态误差增大，甚至使得滤波发散，因此噪声统计特性的变化必须加以考虑.

本文针对隐蔽假数据注入攻击的检测问题，提出一种考虑噪声统计特性变化的自适应卡尔曼滤波器的检测方法来降低噪声的影响. 该方法利用系统可测信息重建状态变量，然后构造状态检测残差与先验阈值做对比，由此给出攻击检测的判断方法. 最后进行仿真实验. 实验结果证明该种攻击检测方法可以有效地检测出隐蔽假数据注入攻击并且可以缩短检测时间.

1　状态空间模型和攻击隐蔽性分析

本节考虑基于三相正弦电压方程的状态空间模型[15]. 在此基础上，分析和给出随机攻击和隐蔽假数据注入攻击模型.

1.1　状态空间模型

文[18]的研究表明，在电力系统中的攻击或故障总是反映在电压、电流或相位的变化上. 以如图 1 所示的 3 电机 6 总线电网为例，在每个总线与发电机组连接处使用相量测量得到电压. 在每个测量点记录电网三相电压信号. 文[15]从三相电压公式推导得到如式(1)的单测量节点的状态空间模型，文[19]在此基础上对该模型进行补充，使之成为有 3 个测量节点的系统模型. 本文假设只有一个测量节点遭受攻击，因此只需使用文[15]中的单测量节点模型.

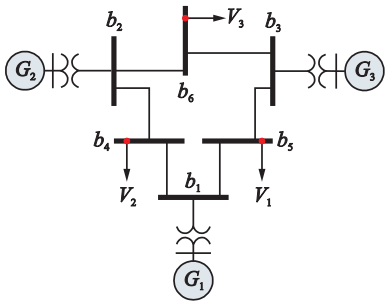

图 1　3 电机 6 总线电网结构图

Fig.1　Diagram of three-generator six-bus power grid

$$\begin{cases}\boldsymbol{x}(k+1)=\begin{pmatrix}1&0\\0&1\end{pmatrix}\boldsymbol{x}(k)+\tilde{\boldsymbol{\omega}}(k)\\ \boldsymbol{z}(k)=(\cos\omega k\ -\sin\omega k)\begin{pmatrix}x_1(k)\\x_2(k)\end{pmatrix}+\boldsymbol{\nu}(k)\end{cases}\tag{1}$$

式中 $\boldsymbol{x}(k)=[A_{\mathrm{V}}\cos\phi\quad A_{\mathrm{V}}\sin\phi]^{\mathrm{T}}$ 是系统状态；$\boldsymbol{z}(k)$是电压量测输出；A_{V} 是电压幅值，ϕ 是相位，ωk 是角频率；$\tilde{\boldsymbol{\omega}}(k)$是过程噪声，$\boldsymbol{\nu}(k)$代表量测噪声，并且假设这些噪声为 0 均值，标准差为 σ 的高斯白噪声. 方程(1)可以改写为

$$\begin{cases}\boldsymbol{x}(k+1)=\boldsymbol{A}\boldsymbol{x}(k)+\tilde{\boldsymbol{\omega}}(k)\\ \boldsymbol{z}(k)=\boldsymbol{H}\boldsymbol{x}(k)+\boldsymbol{\nu}(k)\end{cases}\tag{2}$$

式中 $\boldsymbol{A}=\begin{pmatrix}1&0\\0&1\end{pmatrix}\in\mathrm{R}^{n\times n}$ 为已知的系统矩阵，$\boldsymbol{H}=(\cos\omega k\quad -\sin\omega k)\in\mathrm{R}^{m\times n}$为输出矩阵. 假设 $\tilde{\boldsymbol{\omega}}(k)$，$\boldsymbol{\nu}(k)$为 0 均值，标准差为 σ 的高斯白噪声，且两者不相关. 即满足

$$\begin{cases}E(\tilde{\boldsymbol{\omega}}_k)=0,\ E(\tilde{\boldsymbol{\omega}}_k\tilde{\boldsymbol{\omega}}_j^{\mathrm{T}})=\boldsymbol{Q}_k\delta_{kj}\\ E(\boldsymbol{\nu}_k)=0,\ E(\boldsymbol{\nu}_k\boldsymbol{\nu}_j^{\mathrm{T}})=\boldsymbol{R}_k\delta_{kj}\\ E(\tilde{\boldsymbol{\omega}}_k\boldsymbol{\nu}_k^{\mathrm{T}})=0\end{cases}\tag{3}$$

式中 δ_{kj}是狄拉克函数，即 $\delta_{kj}=\begin{cases}0\ \text{if}\ k\neq j\\1\ \text{if}\ k=j\end{cases}$，$\boldsymbol{Q}_k$ 是过程噪声协方差矩阵，$\boldsymbol{R}_k$ 是测量噪声协方差矩阵.

1.2　攻击描述和问题提出

目前针对智能电网的攻击主要有重放攻击，拒绝服务攻击以及虚假数据注入攻击等. 本文主要考虑虚假数据注入攻击中的随机攻击和隐蔽假数据注入攻击.

假设攻击者可以控制一系列系统中的传感器数据，且攻击持续时间为 T.

1）在随机攻击情况下，攻击者无需精心构建攻击序列来克服检测机制，其主要目的就是破坏电网的正常运行. 最快速有效的方法就是直接攻击传感器导致量测输出错误. 随机攻击模型描述如下

$$\boldsymbol{z}_{\mathrm{a}}(k)=\begin{cases}\boldsymbol{H}\boldsymbol{x}(k)+\boldsymbol{\nu}(k) & k\not\subset T\\ \boldsymbol{H}\boldsymbol{x}_{\mathrm{a}}(k)+\boldsymbol{\nu}(k)+\boldsymbol{y}_{\mathrm{a}}(k) & k\subset T\end{cases}\tag{4}$$

式中 $\boldsymbol{y}_{\mathrm{a}}(k)$可以是任意的攻击信号，在后面的仿真部分，采用的是正弦信号.

2）在隐蔽假数据注入攻击情况下，攻击者需要精心构建攻击序列来克服检测机制，其次需要系统按照攻击者的期望运行. 这就要求攻击者已知系统信息，包括参数矩阵 $\boldsymbol{H}$[20].

考虑假数据注入攻击模型描述如下

$$\boldsymbol{z}_{\mathrm{a}}(k)=\begin{cases}\boldsymbol{H}\boldsymbol{x}(k)+\boldsymbol{\nu}(k) & k\not\subset T\\ \boldsymbol{H}\boldsymbol{x}_{\mathrm{a}}(k)+\boldsymbol{\nu}(k)+\boldsymbol{\Gamma}\boldsymbol{y}_{\mathrm{a}}(k) & k\subset T\end{cases}\tag{5}$$

式中 $\boldsymbol{\Gamma}=\mathrm{diag}(\gamma_1,\cdots,\gamma_n)$代表传感器选择矩阵，当 $\gamma_i=1$ 时代表第 i 个传感器遭受攻击，否则 $\gamma_i=0$. 且 $\boldsymbol{y}_{\mathrm{a}}(k)$是攻击者精心构建的攻击向量序列.

对攻击者克服检测机制的隐蔽性[21]进行分析.

忽略噪声和省略下标，由方程(2)中输出方程可得

$$\boldsymbol{z}=\boldsymbol{H}\boldsymbol{x}\tag{6}$$

式中 $\boldsymbol{z}$ 是系统量测数据的列向量，$\boldsymbol{H}$ 为状态变量和量测输出的线性关系.

无攻击时的残差为

$$\boldsymbol{r}=\boldsymbol{z}-\boldsymbol{H}\hat{\boldsymbol{x}}\tag{7}$$

假设 $\boldsymbol{a}$ 为构建的攻击向量，$\boldsymbol{c}$ 为遭受攻击时的状态误

差向量. 此时残差为

$$\begin{aligned}\boldsymbol{r}_{\mathrm{a}} &= \|(\boldsymbol{z}+\boldsymbol{a})-\boldsymbol{H}(\hat{\boldsymbol{x}}+\boldsymbol{c})\| \\ &= \|(\boldsymbol{z}-\boldsymbol{H}\hat{\boldsymbol{x}})+(\boldsymbol{a}-\boldsymbol{H}\boldsymbol{c})\| \\ &\leqslant \|\boldsymbol{z}-\boldsymbol{H}\hat{\boldsymbol{x}}\|+\|\boldsymbol{a}-\boldsymbol{H}\boldsymbol{c}\| \\ &= \boldsymbol{r}+\boldsymbol{\tau}_{\mathrm{a}} \\ &= \boldsymbol{r} \end{aligned} \tag{8}$$

式中 $\boldsymbol{\tau}_{\mathrm{a}}$ 为由虚假数据注入攻击导致的残差增量.

由式(8)可知, 当 $\boldsymbol{a}=\boldsymbol{H}\boldsymbol{c}$, 即 $\boldsymbol{a}$ 是列向量 $\boldsymbol{H}$ 的线性组合(此时的攻击称为隐蔽型假数据注入攻击), 此时的残差与没有攻击时的残差的误差处于正常范围, 这就导致隐蔽假数据注入攻击很难被成功检测. 本文尝试考虑通过基于自适应卡尔曼滤波器的方法来检测此种攻击.

2 自适应卡尔曼滤波器

本节首先对标准卡尔曼滤波器的缺陷进行分析, 然后在此基础上提出采用一种基于 $\boldsymbol{R}$, $\boldsymbol{Q}$ 动态估计的自适应卡尔曼滤波算法.

2.1 基于 $\boldsymbol{R}$, $\boldsymbol{Q}$ 动态估计的自适应卡尔曼滤波

为了得到系统部分的状态变量实现攻击检测, 考虑利用已知的输入与输出量, 通过标准卡尔曼滤波器对系统状态变量进行估计. 但是标准卡尔曼滤波的前提是噪声的统计特性必须已知, 但在结合实际时发现, 一般只能得到近似的统计特性, 这样就会导致观测系统的状态误差增大, 有时甚至会导致滤波发散.

因此基于上述分析, 本文提出采用一种基于 $\boldsymbol{R}$, $\boldsymbol{Q}$ 动态估计的自适应卡尔曼滤波算法, 该算法考虑了噪声统计特性的变化, 并利用观测数据实时更新, 从而达到实时修正噪声模型的目的. 目前已经有很多根据实际应用提出的自适应卡尔曼滤波算法, 本文采用一种改进的基于 $\boldsymbol{R}$, $\boldsymbol{Q}$ 动态估计的卡尔曼滤波算法. 本文首先假定 $\boldsymbol{R}$, $\boldsymbol{Q}$ 是未知的, 利用观测数据对系统噪声变化特性进行实时估计.

考虑系统的噪声统计特性是时变的, 即满足

$$\begin{cases} E(\tilde{\boldsymbol{\omega}}_k)=\boldsymbol{q}_k,\ E(\tilde{\boldsymbol{\omega}}_k\tilde{\boldsymbol{\omega}}_j^{\mathrm{T}})=\boldsymbol{Q}_k\delta_{kj} \\ E(\boldsymbol{\nu}_k)=\boldsymbol{r}_k,\ E(\boldsymbol{\nu}_k\boldsymbol{\nu}_j^{\mathrm{T}})=\boldsymbol{R}_k\delta_{kj} \\ E(\tilde{\boldsymbol{\omega}}_k\boldsymbol{\nu}_k^{\mathrm{T}})=0 \end{cases} \tag{9}$$

因此下一步需要求得时变噪声的 $\hat{\boldsymbol{q}}_k$, $\hat{\boldsymbol{Q}}_k$, $\hat{\boldsymbol{r}}_k$, $\hat{\boldsymbol{R}}_k$.

文[22]给出了次优无偏常值噪声统计估计器

$$\hat{\boldsymbol{q}}_k = \frac{1}{k}\sum_{i=0}^{k}[\hat{\boldsymbol{x}}_i-\boldsymbol{A}\hat{\boldsymbol{x}}_{i-1}] \tag{10}$$

$$\hat{\boldsymbol{Q}}_k = \frac{1}{k}\sum_{i=0}^{k}[\boldsymbol{L}_i\bar{\boldsymbol{z}}_i\bar{\boldsymbol{z}}_i^{\mathrm{T}}\boldsymbol{L}_i^{\mathrm{T}}+\boldsymbol{P}_i-\boldsymbol{A}\boldsymbol{P}_{i-1}\boldsymbol{A}^{\mathrm{T}}] \tag{11}$$

$$\hat{\boldsymbol{r}}_k = \frac{1}{k}\sum_{i=0}^{k}[\boldsymbol{z}_{i+1}-\boldsymbol{H}\hat{\boldsymbol{x}}_{i|i-1}] \tag{12}$$

$$\hat{\boldsymbol{R}}_k = \frac{1}{k}\sum_{i=0}^{k}[\bar{\boldsymbol{z}}\bar{\boldsymbol{z}}^{\mathrm{T}}-\boldsymbol{H}\boldsymbol{P}_{i|i-1}\boldsymbol{H}^{\mathrm{T}}] \tag{13}$$

式中 $\bar{\boldsymbol{z}}=\boldsymbol{z}_k-\boldsymbol{H}\hat{\boldsymbol{x}}_{k|k-1}$ 称为新息. 由上式中的每项权系数均为 $1/k$ 可知上式都为算术平均. 但是当噪声为时变时, 应强调新近数据的作用, 对此本文考虑采用指数加权的方法来实现.

选取加权系数 $\{\beta_i\}$ 满足:

$$\beta_i=\beta_{i-1}b,\quad 0<b<1,\ \sum_{i=0}^{k}\beta_i=1 \tag{14}$$

于是有

$$\begin{cases} \beta_i=d_kb^{i-1} \\ d_k=\dfrac{1-b}{1-b^{k+1}} \end{cases} \tag{15}$$

式中 b 是遗忘因子, 对式(10)~(13)用 β_{k+1-i} 代替 $1/k$ 便可以得到时变噪声估计器, 本文以 $\hat{\boldsymbol{q}}_k$ 为例.

$$\begin{aligned}\hat{\boldsymbol{q}}_k &= \beta_{k+1-i}\sum_{i=0}^{k}[\hat{\boldsymbol{x}}_i-\boldsymbol{A}\hat{\boldsymbol{x}}_{i-1}] \\ &= \sum_{i=0}^{k}d_kb^{k-i}[\hat{\boldsymbol{x}}_i-\boldsymbol{A}\hat{\boldsymbol{x}}_{i-1}] \\ &= d_k\sum_{i=0}^{k}b^{k-i}[\hat{\boldsymbol{x}}_i-\boldsymbol{A}\hat{\boldsymbol{x}}_{i-1}] \\ &= d_k[\hat{\boldsymbol{x}}_k-\boldsymbol{A}\hat{\boldsymbol{x}}_{k-1}]+d_k\sum_{i=0}^{k-1}b^{k-i}[\hat{\boldsymbol{x}}_i-\boldsymbol{A}\hat{\boldsymbol{x}}_{i-1}] \\ &= d_k[\hat{\boldsymbol{x}}_k-\boldsymbol{A}\hat{\boldsymbol{x}}_{k-1}]+\frac{d_kb}{d_{k-1}}\sum_{i=0}^{k-1}d_{k-1}b^{k-i-1}[\hat{\boldsymbol{x}}_i-\boldsymbol{A}\hat{\boldsymbol{x}}_{i-1}] \\ &= d_k[\hat{\boldsymbol{x}}_k-\boldsymbol{A}\hat{\boldsymbol{x}}_{k-1}]+(1-d_k)\sum_{i=0}^{k-1}d_{k-1}b^{k-i-1}[\hat{\boldsymbol{x}}_i-\boldsymbol{A}\hat{\boldsymbol{x}}_{i-1}] \\ &= (1-d_k)\hat{\boldsymbol{q}}_{k-1}+d_k[\hat{\boldsymbol{x}}_k-\boldsymbol{A}\hat{\boldsymbol{x}}_{k-1}] \end{aligned} \tag{16}$$

同理可得

$$\hat{\boldsymbol{Q}}_k=(1-d_k)\hat{\boldsymbol{Q}}_{k-1}+d_k[\boldsymbol{L}_i\bar{\boldsymbol{z}}\bar{\boldsymbol{z}}^{\mathrm{T}}\boldsymbol{L}_i^{\mathrm{T}}+\boldsymbol{P}_i-\boldsymbol{A}\boldsymbol{P}_{i-1}\boldsymbol{A}^{\mathrm{T}}] \tag{17}$$

$$\hat{\boldsymbol{r}}_k=(1-d_k)\hat{\boldsymbol{r}}_{k-1}+d_k[\hat{\boldsymbol{z}}_k-\boldsymbol{H}\boldsymbol{x}_{k|k-1}] \tag{18}$$

$$\hat{\boldsymbol{R}}_k=(1-d_k)\hat{\boldsymbol{R}}_{k-1}+d_k[\bar{\boldsymbol{z}}\bar{\boldsymbol{z}}^{\mathrm{T}}-\boldsymbol{H}\boldsymbol{P}_{i|i-1}\boldsymbol{H}^{\mathrm{T}}] \tag{19}$$

但在实际应用中发现, 每一步都进行迭代来估计噪声的统计特性会使得计算过程变得过于复杂. 因此考虑只在滤波发生异常时进行迭代, 滤波正常时则采用上一步的估计值.

由文[23], 得到根据新息来设置的滤波异常判据

$$\bar{\boldsymbol{z}}^{\mathrm{T}}\bar{\boldsymbol{z}}>\lambda\mathrm{tr}(E[\bar{\boldsymbol{z}}\bar{\boldsymbol{z}}^{\mathrm{T}}]) \tag{20}$$

式中 λ 为误差系数, 且 $\lambda\geqslant 1$, 当 $\lambda=1$ 时为严格收敛判据; tr 表示矩阵的迹. 当式(20)成立时, 表示此时的误差将会超过理论值的 λ 倍, 此时滤波会发散. 则需要对 $\hat{\boldsymbol{Q}}_k$ 和 $\hat{\boldsymbol{R}}_k$ 进行估计使之适应当前的滤波器. 反之, 式(20)不成立, 即此时的滤波处于正常情况, 无需对 $\hat{\boldsymbol{Q}}_k$ 和 $\hat{\boldsymbol{R}}_k$ 重新估计, 即 $\hat{\boldsymbol{Q}}_k=\hat{\boldsymbol{Q}}_{k-1}$, $\hat{\boldsymbol{R}}_k=\hat{\boldsymbol{R}}_{k-1}$.

最后综合标准卡尔曼滤波和时变噪声估计器就得到可以对时变噪声进行统计估计的自适应卡尔曼滤波算法.

3 攻击检测

本节分析常用的 χ^2 检测和检测残差法对隐蔽假数据注入攻击的检测原理.

3.1 χ^2 检测

文[24]中, χ^2 检测利用残差信号构建目标检测函数 $f(k)$, 并与设定的阈值 $\bar{f}$ 相比较.

χ^2 检测的目标函数给定为

$$f(k)=\boldsymbol{r}(k)^{\mathrm{T}}\boldsymbol{d}(k)\boldsymbol{r}(k) \tag{21}$$

式中 $\boldsymbol{r}(k)=\boldsymbol{z}_k-\hat{\boldsymbol{z}}_{k|k}$ 是残差信号，$\boldsymbol{d}(k)$ 是残差 $\boldsymbol{r}(k)$ 的协方差矩阵，根据协方差的定义可求得

$$\begin{aligned}\boldsymbol{d}(k)&=E[\boldsymbol{r}_k\boldsymbol{r}_k^{\mathrm{T}}]\\&=(\boldsymbol{H}-\boldsymbol{H}\boldsymbol{L}_k\boldsymbol{H})\boldsymbol{P}_k^{-}(\boldsymbol{H}-\boldsymbol{H}\boldsymbol{L}_k\boldsymbol{H})^{\mathrm{T}}+\\&\quad(\boldsymbol{I}-\boldsymbol{H}\boldsymbol{L}_k)\boldsymbol{R}(\boldsymbol{I}-\boldsymbol{H}\boldsymbol{L}_k)^{\mathrm{T}}\end{aligned}\tag{22}$$

在使用 χ^2 检测的情况下，为了最小化由系统噪声引起的误报率，设定阈值在 3σ（σ 是噪声的标准差），这样可以滤去 99.73% 的误报率[22]．然后根据式(23)判断攻击的存在

$$\begin{cases}f(k)\geqslant\bar{f}\rightarrow\text{有攻击}\rightarrow\text{触发警报}\\f(k)<\bar{f}\rightarrow\text{无攻击}\end{cases}\tag{23}$$

文[15]中，假数据注入攻击可以描述为：攻击持续时间 T 内，在攻击序列 $\boldsymbol{y}_{\mathrm{a}}(k)$ 的作用下，使得 $\lim\limits_{k\to T}\|\boldsymbol{x}_{\mathrm{a}}(k)-\boldsymbol{x}(k)\|\geqslant\alpha(\alpha>0)$，但是残差信号 $\lim\limits_{k\to T}\|\boldsymbol{r}_{\mathrm{a}}(k)\|\leqslant\varepsilon(0<\varepsilon<1)$，其中 $\boldsymbol{x}_{\mathrm{a}}(k)$ 和 $\boldsymbol{r}_{\mathrm{a}}(k)$ 是遭受攻击时系统的状态变量和残差．这就导致 χ^2 无法检测针对传感器的隐蔽假数据注入攻击．因此，在下一小节中从系统状态进行考虑分析．

3.2　检测残差分析

考虑自适应卡尔曼滤波器公式

$$\hat{\boldsymbol{x}}_{k|k}=\boldsymbol{A}\hat{\boldsymbol{x}}_{k-1|k-1}+\boldsymbol{B}\boldsymbol{u}_k+\boldsymbol{L}_k(\boldsymbol{z}_k-\boldsymbol{H}\hat{\boldsymbol{x}}_{k|k-1})\tag{24}$$

当有攻击发生时，首先变化的是量测输出 $\boldsymbol{z}_k$，假设它的变化量是 $\boldsymbol{a}$，即式(24)变为

$$\hat{\boldsymbol{x}}_{k|k}=\boldsymbol{A}\hat{\boldsymbol{x}}_{k-1|k-1}+\boldsymbol{B}\boldsymbol{u}_k+\boldsymbol{L}_k(\boldsymbol{z}_k+\boldsymbol{a}-\boldsymbol{H}\hat{\boldsymbol{x}}_{k|k-1})\tag{25}$$

从式(25)可知，由于 $\boldsymbol{a}$ 的注入，导致估计状态 $\hat{\boldsymbol{x}}_{k|k}$ 也发生了变化．假设变化量为 $\boldsymbol{c}$，式(25)变为

$$\begin{aligned}\hat{\boldsymbol{x}}_{k+1|k+1}&=\boldsymbol{A}(\hat{\boldsymbol{x}}_{k|k-1}+\boldsymbol{c})+\boldsymbol{B}\boldsymbol{u}_{k+1}+\boldsymbol{L}_k[\boldsymbol{z}_k+\boldsymbol{a}-\boldsymbol{H}(\hat{\boldsymbol{x}}_{k|k-1}+\boldsymbol{c})]\\&=\boldsymbol{A}\hat{\boldsymbol{x}}_{k|k-1}+\boldsymbol{A}\boldsymbol{c}+\boldsymbol{B}\boldsymbol{u}_{k+1}+\boldsymbol{L}_k(\boldsymbol{z}_k-\boldsymbol{H}\hat{\boldsymbol{x}}_{k|k-1})+\boldsymbol{L}_k(\boldsymbol{a}-\boldsymbol{H}\boldsymbol{c})\\&=\boldsymbol{A}\hat{\boldsymbol{x}}_{k|k-1}+\boldsymbol{B}\boldsymbol{u}_{k+1}+\boldsymbol{L}_k(\boldsymbol{z}_k-\boldsymbol{H}\hat{\boldsymbol{x}}_{k|k-1})+\boldsymbol{A}\boldsymbol{c}+\boldsymbol{L}_k(\boldsymbol{a}-\boldsymbol{H}\boldsymbol{c})\end{aligned}\tag{26}$$

式中 $\boldsymbol{a}=\boldsymbol{H}\boldsymbol{c}$ 在前面已经假定．比较式(24)和式(26)发现，由于攻击向量 $\boldsymbol{a}$ 的隐蔽性，使得基于残差的 χ^2 检测失效．但是因为 $\boldsymbol{A}\boldsymbol{c}$ 的存在，可以通过状态来进行检测．

当攻击向量 $\boldsymbol{a}$ 注入到传感器输出 $\boldsymbol{z}$ 时，观测状态会立即发生改变，但是因为框架结构导致攻击对输入的影响存在延迟[25]，因此系统会保持原来的状态，这就给检测该攻击提供了时间．

因此考虑用如下的基于状态分析的检测残差法来进行检测

$$\boldsymbol{g}(\boldsymbol{x},\hat{\boldsymbol{x}})=\frac{\|\boldsymbol{x}(k)-\hat{\boldsymbol{x}}(k)\|}{\|\boldsymbol{x}(k)\|\|\hat{\boldsymbol{x}}(k)\|}\tag{27}$$

式中 $\|\cdot\|$ 是向量模长．当没有攻击时，观测状态和系统状态是一致的，式(27)趋于0，当有虚假数据注入攻击发生时，即 $\lim\limits_{k\to T}\|\boldsymbol{x}_{\mathrm{a}}(k)-\boldsymbol{x}(k)\|\geqslant\alpha(\alpha>0)$ 时，检测目标函数 $\lim\limits_{k\to\infty}g(\boldsymbol{x},\hat{\boldsymbol{x}})>\rho(\rho>0)$．最后考虑过程噪声和量测噪声影响，阈值 $\bar{f}$ 选择与 χ^2 检测相同的 3σ．

$$\begin{cases}g(x,\hat{x})\geqslant\bar{f}\rightarrow\text{有攻击}\rightarrow\text{触发警报}\\g(x,\hat{x})<\bar{f}\rightarrow\text{无攻击}\end{cases}\tag{28}$$

通过上述分析，证明了从系统状态得到的检测残差可以有效检测虚假数据注入攻击．

4　仿真分析

为了验证所设计的检测残差法的检测效果，在 Matlab/Simulation 环境下进行仿真研究．考虑随机攻击和隐蔽假数据注入攻击，用 χ^2 检测和残差检测法分别进行检测．采用的仿真参数和初始值如表 1 所示．AKF 估计器的输入是频率为 60 Hz 的带有随机高斯白噪声的正弦电压信号，利用 Matlab 函数 randn( )产生正态分布的零均值噪声．

表 1　仿真参数

Tab.1　Simulation parameters

参数	数值	参数	数值
频率	60 Hz	$\boldsymbol{x}_2(0)$	0
幅值	1 V	$\boldsymbol{P}(0\|0)$	单位阵
$\boldsymbol{x}_1(0)$	0	σ	0.4

4.1　在随机攻击下的 χ^2 检测和检测残差法

本文考虑的随机攻击信号为正弦信号．应用 3σ 准则求得先验阈值 $\bar{f}=3\sigma=1.2$.

图 2 给出了系统实际状态和观测状态响应图．从图中可以看出，在刚开始时，两者之间存在一个较大的误差，但是很短时间内趋于一致．这表明设计的自适应卡尔曼滤波器能很好地实时跟随系统的实际状态．

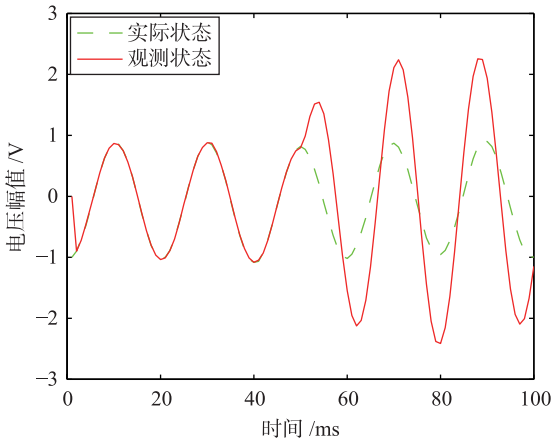

图 2　随机攻击情况下的实际值和观测值

Fig.2　Observed value and actual value under random attack

图 3 给出了随机攻击情况下的 χ^2 检测和检测残差响应图．由图 3 看出，在没有攻击时，χ^2 检测一直都处于 0，而检测残差受到噪声和估计误差的影响而围绕 0 上下小范围的波动．在 $t=50$ ms 时，实际状态和观测状态开始发生偏离，表明此时遭受攻击，且攻击持续时间 $T=50$．从图 3 可知，在随机攻击的情况下，χ^2 检测和状态检测残差法都能快速地超过设定的先验阈值，实现攻击检测．

4.2　隐蔽假数据攻击下的 χ^2 检测和检测残差法

图 4 是在隐蔽假数据注入攻击下的系统实际状态和观测状态响应图．

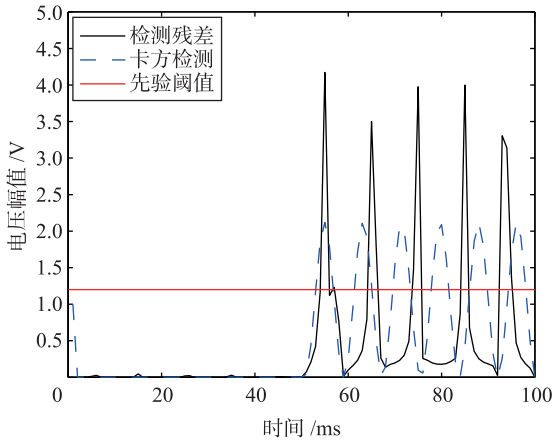

图 3 随机攻击情况下的χ^2 检测和检测残差

Fig.3 χ^2 detect and state residual under random attack

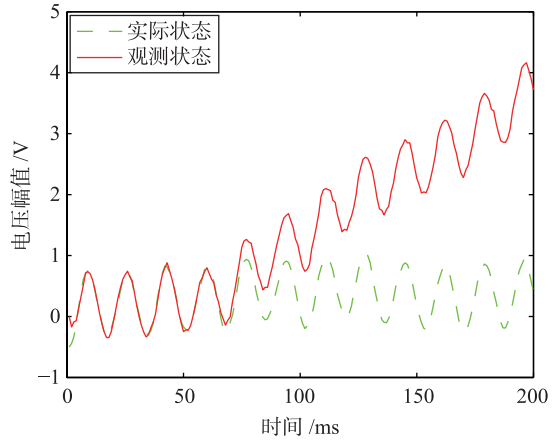

图 4 假数据注入攻击下的实际值和观测值

Fig.4 Observed value and actual value under FDIA

从图 4 中可看出，当 $t=60$ ms 时发生攻击，观测状态开始逐渐偏离实际状态，攻击持续时间 $T=140$. 图 5 给出了隐蔽假数据攻击情况下的χ^2 检测和文[15]中欧几里德检测响应图.

由图 5 看出，χ^2 检测没有明显的变化，且始终没有超过先验阈值，这表明χ^2 检测无法有效地检测出假数据注入攻击. 而基于卡尔曼滤波器构建的检测残差，当攻击发生时，它很快发生变化，且能在较短时间内迅速超过阈值检测出攻击. 但是使用 KF 滤波得到的残差信号受噪声影响很大，容易出现误报，而且检测所需要的时间也较长. 因此考虑使用基于 AKF 滤波的检测残差法来降低噪声影响，同时缩短检测时间.

如图 6 所示，使用本文提出的基于 AKF 滤波得到的残差信号在没有攻击时稳定在 0 附近，这大大降低了误报率，而且在此基础上使用的检测残差法可以在攻击发生后的极短时间内迅速超过设定的阈值，检测出隐蔽假数据攻击的存在，这与文[15]的欧几里德检测相比大大缩短了检测时间. 这验证了本文提出的方法的有效性.

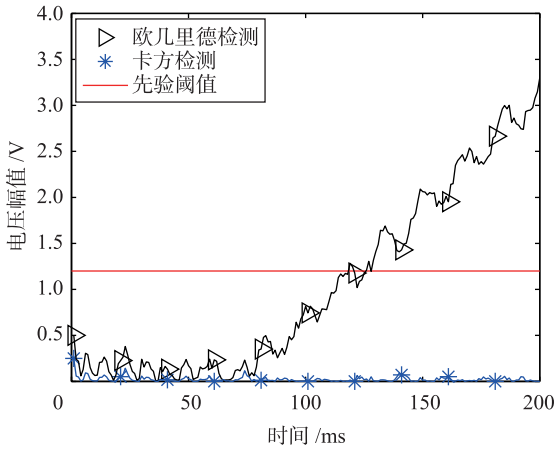

图 5 卡尔曼滤波下的χ^2 检测和欧几里德检测

Fig.5 χ^2 detect and Euclidean detect under FDIA using KF

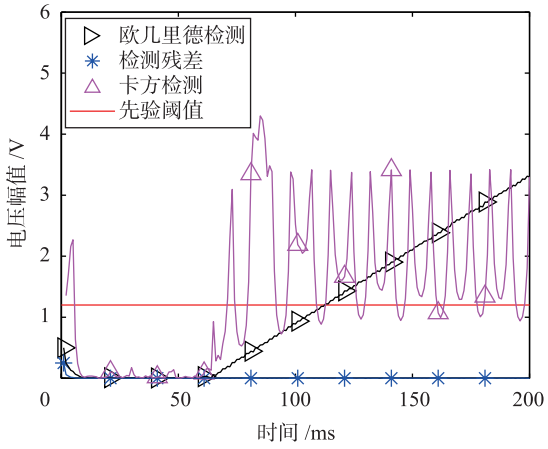

图 6 AKF 下的χ^2 检测和检测残差法和欧几里德检测

Fig.6 χ^2 detect, state residual and Euclidean detect under FDIA using AKF

5 结论

本文提出了一种基于自适应卡尔曼滤波的检测方法来检测隐蔽假数据注入攻击. 首先在三相电压模型的基础上分析了随机攻击和隐蔽假数据注入攻击的特性. 其次考虑用自适应卡尔曼滤波代替标准卡尔曼滤波来估计系统状态的同时降低误报率和缩短检测时间. 最后，比较常用的χ^2 检测，欧几里德检测与检测残差法对攻击检测的实际效果，仿真结果表明提出的检测残差法不仅可以有效地检测随机攻击，还可以快速检测隐蔽假数据注入攻击. 下一步工作考虑利用自适应阈值进一步缩短检测时间，同时从多个传感器中定位出被攻击的传感器.

参考文献

[1] Moslemi R, Mesbahi A, Velni J M. A fast, decentralized covariance selection-based approach to detect cyber attacks in smart grids[J]. IEEE Transactions on Smart Grid, 2017, PP(99): 1 – 12.

[2] 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.
[3] Metke A R, Ekl R L. Security technology for smart grid networks[J]. IEEE Transactions on Smart Grid, 2010, 1(1): 99 – 107.
[4] Pasqualetti F, Dörfler F, Bullo F. Attack detection and identification in cyber-physical systems[J]. IEEE Transactions on Automatic Control, 2013, 58(11): 2715 – 2729.
[5] Douligeris C, Mitrokotsa A D. DoS attacks and defense mechanisms: a classification[C]//IEEE International Symposium on Signal Processing and Information Technology. Berlin, Germany: Springer, 2003: 190 – 193.
[6] Mo Y, Sinopoli B. Secure control against replay attacks[C]//Allerton Conference on Communication, Control, and Computing. Pscataway, NJ, USA: IEEE, 2009: 911 – 918.
[7] 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, 2011, 14(1): 1 – 34.
[8] Li Z, Shahidehpour M, Alabdulwahab A, et al. Bilevel model for analyzing coordinated cyber-physical attacks on power systems[J]. IEEE Transactions on Smart Grid, 2016, 7(5): 2260 – 2272.
[9] 黄心，王清心，丁家满. 基于概率盒理论的电网规划方案中指标不确定性建模[J]. 信息与控制，2016，45(3)：272 – 277.
Huang X, Wang Q X, Ding J M. Index uncertainty modeling in grid planning based on probability box theory[J]. Information and Control, 2016, 45(3): 272 – 277.
[10] Sandberg H, Teixeira A, Johansson K H. On security indices for state estimators in power networks[C]//Preprints of the First Workshop on Secure Control Systems. Stockholm, Sweden: CPSWEEK, 2010: 1 – 6.
[11] Teixeira A, Dán G, Sandberg H, et al. A cyber security study of a SCADA energy management system: Stealthy deception attacks on the state estimator[J]. IFAC Proceedings Volumes, 2011, 44(1): 11271 – 11277.
[12] Deng R, Xiao G, Lu R, et al. False data injection on state estimation in power systems-attacks, impacts, and defense: A Survey[J]. IEEE Transactions on Industrial Informatics, 2017, 13(2): 411 – 423.
[13] Bobba R B, Rogers K M, Wang Q, et al. Detecting false data injection attacks on DC State estimation[C]//Preprints of the First Workshop on Secure Control Systems. Stockholm, Sweden: CPSWEEK, 2010: 1 – 9.
[14] Gu Y, Liu T, Wang D, et al. Bad data detection method for smart grids based on distributed state estimation[C]//2013 IEEE International Conference on Communications. Piscataway, NJ, USA: IEEE, 2013: 4483 – 4487.
[15] Manandhar K, Cao X, Hu F, et al. Detection of faults and attacks including false data injection attack in smart grid using Kalman filter[J]. IEEE Transactions on Control of Network Systems, 2014, 1(4): 370 – 379.
[16] Kwon C, Hwang I. Reachability analysis for safety assurance of cyber-physical systems against cyber attacks[J]. IEEE Transactions on Automatic Control, 2017, PP(99): 1 – 1.
[17] 夏小虎，刘明. 基于交互式多模型卡尔曼滤波的电池荷电状态估计[J]. 信息与控制，2017，46(5)：519 – 524.
Xia X H, Liu M. Battery state-of-charge estimation using interactive multiple-model Kalman filter[J]. Information and Control, 2017, 46(5): 519 – 524.
[18] Qi H, Wang X, Tolbert L M, et al. A resilient real-time system design for a secure and reconfigurable power grid[J]. IEEE Transactions on Smart Grid, 2011, 2(4): 770 – 781.
[19] Rigatos G, Serpanos D, Zervos N. Detection of attacks against power grid sensors using Kalman filter and statistical decision making[J]. IEEE Sensors Journal, 2017, PP(99): 1 – 1.
[20] Mo Y, Garone E, Casavola A, et al. False data injection attacks against state estimation in wireless sensor networks[C]//49th IEEE Conference on Decision and Control. Piscatawy, NJ, USA: IEEE, 2011: 5967 – 5972.
[21] Jie Z, Zhang G, Tao W, et al. Overview of fraudulent data attack on power system state estimation and defense mechanism[J]. Power System Technology, 2016, 40(8): 2406 – 2415.
[22] 赵琳，王小旭，孙明，等. 基于极大后验估计和指数加权的自适应 UKF 滤波算法[J]. 自动化学报，2010，36(7)：1007 – 1019.
Zhao L, Wang X X, Sun M, et al. An adaptive UKF filtering algorithm based on maximum a posteriori estimation and exponential weighting [J]. Acta Automatica Sinica, 2010, 36(7): 1007 – 1019.
[23] Sun J, Xu X, Liu Y, et al. FOG random drift signal denoising based on the improved AR model and modified sage-Husa adaptive Kalman filter [J]. Sensors, 2016, 16(7): 1073.
[24] Dixon W J, Massey F J J. An introduction to statistical analysis[M]. 4th ed. NewYork, USA: McGraw-Hill, 1983: 426 – 441.
[25] Kashima K, Inoue D. Replay attack detection in control systems with quantized signals[C]//2015 European Control Conference. Piscataway, NJ, USA: IEEE, 2015: 782 – 787.

作者简介

罗小元(1976 –)，男，博士生导师，教授. 研究领域为多机器人移动组网及拓扑优化，信息物理系统的攻击检测与定位.
朱鸣皋(1994 –)，男，硕士生. 研究领域为智能电网的攻击检测.
王新宇(1987 –)，男，博士生. 研究领域为智能电网的攻击检测与定位.