虚假数据注入攻击信号的融合估计

翁品迪 [1,2] 陈 博 [1,2] 俞 立 [1,2]

摘 要 研究了信息物理系统中假数据注入 (False data injection, FDI) 攻击信号的检测问题. 在分布式融合框架下, 首先将 FDI 攻击信号建模为信息物理系统模型中的未知输入, 从而使得攻击信号的检测问题转化为对 FDI 攻击信号的实时估计问题. 其次, 在每个传感器端设计基于自适应卡尔曼滤波的 FDI 攻击信号的局部估计器; 在融合中心端引入补偿因子, 设计分布式信息融合准则以导出攻击信号的融合估计器. 特别地, 当 FDI 攻击信号是时变情况时, 融合过程中补偿因子的引入可以大大提高对攻击信号的估计精度. 最后, 通过两个仿真算例验证所提算法的有效性.

关键词 自适应 Kalman 滤波, 假数据注入攻击, 攻击信号估计, 信息融合

引用格式 翁品迪, 陈博, 俞立. 虚假数据注入攻击信号的融合估计. 自动化学报, 2021, **47**(9): 2292−2300

DOI 10.16383/j.aas.c190045

Fusion Estimate of FDI Attack Signals

WENG Pin-Di [1,2] CHEN Bo [1,2] YU Li [1,2]

Abstract This paper is concerned with the estimation problem of false data injection (FDI) attacks in cyber-physical systems, where the false data is injected into the actuator. Under the distributed fusion framework, the attack signal is modeled as an unknown parameter in the addressed system models, and each local estimate of the attack signal at the sensor is obtained based on adaptive Kalman filtering method. Subsequently, by introducing the compensation factor, an optimal weighing fusion criterion is designed to improve estimation precision of attack signals. Particularly, when the FDI attack signal is time-varying, the introduction of compensation factor during fusing process can indeed improve the estimation precision largely. Finally, two illustrative examples are used to show the effectiveness of the proposed methods.

Key words Adaptive Kalman filter, false data injection (FDI) attacks, estimation of attack signal, information fusion

Citation Weng Pin-Di, Chen Bo, Yu Li. Fusion estimate of FDI attack signals. *Acta Automatica Sinica*, 2021, **47**(9): 2292−2300

信息物理系统 (Cyber-physical systems, CPSs) 是一个综合了计算、网络和物理环境的多维智能化复杂系统, 它借助有线或无线通信网络将各个关键设施整合在一起, 使得人机和物理进程的交互更加便捷. 随着网络通信、嵌入式系统、计算机控制及相关硬件技术的不断发展, 信息物理系统已引起了工业界的广泛关注[1−3]. 然而在此框架下, 原有系统的封闭性被打破, 使其面临着来自网络攻击的安全威胁. 例如: 2010 年, 伊朗的首座核电站−布什核电站被震网病毒攻击, 使得伊朗的第一座核设施推迟发电, 严重损害了伊朗的工业设施[4]; 2011 年美国伊利诺伊州一处水利控制系统遭到网络攻击, 险些造成大面积的供水中断. 当信息物理系统遭受网络攻击时, 准确、及时地监测到攻击信号对于监测中心采取高效的防御策略是至关重要的[5].

从控制角度来看, 信息物理系统是传统数字控制系统融合通信技术的下一代网络化控制系统[6]. 信息物理系统中被控对象的量测信号和控制信号均通过网络进行传输, 因此无论是传感器到控制器端, 还是控制器到执行器端, 均有受到恶意网络攻击的可能, 从而影响系统的稳定性, 甚至引发系统崩溃, 造成严重的生产事故与经济损失[7−8]. 典型的网络攻击有三种, 分别是拒绝服务攻击 (Denial of service, DoS)、欺骗攻击和重放攻击. 针对拒绝服务攻击, 文献 [9] 将 DoS 攻击建模为一类能量约束问题, 针对攻击者与防御者设计递归分布式卡尔曼估计器进行双边优化; 文献 [10] 假定 DoS 攻击有界的情况下, 通过构建嵌套切换模型得到了 CPS 在基于包控制方法下的稳定性条件. 针对重放攻击, 文献 [11] 从防御者的角度提出一种带补偿策略的数学模型来描述重放攻击和带宽约束, 并在线性最小方差意义下设计了递归分布式卡尔曼融合估计器. 欺骗攻击又称为假数据注入 (False data injection, FDI) 攻击, 它通过向系统注入错误的控制信号或测量信号影响信号数据的准确性. 注意到 FDI 攻击可以通过欺骗攻击检测机制来影响信息物理系统, 从而造成攻击检测器的漏报或虚警[12]. 文献 [13] 中研究了电力系统中针对 FDI 攻击的状态估计问题, 揭示了现有错误测量检测算法中存在的脆弱性. 该研究表明,即使攻击者的资源受限, 依旧可以改变状态估计的结果; 而在假定攻击者可以获取系统参数及所有数据流的基础上, 此类攻击则可以对系统产生一定影响的同时却不被检测到. 目前信息物理系统中的 FDI 攻击主要采用异常信号检测方法进行检测[14], 如基于二元假设的贝叶斯检测[15−18]、基于卡尔曼滤波器的 χ^2 检测[19−21] 和加权最小平方检测[22−24]. 后两种方法均是利用观测值构造新息残差, 然后通过与一个给定的阈值比较来判决是否受到攻击. 攻击检测的阈值对检测精度起着重要作用, 然而阈值通常在检测前根据已有经验进行选择, 这无疑会降低检测精度. 而且对于 CPSs 而言, 特别是电力、医疗系统, 仅仅检测或识别到攻击是不够的, 因为这些基础设施系统无法快速对其关闭、重启、恢复, 以对抗攻击. 为此, 不同于文献 [15−24] 中基于阈值的 FDI 攻击信号检测方法, 本文将对 CPSs 中控制器与执行器间的通信网络遭受的 FDI 攻击信号进行实时估计, 不仅可以检测攻击信号是否存在, 而且还可以掌握攻击信号的基本特征, 从而使防御方采取简单有效的补救措施以降低系统性能的损失程度.

为了对 FDI 攻击信号进行实时估计, 本文将攻击信号建模为状态动态方程中的一个未知参数, 然后在每个传感器端利用自适应卡尔曼滤波设计 FDI 攻击信号的局部估计. 由于基于递归最小二乘的自适应 Kalman 滤波方法适应于被估计的参数是时不变或者变化缓慢的, 而恶意的攻击信号往往是时变的. 因此, 本文充分利用多传感器融合所提供的冗余信息, 通过引入补偿因子和设计分布式融合准则, 来提高时变攻击信号的估计精度. 注意到补偿因子是通过影响局部互协方差的信息来提高估计性能, 而局部

收稿日期 2019-01-18 录用日期 2019-07-10
Manuscript received January 18, 2019; accepted July 10, 2019
国家自然科学基金 (61673351, 61973277) 资助
Supported by National Natural Science Foundation of China (61673351, 61973277)
本文责任编委 孙秋野
Recommended by Associate Editor SUN Qiu-Ye
1. 浙江工业大学网络空间安全研究院 杭州 310023 2. 浙江工业大学信息工程学院 杭州 310023
1. Institute of Cyberspace Security, Zhejiang University of Technology, Hangzhou 310023 2. College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023

互协方差矩阵存在于分布式融合结构中. 最后, 通过两个仿真验证所设计的融合算法对 FDI 攻击信号估计的有效性. 特别地, 当攻击信号是时变时, 仿真结果表明所引入的补偿因子可以明显提高攻击信号的融合估计性能.

符号说明. E 表示数学期望, $\mathrm{diag}\{N_1,\cdots,N_n\}$ 表示由 $N_1,\cdots,N_n$ 组成的块对角矩阵, I 表示单位矩阵, $\perp$ 表示正交.

1 问题描述与分析

考虑如下离散线性时变状态空间模型

$$\boldsymbol{x}(k)=A(k)\boldsymbol{x}(k-1)+B(k)\boldsymbol{u}(k-1)+\boldsymbol{w}(k) \tag{1}$$

其中, $\boldsymbol{x}(k)\in\mathbf{R}^n$ 为系统状态, $\boldsymbol{u}(k)\in\mathbf{R}^l$ 为控制输入; $A(k)$ 与 $B(k)$ 分别为系统状态转移矩阵和控制输入矩阵. 当控制输入信号 $\boldsymbol{u}(k)$ 受到 FDI 攻击时, 被攻击后的控制输入信号描述为

$$\boldsymbol{u}_a(k)=\boldsymbol{u}(k)+\boldsymbol{\theta}(k) \tag{2}$$

其中, $\boldsymbol{\theta}(k)\in\mathbf{R}^l$ 表示 FDI 攻击信号. 因此, 控制信号遭受攻击后的系统 (1) 建模为

$$\boldsymbol{x}(k)=A(k)\boldsymbol{x}(k-1)+B(k)(\boldsymbol{u}(k-1)+\boldsymbol{\theta}(k-1))+\boldsymbol{w}(k) \tag{3}$$

然后, 利用 L 个传感器对系统 (3) 的状态进行实时监测, 且每个传感器的量测 $\boldsymbol{y}_i(k)$ 建模为

$$\boldsymbol{y}_i(k)=C_i(k)\boldsymbol{x}(k)+\boldsymbol{v}_i(k),\quad i=1,2,\cdots,L \tag{4}$$

其中, $C_i(k)$ 为量测矩阵. $\boldsymbol{w}(k)\in\mathbf{R}^n$, $\boldsymbol{v}_i(k)\in\mathbf{R}^{m_i}$ 是相互独立的零均值高斯白噪声, 且满足

$$\mathrm{E}\left\{\begin{bmatrix}\boldsymbol{w}(k)\\ \boldsymbol{v}_i(k)\\ \boldsymbol{v}_j(k)\end{bmatrix}\begin{bmatrix}\boldsymbol{w}(k)\\ \boldsymbol{v}_i(k)\\ \boldsymbol{v}_j(k)\end{bmatrix}^{\mathrm{T}}\right\}=\mathrm{diag}\{Q(k),R_i(k),R_j(k)\},\quad i\neq j \tag{5}$$

本文要解决的问题是: 针对控制信号遭受 FDI 攻击篡改的信息物理系统 (3), 根据各个传感器的量测信息, 计算得到 FDI攻击信号的局部估计 $\hat{\boldsymbol{\theta}}_i(k)$; 然后, 基于局部估计 $\hat{\boldsymbol{\theta}}_i(k)$, $i=1,2,\cdots,L$, 设计满足 $\sum_{i=1}^{L}W_i^{\theta}(k)=I$ 最优加权融合矩阵 $W_1^{\theta}(k),\cdots,W_L^{\theta}(k)$, 使得攻击信号的分布式融合估计器

$$\hat{\boldsymbol{\theta}}_0(k)=\sum_{i=1}^{L}W_i^{\theta}(k)\hat{\boldsymbol{\theta}}_i(k) \tag{6}$$

在线性最小方差意义下是最优的.

注 1. 本文的主要思想是将攻击信号 $\boldsymbol{\theta}(k)$ 建模为系统 (3) 中的未知输入, 然后在信息融合框架下利用多组传感器的冗余量测信息, 实现对 $\boldsymbol{\theta}(k)$ 的实时估计. 特别地, 分布式信息融合不仅可以保证所设计估计器的鲁棒性, 而且也能够提高对攻击信号的估计精度. 注意到攻击信号的融合估计器与原闭环系统 (1) 中的控制信号设计是相互独立的, 即: 本文的控制信号是原有闭环控制系统根据未被攻击情况下的期望性能而提前设计给出的, 不依赖于融合中心的状态估计结果. 此外, 由于文中所给出的局部估计算法是以递推形式给出, 所以实现局部的状态估计不需要局部可观性条件. 但是, 为了保证局部估计器可以提供较好的估计性能, 则需要 $(A(k),C_i(k))$ 是可观的.

2 攻击信号的估计–分布式融合策略

在分布式融合结构下, 首先基于每个传感器量测信息 $\{\boldsymbol{y}_i(1),\cdots,\boldsymbol{y}_i(k)\}$, 结合 Kalman 滤波和带有遗忘因子的递归最小二乘估计方法导出攻击信号的局部估计 $\hat{\boldsymbol{\theta}}_i(k)$, 主要结果由定理 1 给出.

定理 1. 给定初始值 $P_i(0|0)=P_0$, $\Upsilon_i(0)=0$, $S_i(0)=\omega I_p$, $\hat{\boldsymbol{\theta}}_i(0)=\boldsymbol{\theta}_0$, $\hat{\boldsymbol{x}}_i(0|0)=\boldsymbol{x}_0$ 和遗忘因子 $0<\lambda_i\leq 1$, 则局部估计 $\hat{\boldsymbol{\theta}}_i(k)$ 由以下递推公式计算

$$\begin{cases}\widetilde{\boldsymbol{y}}_i(k)=\boldsymbol{y}_i(k)-C_i(k)[A(k)\hat{\boldsymbol{x}}_i(k-1|k-1)+\\ \qquad B(k)(\boldsymbol{u}(k-1)+\hat{\boldsymbol{\theta}}_i(k-1))]\\ \hat{\boldsymbol{\theta}}_i(k)=\hat{\boldsymbol{\theta}}_i(k-1)+\Gamma_i(k)\widetilde{\boldsymbol{y}}_i(k)\\ \hat{\boldsymbol{x}}_i(k|k)=A(k)\hat{\boldsymbol{x}}_i(k-1|k-1)+\\ \qquad B(k)(\boldsymbol{u}(k-1)+\hat{\boldsymbol{\theta}}_i(k-1))+\\ \qquad K_i(k)\widetilde{\boldsymbol{y}}_i(k)+\Upsilon_i(k)[\hat{\boldsymbol{\theta}}_i(k)-\\ \qquad \hat{\boldsymbol{\theta}}_i(k-1)]\end{cases} \tag{7}$$

$$\begin{cases}P_i(k|k-1)=A(k)P_i(k-1|k-1)A^{\mathrm{T}}(k)+Q(k)\\ \Sigma_i(k)=C_i(k)P_i(k|k-1)C_i^{\mathrm{T}}(k)+R_i(k)\\ K_i(k)=P_i(k|k-1)C_i^{\mathrm{T}}(k)\Sigma_i^{-1}(k)\\ P_i(k|k)=[I_n-K_i(k)C_i(k)]P_i(k|k-1)\end{cases} \tag{8}$$

$$\begin{cases}\Upsilon_i(k)=[I_n-K_i(k)C_i(k)]A(k)\Upsilon_i(k-1)+\\ \qquad [I_n-K_i(k)C_i(k)]B(k)\\ \Omega_i(k)=C_i(k)A(k)\Upsilon_i(k-1)+C_i(k)B(k)\\ \Lambda_i(k)=[\lambda_i\Sigma_i(k)+\Omega_i(k)S_i(k-1)\Omega_i^{\mathrm{T}}(k)]^{-1}\\ \Gamma_i(k)=S_i(k-1)\Omega_i^{\mathrm{T}}(k)\Lambda_i(k)\\ S_i(k)=\dfrac{1}{\lambda_i}S_i(k-1)-\dfrac{1}{\lambda_i}\Gamma_i(k)\Omega_i(k)S_i(k-1)\end{cases} \tag{9}$$

证明. FDI 攻击信号 $\boldsymbol{\theta}(k)$ 的局部估计器的推导类似于文献 [25] 与文献 [26] 的推导过程, 故在此省略. □

注 2. 局部 FDI 攻击信号估计 $\hat{\boldsymbol{\theta}}_i(k)$ 和局部状态估计 $\hat{\boldsymbol{x}}_i(k|k)$ 由递推式 (7) 计算, 其中, 未知输入 $\boldsymbol{\theta}(k-1)$ 的真实值由 $\hat{\boldsymbol{\theta}}_i(k-1)$ 给出. 当 $\boldsymbol{\theta}(k-1)=\hat{\boldsymbol{\theta}}_i(k-1)$ 时, 则

$$\widetilde{\boldsymbol{y}}_i(k)=C_i(k)A(k)\widetilde{\boldsymbol{x}}_i(k-1|k-1)+C_i(k)\boldsymbol{w}(k)+\boldsymbol{v}_i(k) \tag{10}$$

其中, $\widetilde{\boldsymbol{x}}_i(k-1|k-1)=\boldsymbol{x}(k-1)-\hat{\boldsymbol{x}}_i(k-1|k-1)$. 显然新息 (10) 与标准 Kalman 滤波的新息相同, 此时将 $B(k)\hat{\boldsymbol{\theta}}_i(k-1)$ 作为局部状态估计器的输入信号, 则可直接由 Kalman 滤波得到状态的估计值; 当 $\boldsymbol{\theta}(k-1)\neq\hat{\boldsymbol{\theta}}_i(k-1)$ 时, 将无法避免攻击信号的估计误差, 此时则用 $\Upsilon_i(k)[\hat{\boldsymbol{\theta}}_i(k)-\hat{\boldsymbol{\theta}}_i(k-1)]$ 补偿这一误差. 另一方面, 遗忘因子 λ_i $(0<\lambda_i\leq 1)$ 是用来控制过去量测信息被遗忘的速度, 以体现新近数据的作用, 即: 小的遗忘因子 λ_i 代表过去的量测信息将被快速遗忘. 根据递归最小二乘法的核心思想, 当 $\boldsymbol{\theta}(k)$ 是缓慢变化时, 式 (7) 中的局部估计器 $\hat{\boldsymbol{\theta}}_i(k)$ 的性能随着 λ_i 的减小而提高; 而当 $\boldsymbol{\theta}(k)$ 是时不变参数时, 则式 (7) 中的局部估计器 $\hat{\boldsymbol{\theta}}_i(k)$ 的性能随着 λ_i 的减小而降低[26]. 特别地, 当 $\boldsymbol{\theta}(k)$ 是变化速率快时, 不能通

过调节遗忘因子保证局部估计器的收敛性.

定义

$$\begin{cases}\widetilde{\boldsymbol{\theta}}_i(k)=\boldsymbol{\theta}(k)-\hat{\boldsymbol{\theta}}_i(k)\\ P_{ij}^{\theta}(k)=\mathrm{E}(\widetilde{\boldsymbol{\theta}}_i(k)\tilde{\boldsymbol{\theta}}_j^{\mathrm{T}}(k),\ \ \forall i,j\end{cases}\tag{11}$$

当在 k 时刻 L 个局部估计 $\hat{\boldsymbol{\theta}}_i(k)$ 由定理 1 给出时, 根据文献 [27] 的结论可以导出满足式 (6) 最优融合估计 $\hat{\boldsymbol{\theta}}_0(k)=\sum_{i=1}^{L}W_i^{\theta}(k)\hat{\boldsymbol{\theta}}_i(k)$ 的权重 $W_i^{\theta}(k)$, $i=1,2,\cdots,L$, 由下式计算得到

$$W^{\theta}(k)=\Sigma_{\theta}^{-1}(k)e_{\theta}\left(e_{\theta}^{\mathrm{T}}\Sigma_{\theta}^{-1}(k)e_{\theta}\right)^{-1}\tag{12}$$

其中, $W^{\theta}(k)=[W_1^{\theta}(k),W_2^{\theta}(k),\cdots,W_L^{\theta}(k)]^{\mathrm{T}}$ 与 $e_{\theta}=[I_p,\cdots,I_p]^{\mathrm{T}}$ 均为 $pL\times p$ 的矩阵; $\Sigma_{\theta}(k)=(P_{ij}^{\theta}(k))$ 为 $pL\times pL$ 的对称正定矩阵. 在此情形下, 最优分布式融合估计器 $\hat{\boldsymbol{\theta}}_0(k)$ 的协方差矩阵为

$$P_0^{\theta}(k)=(e_{\theta}^{\mathrm{T}}\Sigma_{\theta}^{-1}(k)e_{\theta})^{-1}\tag{13}$$

且满足 $\mathrm{tr}(P_{ii}^{\theta}(k))\geq\mathrm{tr}(P_0^{\theta}(k))$.

由式 (11) 可知, 最优加权矩阵 $W_i^{\theta}(k)$ $(i=1,2,\cdots,L)$ 的计算需要求解每个协方差矩阵 $P_{ij}^{\theta}(k)$ $(\forall i,j)$. 因此, 定理 2 将给出 $P_{ij}^{\theta}(k)$ 的递推形式. 在给出主要结果之前, 需要定义如下变量

$$\begin{cases}\widetilde{\boldsymbol{x}}_i(k|k)=\boldsymbol{x}(k)-\hat{\boldsymbol{x}}_i(k|k)\\ P_{ij}^{x}(k)=\mathrm{E}(\widetilde{\boldsymbol{x}}_i(k|k)\tilde{\boldsymbol{x}}_j^{\mathrm{T}}(k|k))\\ \Psi_{ij}(k)=\mathrm{E}(\widetilde{\boldsymbol{x}}_i(k|k)\tilde{\boldsymbol{\theta}}_j^{\mathrm{T}}(k))\end{cases}\tag{14}$$

定理 2. 给定补偿因子 $\eta>0$, FDI 攻击信号的协方差矩阵 $P_{ij}^{\theta}(k)$ 由以下的递推公式计算

$$\begin{aligned}P_{ij}^{\theta}(k)=&[I_p-\Gamma_i(k)C_i(k)B(k)]P_{ij}^{\theta}(k-1)\times\\&[I_p-\Gamma_j(k)C_j(k)B(k)]^{\mathrm{T}}+\\&\Gamma_i(k)C_i(k)A(k)P_{ij}^{x}(k-1)\times\\&A^{\mathrm{T}}(k)C_j^{\mathrm{T}}(k)\Gamma_j^{\mathrm{T}}(k)-[I_p-\Gamma_i(k)C_i(k)B(k)]\times\\&\Psi_{ji}^{\mathrm{T}}(k-1)A^{\mathrm{T}}(k)C_j^{\mathrm{T}}(k)\Gamma_j^{\mathrm{T}}(k)-\\&\Gamma_i(k)C_i(k)A(k)\Psi_{ij}(k-1)\times\\&[I_p-\Gamma_j(k)C_j(k)B(k)]^{\mathrm{T}}+\\&\Gamma_i(k)C_i(k)Q(k)C_j^{\mathrm{T}}(k)\Gamma_j^{\mathrm{T}}(k)+\\&\Gamma_i(k)\delta_{ij}R_i\Gamma_j^{\mathrm{T}}(k)+\eta I\end{aligned}\tag{15}$$

其中,

$$\begin{aligned}P_{ij}^{x}(k)=&[I_n-(K_i(k)+\Upsilon_i(k)\Gamma_i(k))C_i(k)]\times\\&[A(k)P_{ij}^{x}(k-1)A^{\mathrm{T}}(k)+B(k)P_{ij}^{\theta}(k-1)\times\\&B^{\mathrm{T}}(k)+A(k)\Psi_{ij}(k-1)B^{\mathrm{T}}(k)+\\&B(k)\Psi_{ji}^{\mathrm{T}}(k-1)A^{\mathrm{T}}(k)+\\&Q(k)][I_n-(K_j(k)+\Upsilon_j(k)\Gamma_j(k))C_j(k)]^{\mathrm{T}}+\\&[K_i(k)+\Upsilon_i(k)\Gamma_i(k)]\delta_{ij}R_i[K_j(k)+\\&\Upsilon_j(k)\Gamma_j(k)]^{\mathrm{T}}\end{aligned}\tag{16}$$

$$\begin{aligned}\Psi_{ij}(k)=&[I_n-(K_i(k)+\Upsilon_i(k)\Gamma_i(k))C_i(k)]\times\\&[A(k)\Psi_{ij}(k-1)+B(k)P_{ij}^{\theta}(k-1)]\times\\&[I_p-\Gamma_j(k)C_j(k)B(k)]^{\mathrm{T}}-\\&[I_n-(K_i(k)+\Upsilon_i(k)\Gamma_i(k))C_i(k)]\times\\&[A(k)P_{ij}^{x}(k-1)A^{\mathrm{T}}(k)+\\&B(k)\Psi_{ji}^{\mathrm{T}}(k-1)A^{\mathrm{T}}(k)+Q(k)]C_j^{\mathrm{T}}(k)\Gamma_j^{\mathrm{T}}(k)+\\&[K_i(k)+\Upsilon_i(k)\Gamma_i(k)]\delta_{ij}R_i\Gamma_j^{\mathrm{T}}(k)\end{aligned}\tag{17}$$

其中, $\delta_{ii}=1,\delta_{ij}=0,\ i\neq j$.

证明. 定义 $\boldsymbol{\phi}(k)=\boldsymbol{\theta}(k)-\boldsymbol{\theta}(k-1)$, 则由式 (7) 可导出攻击信号 $\boldsymbol{\theta}(k)$ 的局部估计误差为

$$\begin{aligned}\widetilde{\boldsymbol{\theta}}_i(k)=&\boldsymbol{\phi}(k)+\boldsymbol{\theta}(k-1)-[\hat{\boldsymbol{\theta}}_i(k-1)+\Gamma_i(k)\widetilde{\boldsymbol{y}}_i(k)]=\\&\widetilde{\boldsymbol{\theta}}_i(k-1)-\Gamma_i(k)C_i(k)\times\\&[A(k)\widetilde{\boldsymbol{x}}_i(k-1|k-1)+B(k)\widetilde{\boldsymbol{\theta}}_i(k-1)]-\\&\Gamma_i(k)C_i(k)\boldsymbol{w}(k)-\Gamma_i(k)\boldsymbol{v}_i(k)+\boldsymbol{\phi}(k)=\\&[I_p-\Gamma_i(k)C_i(k)B(k)]\widetilde{\boldsymbol{\theta}}_i(k-1)-\\&\Gamma_i(k)C_i(k)A(k)\widetilde{\boldsymbol{x}}_i(k-1|k-1)-\\&\Gamma_i(k)C_i(k)\boldsymbol{w}(k)-\Gamma_i(k)\boldsymbol{v}_i(k)+\boldsymbol{\phi}(k)\end{aligned}\tag{18}$$

由式 (3) 和式 (7) 可得状态估计误差为

$$\begin{aligned}\widetilde{\boldsymbol{x}}_i(k|k)=&[I_n-K_i(k)C_i(k)][A(k)\widetilde{\boldsymbol{x}}_i(k-1|k-1)+\\&B(k)\widetilde{\boldsymbol{\theta}}_i(k-1)]-\Upsilon_i(k)[\hat{\boldsymbol{\theta}}_i(k)-\hat{\boldsymbol{\theta}}_i(k-1)]+\\&[I_n-K_i(k)C_i(k)]\boldsymbol{w}(k)-K_i(k)\boldsymbol{v}_i(k)=\\&[I_n-K_i(k)C_i(k)][A(k)\widetilde{\boldsymbol{x}}_i(k-1|k-1)+\\&B(k)\widetilde{\boldsymbol{\theta}}_i(k-1)]+\Upsilon_i(k)[\widetilde{\boldsymbol{\theta}}_i(k)-\\&\widetilde{\boldsymbol{\theta}}_i(k-1)-\boldsymbol{\phi}(k)]+\\&[I_n-K_i(k)C_i(k)]\boldsymbol{w}(k)-K_i(k)\boldsymbol{v}_i(k)\end{aligned}\tag{19}$$

将式 (18) 代入式 (19), 导出

$$\begin{aligned}\widetilde{\boldsymbol{x}}_i(k|k)=&[I_n-(K_i(k)+\Upsilon_i(k)\Gamma_i(k))C_i(k)]\times\\&[A(k)\widetilde{\boldsymbol{x}}_i(k-1|k-1)+B(k)\widetilde{\boldsymbol{\theta}}_i(k-1)]+\\&[I_n-(K_i(k)+\Upsilon_i(k)\Gamma_i(k))C_i(k)]\boldsymbol{w}(k)-\\&[K_i(k)+\Upsilon_i(k)\Gamma_i(k)]\boldsymbol{v}_i(k)\end{aligned}\tag{20}$$

注意到 $\boldsymbol{\phi}(k)$ 是未知变量, 在此假设它与其他统计变量无关, 且它的协方差为 $\eta I=\mathrm{E}(\boldsymbol{\phi}(k)\boldsymbol{\phi}^{\mathrm{T}}(k))$. 在此情况下, 调整参数 η 可以改变所设计融合估计器对攻击信号 $\boldsymbol{\theta}(k)$ 的估计性能, 即通过 η 实现对估计误差 $\widetilde{\boldsymbol{\theta}}_i(k)$ 的协方差进行补偿, 因而 η 称为补偿因子. 根据 $\boldsymbol{w}(k)\perp\widetilde{\boldsymbol{x}}_j(k-1|k-1)$, $\boldsymbol{v}_i(k)\perp\widetilde{\boldsymbol{x}}_j(k-1|k-1)$, $\boldsymbol{\phi}_i(k)\perp\widetilde{\boldsymbol{x}}_j(k-1|k-1)$, $\boldsymbol{w}(k)\perp\widetilde{\boldsymbol{\theta}}_i(k-1)$, $\boldsymbol{v}_i(k)\perp\widetilde{\boldsymbol{\theta}}_i(k-1)$, $\boldsymbol{\phi}_i(k)\perp\widetilde{\boldsymbol{\theta}}_i(k-1)$, 由式 (18) 可得 $P_{ij}^{\theta}(k)$ 由下式计算

$$\begin{aligned}P_{ij}^{\theta}(k)=&[I_p-\Gamma_i(k)C_i(k)B(k)]P_{ij}^{\theta}(k-1)\times\\&[I_p-\Gamma_j(k)C_j(k)B(k)]^{\mathrm{T}}+\\&\Gamma_i(k)C_i(k)A(k)P_{ij}^{x}(k-1)A^{\mathrm{T}}(k)C_j^{\mathrm{T}}(k)\Gamma_j^{\mathrm{T}}(k)+\\&\Gamma_i(k)C_i(k)\mathrm{E}(\boldsymbol{w}(k)\boldsymbol{w}^{\mathrm{T}}(k))C_j^{\mathrm{T}}(k)\Gamma_j^{\mathrm{T}}(k)+\\&\Gamma_i(k)\mathrm{E}(\boldsymbol{v}_i(k)\boldsymbol{v}_j^{\mathrm{T}}(k))\Gamma_j^{\mathrm{T}}(k)+\mathrm{E}(\boldsymbol{\phi}(k)\boldsymbol{\phi}^{\mathrm{T}}(k))-\\&[I_p-\Gamma_i(k)C_i(k)B(k)]\Psi_{ji}^{\mathrm{T}}(k-1)\times\\&A^{\mathrm{T}}(k)C_j^{\mathrm{T}}(k)\Gamma_j^{\mathrm{T}}(k)-\Gamma_i(k)C_i(k)A(k)\times\\&\Psi_{ij}(k-1)[I_p-\Gamma_j(k)C_j(k)B(k)]^{\mathrm{T}}+\\&\Gamma_i(k)C_i(k)\mathrm{E}(\boldsymbol{w}(k)\boldsymbol{v}_j^{\mathrm{T}}(k))\Gamma_j^{\mathrm{T}}(k)+\\&\Gamma_i(k)\mathrm{E}(\boldsymbol{v}_i(k)\boldsymbol{w}^{\mathrm{T}}(k))C_j^{\mathrm{T}}(k)\Gamma_j^{\mathrm{T}}(k)\end{aligned}\tag{21}$$

根据式 (5), 式 (21) 简化为

$$\begin{aligned}P_{ij}^{\theta}(k)=&[I_p-\Gamma_i(k)C_i(k)B(k)]P_{ij}^{\theta}(k-1)\times\\&[I_p-\Gamma_j(k)C_j(k)B(k)]^{\mathrm{T}}+\Gamma_i(k)C_i(k)A(k)\times\\&P_{ij}^{x}(k-1)A^{\mathrm{T}}(k)C_j^{\mathrm{T}}(k)\Gamma_j^{\mathrm{T}}(k)-\\&[I_p-\Gamma_i(k)C_i(k)B(k)]\Psi_{ji}^{\mathrm{T}}(k-1)\times\\&A^{\mathrm{T}}(k)C_j^{\mathrm{T}}(k)\Gamma_j^{\mathrm{T}}(k)-\Gamma_i(k)C_i(k)A(k)\times\\&\Psi_{ij}(k-1)[I_p-\Gamma_j(k)C_j(k)B(k)]^{\mathrm{T}}+\\&\Gamma_i(k)C_i(k)Q(k)C_j^{\mathrm{T}}(k)\Gamma_j^{\mathrm{T}}(k)+\\&\Gamma_i(k)\mathrm{E}(\boldsymbol{v}_i(k)\boldsymbol{v}_j^{\mathrm{T}}(k))\Gamma_j^{\mathrm{T}}(k)+\eta I\end{aligned}\tag{22}$$

从而导出式 (15) 成立. 另一方面, 由式 (20) 可得 $P_{ij}^{x}(k)$ 由下式计算

$$\begin{aligned}P_{ij}^{x}(k)=&[I_n-(K_i(k)+\Upsilon_i(k)\Gamma_i(k))C_i(k)]\times\\&[A(k)P_{ij}^{x}(k-1)A^{\mathrm{T}}(k)+\\&B(k)P_{ij}^{\theta}(k-1)B^{\mathrm{T}}(k)+\\&A(k)\Psi_{ij}(k-1)B^{\mathrm{T}}(k)+\\&B(k)\Psi_{ji}^{\mathrm{T}}(k-1)A^{\mathrm{T}}(k)+\\&\mathrm{E}(\boldsymbol{w}(k)\boldsymbol{w}^{\mathrm{T}}(k))]\times\\&[I_n-(K_j(k)+\Upsilon_j(k)\Gamma_j(k))C_j(k)]^{\mathrm{T}}+\\&[K_i(k)+\Upsilon_i(k)\Gamma_i(k)]\mathrm{E}(\boldsymbol{v}_i(k)\boldsymbol{v}_j^{\mathrm{T}}(k))\times\\&[K_i(k)+\Upsilon_i(k)\Gamma_i(k)]^{\mathrm{T}}+\\&[I_n-(K_i(k)+\Upsilon_i(k)\Gamma_i(k))C_i(k)]\times\\&\mathrm{E}(\boldsymbol{w}(k)\boldsymbol{v}_j^{\mathrm{T}}(k))[K_j(k)+\Upsilon_j(k)\Gamma_j(k)]^{\mathrm{T}}+\\&[K_i(k)+\Upsilon_i(k)\Gamma_i(k)]\mathrm{E}(\boldsymbol{v}_i(k)\boldsymbol{w}^{\mathrm{T}}(k))\times\\&[I_n-(K_j(k)+\Upsilon_j(k)\Gamma_j(k))C_j(k)]^{\mathrm{T}}\end{aligned}\tag{23}$$

根据式 (5), 式 (23) 转换为

$$\begin{aligned}P_{ij}^{x}(k)=&[I_n-(K_i(k)+\Upsilon_i(k)\Gamma_i(k))C_i(k)]\times\\&[A(k)P_{ij}^{x}(k-1)A^{\mathrm{T}}(k)+\\&B(k)P_{ij}^{\theta}(k-1)B^{\mathrm{T}}(k)+\\&A(k)\Psi_{ij}(k-1)B^{\mathrm{T}}(k)+\\&B(k)\Psi_{ji}^{\mathrm{T}}(k-1)A^{\mathrm{T}}(k)+Q(k)]\times\\&[I_n-(K_j(k)+\Upsilon_j(k)\Gamma_j(k))C_j(k)]^{\mathrm{T}}+\\&[K_i(k)+\Upsilon_i(k)\Gamma_i(k)]\mathrm{E}(\boldsymbol{v}_i(k)\boldsymbol{v}_j^{\mathrm{T}}(k))\times\\&[K_j(k)+\Upsilon_j(k)\Gamma_j(k)]^{\mathrm{T}}\end{aligned}\tag{24}$$

故式 (16) 成立. 进一步, 根据式 (18) 和式 (20) 导出 $\Psi_{ij}(k)$ 由下式计算

$$\begin{aligned}\Psi_{ij}(k)=&[I_n-(K_i(k)+\Upsilon_i(k)\Gamma_i(k))C_i(k)]\times\\&[A(k)\Psi_{ij}(k-1)+B(k)P_{ij}^{\theta}(k-1)]\times\\&[I_p-\Gamma_j(k)C_j(k)B(k)]^{\mathrm{T}}-\\&[I_n-(K_i(k)+\Upsilon_i(k)\Gamma_i(k))C_i(k)]\times\\&[A(k)P_{ij}^{x}(k-1)A^{\mathrm{T}}(k)+\\&B(k)\Psi_{ji}^{\mathrm{T}}(k-1)A^{\mathrm{T}}(k)]\times\\&C_j^{\mathrm{T}}(k)\Gamma_j^{\mathrm{T}}(k)+[I_n-(K_i(k)+\\&\Upsilon_i(k)\Gamma_i(k))C_i(k)]\mathrm{E}(\boldsymbol{w}(k)\times\\&\boldsymbol{w}^{\mathrm{T}}(k))C_j^{\mathrm{T}}(k)\Gamma_j^{\mathrm{T}}(k)+[I_n-(K_i(k)+\Upsilon_i(k)\times\\&\Gamma_i(k))C_i(k)]\mathrm{E}(\boldsymbol{w}(k)\boldsymbol{v}_j^{\mathrm{T}}(k))\Gamma_j^{\mathrm{T}}(k)+[K_i(k)+\\&\Upsilon_i(k)\Gamma_i(k)]\mathrm{E}(\boldsymbol{v}_i(k)\boldsymbol{w}^{\mathrm{T}}(k))C_j^{\mathrm{T}}(k)\Gamma_j^{\mathrm{T}}(k)+\\&[K_i(k)+\Upsilon_i(k)\Gamma_i(k)]\mathrm{E}(\boldsymbol{v}_i(k)\boldsymbol{v}_j^{\mathrm{T}}(k))\Gamma_j^{\mathrm{T}}(k)\end{aligned}\tag{25}$$

由式 (5) 和式 (25) 导出

$$\begin{aligned}\Psi_{ij}(k)=&[I_n-(K_i(k)+\Upsilon_i(k)\Gamma_i(k))C_i(k)]\times\\&[A(k)\Psi_{ij}(k-1)+B(k)P_{ij}^{\theta}(k-1)]\times\\&[I_p-\Gamma_j(k)C_j(k)B(k)]^{\mathrm{T}}-\\&[I_n-(K_i(k)+\Upsilon_i(k)\Gamma_i(k))C_i(k)][A(k)\times\\&P_{ij}^{x}(k-1)A^{\mathrm{T}}(k)+B(k)\Psi_{ji}^{\mathrm{T}}(k-1)A^{\mathrm{T}}(k)]\times\\&C_j^{\mathrm{T}}(k)\Gamma_j^{\mathrm{T}}(k)+[I_n-(K_i(k)+\Upsilon_i(k)\Gamma_i(k))\times\\&C_i(k)]\mathrm{E}(\boldsymbol{w}(k)\boldsymbol{w}^{\mathrm{T}}(k))C_j^{\mathrm{T}}(k)\Gamma_j^{\mathrm{T}}(k)+\\&[K_i(k)+\Upsilon_i(k)\Gamma_i(k)]\mathrm{E}(\boldsymbol{v}_i(k)\boldsymbol{v}_j^{\mathrm{T}}(k))\Gamma_j^{\mathrm{T}}(k)\end{aligned}\tag{26}$$

故式 (17) 成立. □

注 3. 当补偿因子 $\eta=0$ 时, 意味着 $\widetilde{\boldsymbol{\theta}}_i(k)$ 的协方差信息无法得到补偿. 在此情况下, 如果攻击信号 $\boldsymbol{\theta}(k)$ 恒等于一个常值, 则所设计的融合估计器将提供满意的估计性能; 但是如果攻击信号 $\boldsymbol{\theta}(k)$ 是时变的, 则 η 取 0 时相应的融合估计性能将会因为无法得到协方差的补偿信息而变差. 由于 $\boldsymbol{\theta}(k)$ 是未知的, 所以补偿因子 η 作为一个可调参数可以

提高对攻击信号的估计精度. 但是如何设计补偿因子的选取准则以保证满意的融合性能非常具有挑战性, 将作为今后在攻击信号实时估计方面的重要研究方向之一.

注 4. 当获得 FDI 攻击信号的估计值时, 攻击对系统性能的破坏仍然存在. 为此, 可以采取在系统 (1) 中直接减去攻击信号融合估计值的方法来尽可能地降低攻击所带来的损失, 即带有补偿策略的系统 (1) 由如下动态方程描述

$$\begin{aligned}\boldsymbol{x}(k) = {} & A(k)\boldsymbol{x}(k-1) + B(k)\boldsymbol{u}_a(k-1) + \\ & \boldsymbol{w}(k) - B(k)\hat{\theta}_0(k-1) = \\ & A(k)\boldsymbol{x}(k-1) + B(k)\boldsymbol{u}(k-1) + \\ & \boldsymbol{w}(k) + B(k)(\theta(k-1) - \hat{\theta}_0(k-1))\end{aligned} \tag{27}$$

其中, $\hat{\theta}_0(k-1)$ 是攻击信号的融合估计值. 由式 (27) 可知, 融合估计器精度越高, 系统所受到的破坏就越小.

根据定理 1 和定理 2, 给定遗忘因子 $\lambda_1, \cdots, \lambda_L$ 和补偿因子 η, 则攻击信号的融合估计 $\hat{\boldsymbol{\theta}}_0(k)$ 实现的算法如下:

算法 1. 最优融合估计算法

步骤 1. 根据式 (8), 计算卡尔曼增益 $K_i(k)$, 然后根据式 (9), 计算参数估计增益 $\Gamma_i(k)$;

步骤 2. 将步骤 1 计算的结果代入式 (7), 得到攻击信号的局部估计 $\hat{\boldsymbol{\theta}}_i(k)$ 和局部状态估计 $\hat{\boldsymbol{x}}_i(k|k)$;

步骤 3. 根据式 (15) 计算协方差矩阵 $P_{ii}^{\theta}(k)$ 和互协方差矩阵 $P_{ij}^{\theta}(k)$;

步骤 4. 将步骤 3 计算的结果代入式 (12) 得到融合权重 $W_i^{\theta}(k)$;

步骤 5. 将步骤 2 和步骤 4 的结果代入式 (6), 计算得到融合估计 $\hat{\boldsymbol{\theta}}_0(k)$;

步骤 6. 重复步骤 1~5, 得到 $k+1$ 时刻的最优融合估计 $\hat{\boldsymbol{\theta}}_0(k+1)$.

3 仿真算例

算例 1. 考虑一个 IEEE 四路配电线路的电网系统. 本文采用互连的分布式发电器 (Distributed energy generator, DEG) 模型, 4 个 DEG 被建模为电压源. 4 个 DEG 通过对应的公共耦合点 (Point of common coupling, PCC) 与主电网连接, PCC 的电压为 $\boldsymbol{v}_s = (v_1, v_2, v_3, v_4)^{\mathrm{T}}$, 其中 v_i 为第 i 个 PCC 电压. 为了保持 DEG 的正常工作, 这些 PCC 电压需要保持在一定的参考值. 在每个 DEG 和其余的电力网络之间存在耦合电感. 然后, 节点电压方程可以用下述线性状态空间模型表示为[3]

$$\dot{\boldsymbol{x}}(t) = \begin{bmatrix} 175.9 & 176.8 & 511.0 & 1\,036.0 \\ -350.0 & 0 & 0 & 0 \\ -544.2 & -474.8 & -408.8 & -828.8 \\ -119.7 & -554.6 & -968.8 & -1\,077.5 \end{bmatrix} \boldsymbol{x}(t) \tag{28}$$

其中, $\boldsymbol{x}(t) = \boldsymbol{v}_s(t) - \boldsymbol{v}_{\mathrm{ref}}(t)$ 为 PCC 状态电压偏差, $\boldsymbol{v}_{\mathrm{ref}}(t)$ 为 PCC 参考电压; 针对系统 (28), 当采样周期取 $T_0 = 0.001$ s 时, 则离散化系统的状态转移矩阵为

$$A_d = \begin{bmatrix} 1.0779 & -0.0691 & 0.1412 & 0.6022 \\ -0.3665 & 1.0012 & -0.0417 & -0.1247 \\ -0.4245 & -0.2761 & 0.8447 & -0.6105 \\ 0.1515 & -0.2261 & -0.5207 & 0.5588 \end{bmatrix}$$

考虑到系统中存在的噪声, 得到如下离散系统方程:

$$\boldsymbol{x}(k) = A_d\boldsymbol{x}(k-1) + \boldsymbol{w}(k) \tag{29}$$

注意到在没有控制信号输入的情况下, 系统 (29) 无法稳定. 因此, 将设计控制器以确保系统稳定. 考虑到控制输入信号受到 FDI 攻击, 利用两组传感器对系统进行量测, 得到如下动态方程

$$\begin{aligned}\boldsymbol{x}(k) = {} & A_d\boldsymbol{x}(k-1) + B_d(\boldsymbol{u}(k-1) + \\ & \boldsymbol{\theta}(k-1)) + \boldsymbol{w}(k)\end{aligned} \tag{30}$$

$$\boldsymbol{y}_i(k) = C_i\boldsymbol{x}(k) + \boldsymbol{v}_i(k), \quad i = 1, 2 \tag{31}$$

其中, $B_d = [2.6890, -4.3035, -8.3410, -7.0725]^{\mathrm{T}}$, $\boldsymbol{u}(k) = -[5.7495, 3.4187, -7.3869, 8.3114]\boldsymbol{x}(k)$ 是通过极点配置设计的稳定控制器. 系统过程噪声 $\boldsymbol{w}(k)$ 的方差取 $Q = \mathrm{diag}\{0.1, 0.2, 0.2, 0.1\}$. 量测矩阵为

$$C_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

$$C_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

量测噪声的协方差矩阵分别取 $R_1 = \mathrm{diag}\{0.5, 0.6, 0.3, 0.2\}$, $R_2 = \mathrm{diag}\{0.8, 0.3, 0.5, 0.9\}$. 在仿真中考虑以下两种情况:

情况 1. 攻击信号保持常值不变, 取

$$\boldsymbol{\theta}(k) \equiv 1 \tag{32}$$

令补偿因子 $\eta = 0$ 及遗忘因子 $\lambda = 0.95$. 执行算法 1, 仿真结果如图 1 和图 2 所示.

由图 1 可知, 当遭受的攻击信号保持常值不变时, 所提出的融合算法可以很好地估计到攻击信号. 图 2 给出了 300 次 Monte Carlo 实验的局部估计器和融合估计器的均方误差 (Mean square errors, MSEs) 曲线. 由此图可以看到融合估计器的性能优于局部估计器的性能, 从而验证了基于多传感器信息融合的攻击信号估计方法优于基于单一传感器的估计方法.

情况 2. 攻击信号 $\boldsymbol{\theta}(k)$ 是时变的, 这里假设 $\boldsymbol{\theta}(k)$ 按如下正弦函数变化

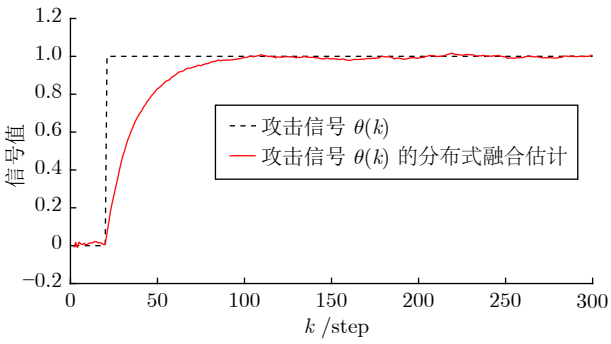

图 1 算例 1: 情况 1 中攻击信号和融合估计的轨迹

Fig.1 Example 1: The trajectories of attack signal and its fusion estimation under Case 1

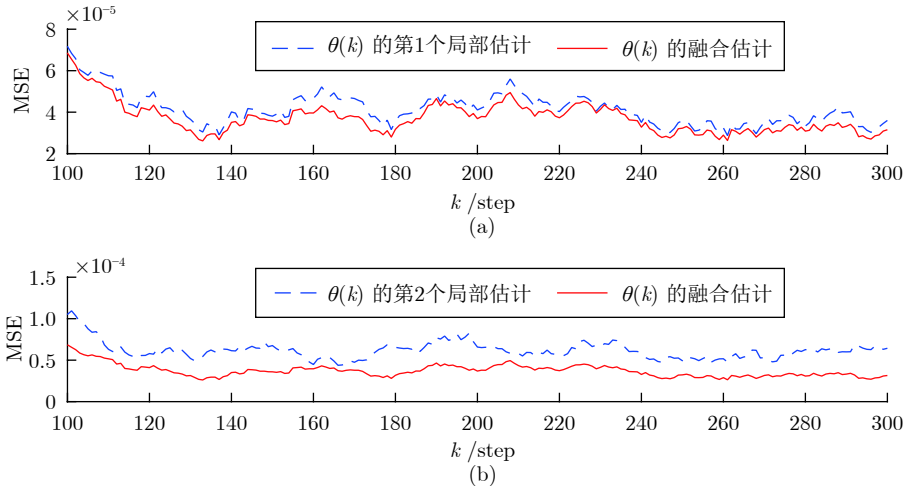

图 2　算例 1: 情况 1 中攻击信号的局部估计器与融合估计器之间的性能比较

Fig.2　Example 1: The performance comparison between local estimators and fusion estimators under Case 1

$$\boldsymbol{\theta}(k) = \sin(0.3k) \tag{33}$$

令补偿因子 $\eta = 0.2$ 和遗忘因子 $\lambda = 0.7$. 执行算法 1, 仿真结果如图 3~6 所示.

由图 3 可知, 当攻击信号是时变的情况时, 所设计的分布式融合器仍可以准确地估计攻击信号. 图 4 给出了 100 次 Monte Carlo 实验的局部估计器和融合估计器的 MSE 曲线. 由此图可知, 融合估计性能明显优于局部估计性能, 进一步说明了融合策略可以提高攻击信号的估计精度. 另一方面, 当取不同的补偿因子 η 时, 相应的融合估计性能如图 5 和图 6 所示, 即 $\eta = 0.2$ 和 $\eta = 0.5$ 的融合性能明显优于 $\eta = 0$ 时的融合估计器, 说明了当 FDI 攻击信号是时变的情形时, 补偿因子 η 对融合估计性能发挥着重要的作用, 即: 不考虑补偿策略的融合方法, 相应的估计性能会变得很差.

算例 2. 考虑一个由两个传感器监控的机动目标, 其运

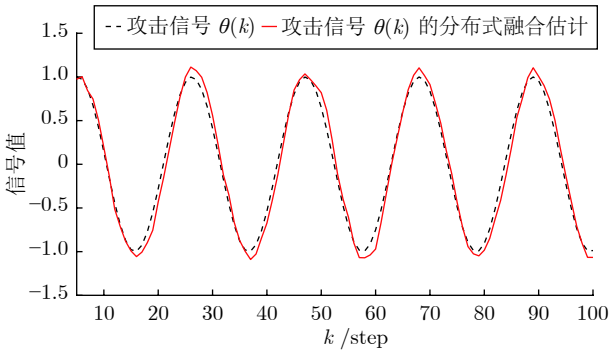

图 3　算例 1: 情况 2 中攻击信号和融合估计轨迹

Fig.3　Example 1: The trajectories of attack signal and its fusion estimation under Case 2

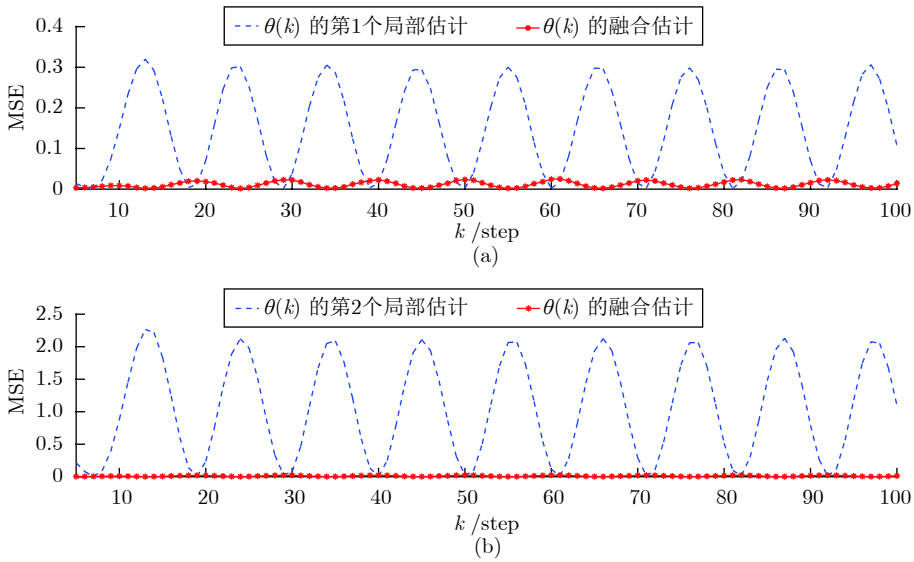

图 4　算例 1: 情况 2 中攻击信号的局部估计器与融合估计器之间的性能比较

Fig.4　Example 1: The performance comparison between local estimators and fusion estimators under Case 2

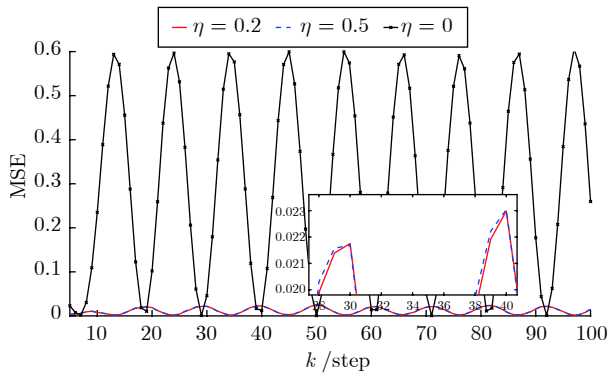

图 5 算例 1: 情况 2 中不同补偿因子下攻击信号融合估计性能的比较

Fig.5 Example 1: The comparison of fusion estimation performance of attack signal under different compensation factors under Case 2

动可以由它的位置和速度矢量来描述. 定义状态向量 $\boldsymbol{x}(k) = \mathrm{col}\{X_s(k), \dot{X}_s(k), Y_s(k), \dot{Y}_s(k)\}$, 其中, $(X_s(k), Y_s(k))$ 分别为 X 与 Y 轴的位置坐标, 同时 $(\dot{X}_s(k), \dot{Y}_s(k))$ 是对应的速度. 引入控制信号改变机动目标的速度以控制移动目标的运动, 考虑控制信号遭受 FDI 攻击, 机动目标的运动轨迹可由如下时变状态空间模型描述[28]:

$$\boldsymbol{x}(k) = \begin{bmatrix} 1 & f(k) & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & f(k) \\ 0 & 0 & 0 & 1 \end{bmatrix} \boldsymbol{x}(k-1) + \boldsymbol{w}(k) + \begin{bmatrix} 0 \\ 10f(k) \\ 0 \\ 10f(k) \end{bmatrix} (\boldsymbol{u}(k-1) + \boldsymbol{\theta}(k-1)) \tag{34}$$

$$\boldsymbol{y}_i(k) = C_i\boldsymbol{x}(k) + \boldsymbol{v}_i(k), \quad i = 1, 2 \tag{35}$$

其中, $f(k)$ 表示采样周期取 $f(k) = 0.9 + 0.1\sin(k)$; 控制信号 $\boldsymbol{u}(k)$ 取 $\boldsymbol{u}(k) = -[0, 0.01, 0, 0.01]x(k-1)$, 表示改变机动目标的速度; $\boldsymbol{\theta}(k)$ 表示未知的 FDI 攻击信号, 且取

$$\boldsymbol{\theta}(k) = \boldsymbol{\theta}(k-1) + c(k) \tag{36}$$

其中, $c(k)$ 为方差为 1 的零均值高斯随机信号. 系统过程噪声 $\boldsymbol{w}(k)$ 和量测噪声 $\boldsymbol{v}_1(k)$, $\boldsymbol{v}_2(k)$ 的方差分别取 $Q = \mathrm{diag}\{0.1,\ 0.2,\ 0.3,\ 0.2\}$, $R_1 = \mathrm{diag}\{0.3,\ 0.2\}$, $R_2 = \mathrm{diag}\{0.2,\ 0.5\}$. 量测矩阵取[28]

$$C_1 = \begin{bmatrix} 0.5 & 1 & 0 & 0 \\ 0 & 1 & 0.9 & 0.6 \end{bmatrix}$$

$$C_2 = \begin{bmatrix} 0.9 & 0.8 & 0 & 0 \\ 0 & 1 & 0.5 & 0.1 \end{bmatrix}$$

最后, 令遗忘因子 $\lambda = 0.6$, 补偿因子 $\eta_1 = 1, \eta_2 = 0$. 执行算法 1, 仿真结果如图 7~9 所示.

由图 7 可知, 系统参数时变的情况时, 所设计的估计算法仍可以很好地估计攻击信号. 图 8 基于 Monte Carlo 实验给出了不同补偿因子下融合估计器的 MSE 曲线. 图 7 和图 8 表明, 补偿因子 $\eta = 1$ 的情况下, 算法的性能优于 $\eta = 0$

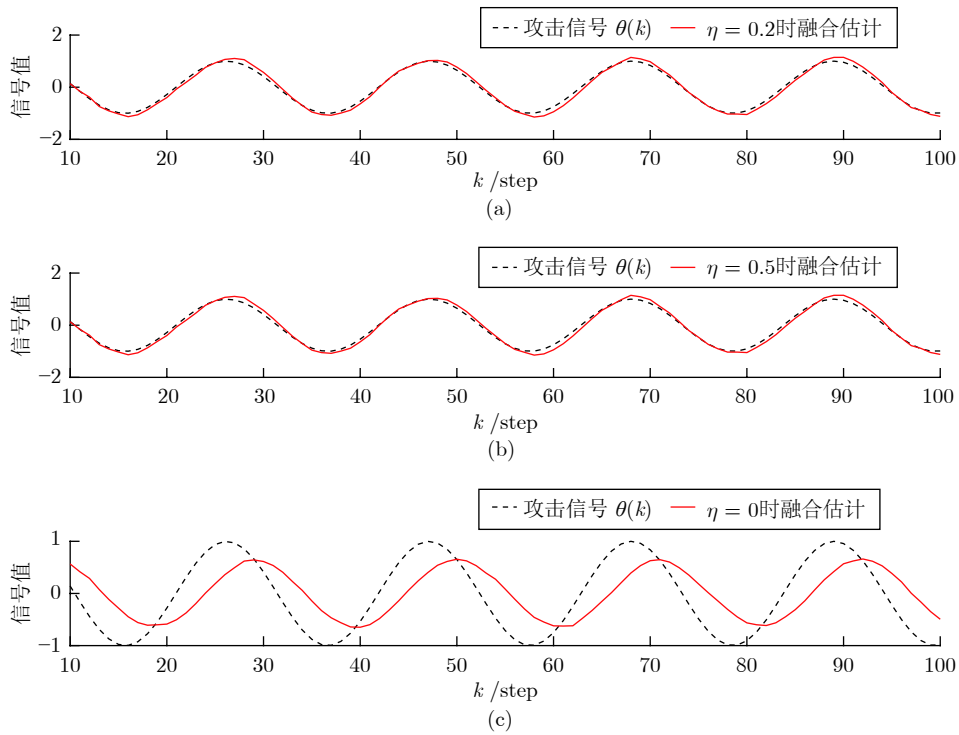

图 6 算例 1: 情况 2 中不同补偿因子下攻击信号融合估计的轨迹

Fig.6 Example 1: The trajectories of fusion estimation of attack signal under different compensation factors under Case 2

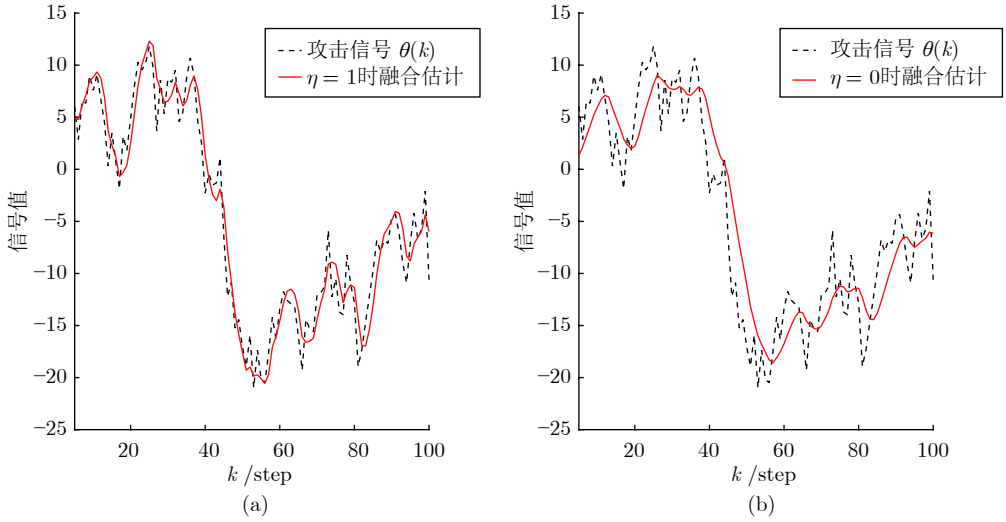

图 7　算例 2: 不同补偿因子下攻击信号融合估计的轨迹

Fig.7　Example 2: The trajectories of fusion estimation of attack signal under different compensation factors

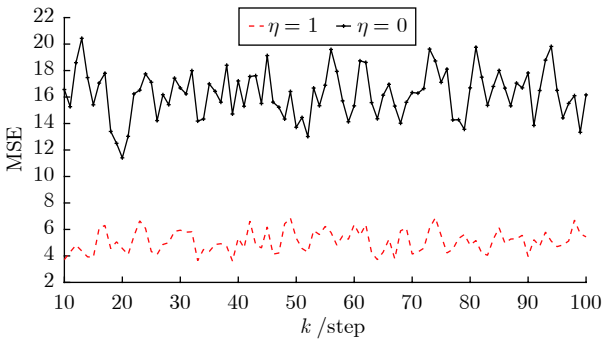

图 8　算例 2: 不同补偿因子下攻击信号融合估计性能的比较

Fig.8　Example 2: The comparison of fusion estimation performance of attack signal under different compensation factors

的情况 (无补偿的情况). 图 9 基于 Monte Carlo 实验给出了局部估计器和融合估计器的 MSE 曲线. 图 9 表明融合性能明显优于局部估计器. 仿真结果验证了算法的有效性.

4　结束语

本文研究了信息物理系统中遭受的 FDI 攻击信号的分布式融合估计问题. 首先, 根据 FDI 攻击策略对系统的影响, 将攻击信号建模为系统状态方程中的一个未知输入, 然后基于自适应卡尔曼滤波, 利用量测信息对系统状态和攻击参数进行联合估计, 从而得到攻击参数的局部变化轨迹. 在此基础上, 在分布式融合结构下, 引入补偿因子并设计了最优融合估计方法, 提高了攻击信号估计的精度. 最后, 通过例子验证了所提算法的有效性. 今后的研究方向包括如下几点: 首先, 选择合适的补偿因子对本文算法估计精度的提高有着显著的影响, 如何选择补偿因子将是未

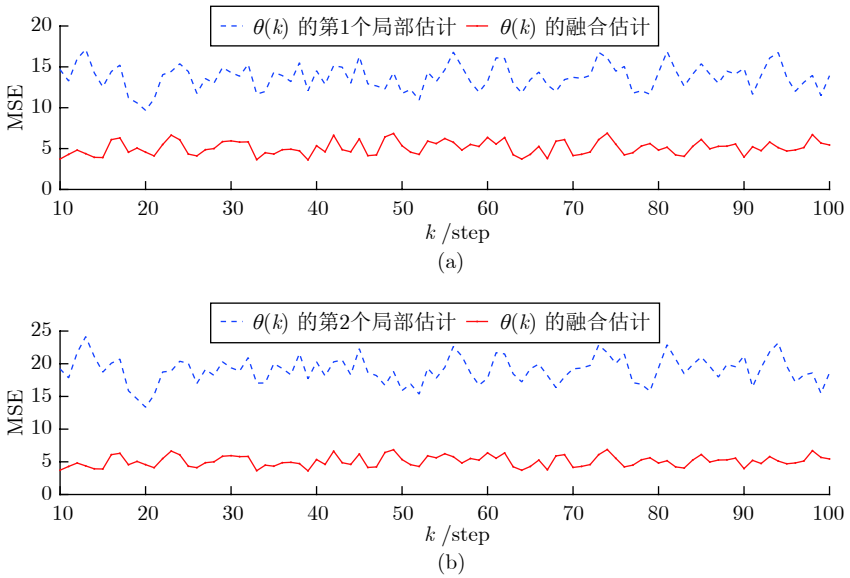

图 9　算例 2: 攻击信号的局部估计器与融合估计器之间的性能比较

Fig.9　Example 2: The performance comparison of attack signal between local estimators and fusion estimators

来研究的重要方向之一; 此外, 许多实际系统无法简单地利用线性模型描述, 因此将该算法推广至非线性情况的攻击信号估计具有重要的意义; 最后, 研究控制器与攻击信号融合估计器及攻击信号实时补偿策略的协同设计, 以降低攻击所带来的损失是另一个重要的研究方向.

References

1 Johansson K H, Pappas G J, Tabuada P, Tomlin C J. Guest editorial special issue on control of cyber-physical systems. *IEEE Transactions on Automatic Control*, 2014, **59**(12): 3120–3121

2 Fink J, Ribeiro A, Kumar V. Robust control for mobility and wireless communication in cyber-physical systems with application to robot teams. *Proceedings on the IEEE*, 2012, **100**(1): 164–178

3 Li H S, Lai L F, Poor H V. Multicast routing for decentralized control of cyber physical systems with an application in smart grid. *IEEE Journal on Selected Areas in Communications*, 2012, **30**(6): 1097–1107

4 Langner R. Stuxnet: Dissecting a cyberwarfare weapon. *IEEE Security and Privacy*, 2011, **9**(3): 49–51

5 Hu L, Wang Z D, Han Q L , Liu X H. State estimation under false data injection attacks: Security analysis and system protection. *Automatica*, 2018, **87**: 176–183

6 Einar Andel. Cyber-physical systems and industry 4.0. *Intelligent Manufacturing*, 2015, **9**: 10–12
(艾纳·安德尔. 信息物理系统和工业4.0. 智能制造, 2015, **9**: 10–12)

7 Pasqualetti F, Dorfler F, Bullo F. Attack detection and identification in cyber-physical systems. *IEEE Transactions on Automatic Control*, 2013, **58**(11): 2715–2729

8 Pang Z H, Liu G P, Zhou D H, Hou F Y, Sun D H. Two-channel false data injection attacks against output tracking control of networked systems. *IEEE Transactions on Industrial Electronics*, 2016, **63**(5): 3242–3251

9 Chen B, Ho D W C, Zhang W A, Yu L. Distributed dimensionality reduction fusion estimation for cyber-physical systems under DoS attacks. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2019, **49**(2): 455–468

10 Lai S Y, Chen B, Li T X, Yu L. Packet-based state feedback control under DoS attacks in cyber-physical systems. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2019. **66**(8): 1421–1425

11 Chen B, Ho D W C, Hu G Q, Yu L. Secure fusion estimation for bandwidth constrained cyber-physical systems under replay attacks. *IEEE Transactions on Cybernetics*, 2018, **48**(6): 1862–1876

12 Ding D R, Han Q L, Xiang Y, Ge X H, Zhang X M. A survey on security control and attack detection for industrial cyber-physical systems. *Neurocomputing*, 2018, **275**: 1674–1683

13 Liu Y, Ning P, Reiter M K. False data injection attacks against state estimation in electric power grids. *ACM Transactions on Information and System Security*, 2011, **14**(1): 1–33

14 Ye N, Zhang Y B, Borror C M. Robustness of the Markov-chain model for cyber-attack detection. *IEEE Transactions on Reliability*, 2004, **53**(1): 116–123

15 Chorppath A K, Alpcan T, Boche H. Bayesian mechanisms and detection methods for wireless network with malicious users. *IEEE Transactions on Mobile Computing*, 2016, **15**(10): 2452–2465

16 Kailkhura B, Han Y S, Brahma S, Varshney P K. Distributed Bayesian detection in the presence of Byzantine data. *IEEE Transactions on Signal Processing*, 2015, **63**(9): 5250–5263

17 Kailkhura B, Han Y S, Brahma S, Varshney P K. Asymptotic analysis of distributed Bayesian detection with Byzantine data. *IEEE Signal Processing Letters*, 2015, **22**(5): 608–612

18 Rawat A S, Anand P, Chen H, Varshney P K. Collaborative spectrum sensing in the presence of Byzantine attacks in cognitive radio networks. *IEEE Transactions on Signal Processing*, 2011, **59**(2): 774–786

19 Manandhar K, Cao X J, Hu F, Liu Y. Detection of faults and attacks including false data injection attack in smart grid using Kalman filter. *IEEE Transactions on Control of Network Systems*, 2014, **1**(4): 370–379

20 Mo Y L, Chabukswar R, Sinopoli B. Detecting integrity attacks on SCADA systems. *IEEE Transactions on Control Systems Technology*, 2014, **22**(4): 1396–1407

21 Rawat D B, Bajracharya C. Detection of false data injection attacks in smart grid communication systems. *IEEE Signal Processing Letters*, 2015, **22**(10): 1652–1656

22 Liu L C, Esmalifalak M, Ding Q F, Emesih V A, Han Z. Detecting false data injection attacks on power grid by sparse optimization. *IEEE Transactions on Smart Grid*, 2014, **5**(2): 612–621

23 Deng R L, Xiao G X, Lu R X. Defending against false data injection attacks on power system state estimation. *IEEE Transactions on Industrial Informatics*, 2017, **13**(1): 198–207

24 Huang Y, Tang J, Cheng Y, Li H S, Campbell K A, Han Z. Real-time detection of false data injection in smart grid networks: An adaptive CUSUM method and analysis. *IEEE Systems Journal*, 2016, **10**(2): 532–543

25 Zhang Q H, Basseville M. Statistical detection and isolation of additive faults in linear time-varying systems. *Automatica*, 2014, **50**(10): 2527–2538

26 Zhang Q H. Adaptive Kalman filter for actuator fault diagnosis. *Automatica*, 2018, **93**: 333–342

27 Sun S L, Deng Z L. Multi-sensor optimal information fusion Kalman filter. *Automatica*, 2004, **40**(6): 1017–1023

28 Chen B, Ho D W C, Zhang W A, Yu L. Networked fusion estimation with bounded noises. *IEEE Transactions on Automatic Control*, 2017, **62**(10): 5415–5421

翁品迪 浙江工业大学硕士研究生. 主要研究方向为信息物理系统中攻击信号的融合检测.
E-mail: pwd2gg@aliyun.com
(**WENG Pin-Di** Master student at Zhejiang University of Technology. His research interest covers fusion detection of attack signal in cyber physical system.)

陈 博 浙江工业大学教授. 主要研究方向为信息融合, 安全估计与控制, 信息物理系统. 本文通信作者.
E-mail: bchen@aliyun.com
(**CHEN Bo** Professor at Zhejiang University of Technology. His research interest covers information fusion, security estimate and control, and cyber physical system. Corresponding author of this paper.)

俞 立 浙江工业大学教授. 主要研究方向为网络化控制, 信息融合, 信息物理系统. E-mail: lyu@zjut.edu.cn
(**YU Li** Professor at Zhejiang University of Technology. His research interest covers networked control, information fusion, and cyber physical system.)