隐蔽攻击下信息物理系统的安全输出反馈控制

张淇瑞[1] 孟思琪[1] 王兰豪[2] 刘 坤[3] 代 伟[1, 4]

摘 要 研究了受到隐蔽攻击的信息物理系统 (Cyber-physical system, CPS) 安全控制问题. 采用 KL (Kullback-Leibler) 散度描述攻击的隐蔽性, 并设计动态输出反馈控制器, 使系统可达集始终保持在安全区域内, 其中可达集定义为系统状态以一定概率属于的集合. 首先, 给出隐蔽攻击下检测器残差所在范围的一个外椭球近似集; 其次, 根据该近似集和噪声的范围给出控制器参数与系统椭球形不变可达集的关系; 然后, 通过设计可逆线性变换并构造凸优化问题, 求解安全控制器参数和相应的不变可达集; 最后, 使用弹簧−质量−阻尼系统进行仿真, 验证了所提控制方法的有效性.

关键词 信息物理系统, 隐蔽攻击, 安全控制, KL 散度, 可达集

引用格式 张淇瑞, 孟思琪, 王兰豪, 刘坤, 代伟. 隐蔽攻击下信息物理系统的安全输出反馈控制. 自动化学报, 2024, **50**(7): 1363−1372

DOI 10.16383/j.aas.c220893

Secure Output-feedback Control for Cyber-physical Systems Under Stealthy Attacks

ZHANG Qi-Rui[1] MENG Si-Qi[1] WANG Lan-Hao[2] LIU Kun[3] DAI Wei[1, 4]

Abstract Secure control problem of cyber-physical systems (CPS) under stealthy attacks is studied. The Kullback-Leibler (KL) divergence is adopted to describe the attack′s stealthiness. The aim is to design a secure dynamic output-feedback controller such that the reachable set, which is defined as the set that the system′s state resides in with a certain probability, resides in a safe set. Firstly, an ellipsoidal outer approximation for the set of residual under stealthy attacks is given. Secondly, based on the approximation and the ranges of noises, the relationship between controller′s parameters and the ellipsoidal invariant reachable set is analyzed. Thirdly, a convex optimization problem is constructed by designing an invertible linear transformation. Parameters of the secure controller and the corresponding invariant reachable set are obtained by solving the problem. Finally, a simulation of the spring-mass-damping system is given to verify the effectiveness of the proposed controller.

Key words Cyber-physical system (CPS), stealthy attack, secure control, KL (Kullback-Leibler) divergence, reachable set

Citation Zhang Qi-Rui, Meng Si-Qi, Wang Lan-Hao, Liu Kun, Dai Wei. Secure output-feedback control for cyber-physical systems under stealthy attacks. *Acta Automatica Sinica*, 2024, **50**(7): 1363−1372

近年来, 以计算机技术、通信技术和控制技术为核心的信息物理系统 (Cyber-physical system, CPS)[1−2] 飞速发展并广泛应用于国防军事、工业生产、智能交通、健康医疗等诸多领域, 推动着经济发展和社会进步, 对人类的生产和生活产生巨大影响. CPS 中各节点通过通信网络相互连接, 这使系统内部信息资源可以进行共享, 系统设计成本降低, 便于扩展和维护, 提高系统工作效率. 然而, 系统的运行环境也因此由封闭和隔离变得开放和互联, 导致系统容易受到各种恶意攻击的影响. CPS 一旦受到恶意攻击, 系统本身会受到严重破坏, 并随之带来

收稿日期 2022-11-16 录用日期 2023-03-21
Manuscript received November 16, 2022; accepted March 21, 2023
江苏省自然科学基金 (BK20231062, BK20200086), 中央高校基本科研业务费专项资金 (2023QN1074), 国家自然科学基金 (62373361, 61973306, 62273041, 52304309), 流程工业综合自动化国家重点实验室联合开放基金 (2020-KF-21-10, 2021-KF-21-05), 矿冶过程自动控制技术国家重点实验室开放基金 (BGRIMM-KZSKL-2022-7), 江苏省研究生科研与实践创新计划 (KYCX23_2717), 中国矿业大学研究生创新计划项目 (2023WLJCRCZL117) 资助
Supported by Natural Science Foundation of Jiangsu Province (BK20231062, BK20200086), Fundamental Research Funds for the Central Universities (2023QN1074), National Natural Science Foundation of China (62373361, 61973306, 62273041, 52304309), Joint Open Foundation of State Key Laboratory of Synthetical Automation for Process Industries (2020-KF-21-10, 2021-KF-21-05), Open Foundation of State Key Laboratory of Process Automation in Mining and Metallurgy (BGRIMM-KZSKL-2022-7), Postgraduate Research & Practice Innovation Program of Jiangsu Province (KYCX23_2717), and Graduate Innovation Program of China University of Mining and Technology (2023WLJCRCZL117)
本文责任编委 曹向辉
Recommended by Associate Editor CAO Xiang-Hui
1. 中国矿业大学信息与控制工程学院 徐州 221116 2. 中国矿业大学国家煤加工与洁净化工程技术研究中心 徐州 221116 3. 北京理工大学自动化学院 北京 100081 4. 中国矿业大学人工智能研究院 徐州 221116
1. School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116 2. National Engineering Research Center of Coal Preparation and Purification, China University of Mining and Technology, Xuzhou 221116 3. School of Automation, Beijing Institute of Technology, Beijing 100081 4. Artificial Intelligence Research Institute, China University of Mining and Technology, Xuzhou 221116

巨大损失. 因此, 对 CPS 安全问题的研究具有重要意义和极大的紧迫性[3−5].

攻击检测[6−9] 是安全防护极为重要的环节, 一旦攻击被检测器发现, 防御者就可以设计弹性滤波器和控制器[3−5], 来抵消该攻击的影响. 有些具有足够系统信息和攻击能力的攻击者可以设计隐蔽攻击. 为了评估隐蔽攻击的危险性, 许多工作研究了隐蔽性攻击下系统的性能损失, 如控制性能[10−11]、估计性能[12−13] 和可达集[14−27] 等.

文献 [14] 最早研究了隐蔽攻击下系统可达集, 其采用检测器残差变化量的二范数来描述攻击的隐蔽性, 并给出只有传感器受到隐蔽攻击情况下的系统可达集有界的充要条件. 文献 [15] 在文献 [14] 基础上, 进一步考虑传感器和执行器均受到攻击的情况, 但只给出可达集有界的充分条件, 文献 [16] 则给出了充要条件. 文献 [17] 将文献 [16] 的结论扩展到分布式系统, 并且分别讨论采用静态滤波器和动态滤波器构造检测器残差的情况. 文献 [18] 进一步给出在稀疏传感器隐蔽攻击下可达集有界的充要条件. 文献 [19] 在文献 [18] 的基础上, 进一步研究了传感器信号和滤波器的估计值同时受到攻击的情况. 文献 [20] 研究了攻击者在部分时刻无法攻击时可达集的有界性. 文献 [21] 在上述文献考虑的隐蔽性的基础上, 额外要求检测器残差的变化量逐渐趋于 0, 并设计使可达集无界的最优隐蔽攻击.

对于有界的可达集, 需要进一步研究可达集的近似计算方法. 文献 [22] 提出一种可达集的外椭球和内椭球近似集的迭代计算方法; 文献 [23] 通过构造凸优化问题以求解可达集的外椭球近似集; 文献 [24] 计算精确的椭球可达集, 但只适用于 SPRT 检测器; 文献 [25] 研究了可达集的凸多面体近似算法.

以上研究均只考虑攻击的能量, 未考虑攻击的随机特性对隐蔽性的影响. 为解决这个问题, 文献 [28–29] 提出可以采用一种描述 2 个概率密度分布函数间距离指标——KL (Kullback-Leibler)[30] 散度来描述隐蔽性; 文献 [26] 采用 KL 散度作为隐蔽性指标, 给出可达集有界的充要条件, 并提出可达集的外椭球近似算法.

研究可达集的有界性以及近似计算方法[14−26] 可以用来分析系统的状态是否在安全范围以内, 以评估系统安全风险, 当系统不安全时, 则需要重新设计系统参数. 文献 [23] 提出可以通过减小可达集体积来降低系统风险, 并给出相应控制器增益设计方法; 文献 [27] 进一步协同设计滤波器和控制器增益. 但是文献 [23, 27] 提出的方法与系统安全范围无关, 设计后的参数无法保证系统状态能离开危险区域.

因此, 本文研究隐蔽攻击下 CPS 的安全控制方法, 旨在设计动态输出反馈控制器, 使系统状态始终保持在安全范围之内. 本文的主要贡献如下:

1) 本文采用 KL 散度作为攻击的隐蔽性指标, 通过构造优化问题, 给出动态输出反馈控制器参数与系统椭球形不变可达集的关系.

2) 本文设计一种可逆的线性变换构造凸优化问题, 求解该凸优化问题以获得控制器参数, 使系统状态始终保持在安全范围之内.

描述. $\mathbf{R}^n$ 表示 n 维欧几里得空间; I_n 表示 n 阶单位矩阵; $A \succ 0$ ($A \succeq 0$) 表示矩阵 A 是正定 (半正定) 矩阵; $\mathrm{Tr}(A)$ 表示矩阵 A 的迹; $A^{\dagger}$ 表示矩阵 A 的 Moore-Penrose 广义逆; $\boldsymbol{x} \sim \mathrm{N}(\boldsymbol{a}, \Sigma)$ 表示向量 $\boldsymbol{x}$ 服从均值为 $\boldsymbol{a}$、协方差为 Σ 的高斯分布; $\mathrm{diag}\{A_1, \cdots, A_n\}$ 表示主对角线上为矩阵 $A_1, \cdots, A_n$ 的分块对角矩阵; $\mathbf{0}$ 表示所有元素均为 0 的矩阵.

1 问题描述

考虑如下的线性时不变系统:

$$\boldsymbol{x}_{k+1} = A\boldsymbol{x}_k + B\boldsymbol{u}_k + D_1\boldsymbol{u}_k^a + \boldsymbol{w}_k \tag{1}$$

$$\boldsymbol{y}_k = C\boldsymbol{x}_k + D_2\boldsymbol{y}_k^a + \boldsymbol{v}_k \tag{2}$$

式中, $\boldsymbol{x}_k \in \mathbf{R}^n$ 是被控对象的状态, $\boldsymbol{u}_k \in \mathbf{R}^l$ 是控制输入, $\boldsymbol{u}_k^a \in \mathbf{R}^p$ 是针对执行器的欺骗攻击, $\boldsymbol{w}_k \in \mathbf{R}^n$ 是过程噪声, $\boldsymbol{y}_k \in \mathbf{R}^m$ 是系统输出, $\boldsymbol{y}_k^a \in \mathbf{R}^q$ 是针对传感器的欺骗攻击, $\boldsymbol{v}_k \in \mathbf{R}^m$ 是量测噪声, A、B、C、D_1 和 D_2 为合适维数的矩阵, $\boldsymbol{w}_k$ 和 $\boldsymbol{v}_k$ 均为独立同分布的零均值高斯噪声即 $\boldsymbol{w}_k \sim \mathrm{N}(0, \Sigma_w)$ 和 $\boldsymbol{v}_k \sim \mathrm{N}(0, \Sigma_v)$, $\Sigma_w \succ 0$ 和 $\Sigma_v \succ 0$ 分别为 $\boldsymbol{w}_k$ 和 $\boldsymbol{v}_k$ 的协方差矩阵, $\boldsymbol{w}_k$ 和 $\boldsymbol{v}_k$ 相互独立, (A, B) 是可控的, (C, A) 是可观的, D_1 和 D_2 均是列满秩的.

系统采用卡尔曼滤波器来估计状态 $\boldsymbol{x}_k$. 假设卡尔曼滤波器的增益收敛到稳态值, 即:

$$\hat{\boldsymbol{x}}_k = \hat{\boldsymbol{x}}_k^- + K(\boldsymbol{y}_k - C\hat{\boldsymbol{x}}_k^-) \tag{3}$$

$$\hat{\boldsymbol{x}}_k^- = A\hat{\boldsymbol{x}}_{k-1} + B\boldsymbol{u}_{k-1} \tag{4}$$

式中, $\hat{\boldsymbol{x}}_k$ 为对状态 $\boldsymbol{x}_k$ 的估计, 滤波器增益为 $K = PC^{\mathrm{T}}(CPC^{\mathrm{T}} + \Sigma_v)^{-1}$, $P = APA^{\mathrm{T}} + \Sigma_w - APC^{\mathrm{T}} \times (CPC^{\mathrm{T}} + \Sigma_v)^{-1}CPA^{\mathrm{T}}$. 由于系统是可观的, 因此 $A - KCA$ 是稳定的.

分别定义卡尔曼滤波器的估计误差和残差为 $\boldsymbol{e}_k = \boldsymbol{x}_k - \hat{\boldsymbol{x}}_k$ 和 $\boldsymbol{r}_k = \boldsymbol{y}_k - C\hat{\boldsymbol{x}}_k^-$. 将系统不受攻击时 (即 $\boldsymbol{u}_{i-1}^a = \mathbf{0}$, $\boldsymbol{y}_i^a = \mathbf{0}$, $\forall i \le k$) 的残差记作 $\bar{\boldsymbol{r}}_k$, 其服从独立同分布的高斯分布 $\mathrm{N}(0, Q)$, 其中 $Q = CPC^{\mathrm{T}} + \Sigma_v$. 攻击检测器利用残差 $\boldsymbol{r}_k$ 根据某种规则来判断系统是否受到攻击.

为保证系统安全, 采用如下形式的动态输出反馈控制器:

$$\boldsymbol{z}_{k+1} = E\boldsymbol{z}_k + F\boldsymbol{y}_{k+1} \tag{5}$$

$$\boldsymbol{u}_k = G\boldsymbol{z}_k \tag{6}$$

式中, $\boldsymbol{z}_k \in \mathbf{R}^n$ 是控制器的状态, E、F 和 G 为待设计的控制器参数.

KL 散度是一种描述 2 个随机变量之间距离的物理量, 在攻击检测研究领域中, 它可以被用来描述攻击相对于检测器的隐蔽性. 下面给出 KL 散度的定义.

定义 1[28, 30]**.** 假设 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 为 2 个维数相同的随机变量, 它们的概率密度分布函数分别为 $f(\boldsymbol{x};\boldsymbol{\xi})$ 和 $f(\boldsymbol{y};\boldsymbol{\xi})$, 那么 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 之间的 KL 散度为:

$$D(\boldsymbol{x}||\boldsymbol{y}) = \int_{\{\boldsymbol{\xi}|f(\boldsymbol{x};\boldsymbol{\xi})>0\}} f(\boldsymbol{x};\boldsymbol{\xi}) \ln \frac{f(\boldsymbol{x};\boldsymbol{\xi})}{f(\boldsymbol{y};\boldsymbol{\xi})} \mathrm{d}\boldsymbol{\xi}$$

注 1. 定义 1 中对向量 $\boldsymbol{\xi} = [\xi_1, \xi_2, \cdots, \xi_n]^{\mathrm{T}}$ 的积分是指对每个元素 $\xi_1, \xi_2, \cdots, \xi_n$ 进行逐个积分.

采用未受攻击和受到攻击时卡尔曼滤波器的残差间的 KL 散度 $D(\bar{\boldsymbol{r}}_k||\boldsymbol{r}_k)$ 来描述攻击的隐蔽性. 为保证隐蔽性, 攻击者需要使 $D(\bar{\boldsymbol{r}}_k||\boldsymbol{r}_k)$ 不大于一个正的阈值 δ, 即:

$$D(\bar{\boldsymbol{r}}_k||\boldsymbol{r}_k) \leq \delta$$

KL 散度 $D(\bar{\boldsymbol{r}}_k||\boldsymbol{r}_k)$ 与检测器漏警率收敛到 0 的速度上界有关[29], 一般 $D(\bar{\boldsymbol{r}}_k||\boldsymbol{r}_k)$ 越小, 漏警率收敛到 0 的速度越快, 攻击越难以保持隐蔽. 因此, 可根据系统所能容忍的漏警率收敛速度来确定阈值 δ.

不失一般性, 假设执行器攻击 $\boldsymbol{u}_k^a$ 和传感器攻击 $\boldsymbol{y}_k^a$ 分别从时刻 0 和时刻 1 开始. 此外, 根据文献 [10−29], 还对攻击者作出以下假设.

假设 1. 攻击者获得足够的系统信息 (系统信息指 A、B、C、D_1、D_2、E、F、G、K、Σ_w、Σ_v 和 δ) 使其能够设计隐蔽攻击.

假设 2. 受攻击时的卡尔曼滤波器的残差服从分布 $\mathrm{N}(\boldsymbol{\eta}_k, \Sigma_k^{-1})$, 其中, 均值 $\boldsymbol{\eta}_k \in \mathbf{R}^m$, 协方差 $\Sigma_k^{-1} \succ 0$.

注 2. 为保证攻击的隐蔽性, 攻击者往往需要获取系统中的关键参数, 假设 1 允许攻击者获得足够的系统信息设计隐蔽攻击. 假设 2 包含常见的攻击形式, 例如文献 [14−19, 21] 中的非随机攻击和文献 [10−13, 28−29] 中服从高斯分布的攻击. 需要指出的是, 由于攻击可能与系统的输入和输出数据相关, 因此残差的协方差不一定满足 $\Sigma_k^{-1} \succeq Q$.

为给出闭环系统的状态方程, 定义闭环系统的状态 $\boldsymbol{\zeta}_k = [\boldsymbol{x}_k^{\mathrm{T}} \ \ \boldsymbol{z}_k^{\mathrm{T}} \ \ (\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k)^{\mathrm{T}}]^{\mathrm{T}}$, 定义攻击信号 $\boldsymbol{\xi}_k = \begin{bmatrix} \boldsymbol{u}_k^a \\ \boldsymbol{y}_{k+1}^a \end{bmatrix}$, 根据式 (1) ~ (6), 可得:

$$\boldsymbol{\zeta}_{k+1} = \bar{A}\boldsymbol{\zeta}_k + \bar{B}_w\boldsymbol{w}_k + \bar{B}_v\boldsymbol{v}_{k+1} + \bar{D}_1\boldsymbol{\xi}_k \tag{7}$$

$$\boldsymbol{r}_{k+1} = \bar{C}\boldsymbol{\zeta}_k + C\boldsymbol{w}_k + \boldsymbol{v}_{k+1} + \bar{D}_2\boldsymbol{\xi}_k \tag{8}$$

式中,

$$\bar{A} = \begin{bmatrix} A & BG & \mathbf{0} \\ FCA & E+FCBG & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & A-KCA \end{bmatrix}$$

$$\bar{B}_w = \begin{bmatrix} I_n \\ FC \\ I_n - KC \end{bmatrix}, \quad \bar{B}_v = \begin{bmatrix} \mathbf{0} \\ F \\ -K \end{bmatrix}$$

$$\bar{C} = [\mathbf{0} \quad \mathbf{0} \quad CA]$$

$$\bar{D}_1 = \begin{bmatrix} D_1 & \mathbf{0} \\ FCD_1 & FD_2 \\ D_1 - KCD_1 & -KD_2 \end{bmatrix}$$

$$\bar{D}_2 = [CD_1 \quad D_2]$$

此外, 根据文献 [26], 为保证可达集的有界性, 假设矩阵 $\bar{D}_2$ 是列满秩的. 那么根据式 (7)、式 (8), 可得:

$$\boldsymbol{\zeta}_{k+1} = \mathcal{A}\boldsymbol{\zeta}_k + \mathcal{B}_w\boldsymbol{w}_k + \mathcal{B}_v\boldsymbol{v}_{k+1} + \mathcal{B}_r\boldsymbol{r}_{k+1} \tag{9}$$

式中

$$\mathcal{A} = \begin{bmatrix} A & BG & -\tilde{D}\bar{D}_2^{\dagger}CA \\ FCA & E+FCBG & -F\bar{D}_2\bar{D}_2^{\dagger}CA \\ \mathbf{0} & \mathbf{0} & A-\tilde{D}\bar{D}_2^{\dagger}CA \end{bmatrix}$$

$$\mathcal{B}_w = \begin{bmatrix} I_n - \tilde{D}\bar{D}_2^{\dagger}C \\ FC - F\bar{D}_2\bar{D}_2^{\dagger}C \\ I_n - \tilde{D}\bar{D}_2^{\dagger}C \end{bmatrix}, \quad \mathcal{B}_v = \begin{bmatrix} -\tilde{D}\bar{D}_2^{\dagger} \\ F - F\bar{D}_2\bar{D}_2^{\dagger} \\ -\tilde{D}\bar{D}_2^{\dagger} \end{bmatrix}$$

$$\mathcal{B}_r = \begin{bmatrix} \tilde{D}\bar{D}_2^{\dagger} \\ F\bar{D}_2\bar{D}_2^{\dagger} \\ \tilde{D}\bar{D}_2^{\dagger} - K \end{bmatrix}, \quad \tilde{D} = [D_1 \quad \mathbf{0}]$$

注意到噪声 $\boldsymbol{w}_k$、$\boldsymbol{v}_k$ 和残差 $\boldsymbol{r}_k$ 服从高斯分布, 这意味着 $\boldsymbol{\zeta}_k$ 可以为任意向量, 因此考虑隐蔽攻击下 $\boldsymbol{\zeta}_k$ 以一定概率属于的集合[26−27]. 根据系统的随机特性, 定义系统的可达集为:

$$\begin{aligned} \mathcal{R}(\boldsymbol{\zeta}_k) = \{&\boldsymbol{\zeta}_k | (\boldsymbol{r}_i - \boldsymbol{\eta}_i)^{\mathrm{T}}\Sigma_i(\boldsymbol{r}_i - \boldsymbol{\eta}_i) \leq a, \\ &\boldsymbol{w}_{i-1}^{\mathrm{T}}\Sigma_w^{-1}\boldsymbol{w}_{i-1} \leq b, \boldsymbol{v}_i^{\mathrm{T}}\Sigma_v^{-1}\boldsymbol{v}_i \leq c, \\ &D(\bar{\boldsymbol{r}}_i||\boldsymbol{r}_i) \leq \delta, i \leq k\} \end{aligned} \tag{10}$$

式中, $a > 0$, $b > 0$ 和 $c > 0$ 为常数.

除隐蔽性约束 $D(\bar{\boldsymbol{r}}_i||\boldsymbol{r}_i) \leq \delta$ 外, 可达集 $\mathcal{R}(\boldsymbol{\zeta}_k)$ 包含额外 3 个约束, 即残差 $\boldsymbol{r}_k$、过程噪声 $\boldsymbol{w}_{k-1}$ 和量测噪声 $\boldsymbol{v}_k$ 分别属于椭球 $(\boldsymbol{r}_k - \boldsymbol{\eta}_k)^{\mathrm{T}}\Sigma_k(\boldsymbol{r}_k - \boldsymbol{\eta}_k) \leq a$、$\boldsymbol{w}_{k-1}^{\mathrm{T}}\Sigma_w^{-1}\boldsymbol{w}_{k-1} \leq b$ 和 $\boldsymbol{v}_k^{\mathrm{T}}\Sigma_v^{-1}\boldsymbol{v}_k \leq c$. 注意到 $(\boldsymbol{r}_k - \boldsymbol{\eta}_k)^{\mathrm{T}}\Sigma_k(\boldsymbol{r}_k - \boldsymbol{\eta}_k)$、$\boldsymbol{w}_{k-1}^{\mathrm{T}}\Sigma_w^{-1}\boldsymbol{w}_{k-1}$ 和 $\boldsymbol{v}_k^{\mathrm{T}}\Sigma_v^{-1}\boldsymbol{v}_k$ 分别服从自由度为 m、n 和 m 的卡方分布, 那么 $\boldsymbol{r}_k$、$\boldsymbol{w}_{k-1}$ 和 $\boldsymbol{v}_k$ 属于对应椭球中的概率分别为 $F(a, m)$、$F(b, n)$ 和 $F(c, m)$, 其中 F 为卡方分布的累积分布函数. 此外, 由于 m 和 n 已知, 那么参数 a、b 和 c 可通过防御者期望考察的概率 $F(a, m)$、$F(b, n)$ 和 $F(c, m)$ 来

预先确定, 如参数 a 可通过求解下列方程获得:

$$\gamma\left(\frac{m}{2},\frac{a}{2}\right)=(1-F(a,m))\,\bar{\gamma}\left(\frac{m}{2}\right)$$

式中, $\bar{\gamma}(\cdot)$ 和 $\gamma(\cdot)$ 分别表示伽马函数和不完全伽马函数.

在实际应用中, 通常关注物理对象状态 $\boldsymbol{x}_k$ 是否在安全范围内. 因此, 类似式 (10), 进一步定义状态 $\boldsymbol{x}_k$ 的可达集为:

$$\begin{aligned}\mathcal{R}(\boldsymbol{x}_k)=\{\boldsymbol{x}_k|&(\boldsymbol{r}_i-\boldsymbol{\eta}_i)^{\mathrm{T}}\Sigma_i(\boldsymbol{r}_i-\boldsymbol{\eta}_i)\le a,\\&\boldsymbol{w}_{i-1}^{\mathrm{T}}\Sigma_w^{-1}\boldsymbol{w}_{i-1}\le b,\,\boldsymbol{v}_i^{\mathrm{T}}\Sigma_v^{-1}\boldsymbol{v}_i\le c,\\&D(\bar{\boldsymbol{r}}_i||\boldsymbol{r}_i)\le\delta,\,i\le k\}\end{aligned}\tag{11}$$

当 $\mathcal{R}(\boldsymbol{\zeta}_k)$ 已知时, 将 $\mathcal{R}(\boldsymbol{\zeta}_k)$ 投影到 $\boldsymbol{x}_k$ 超平面, 即可获得 $\mathcal{R}(\boldsymbol{x}_k)$[31].

假设系统状态的安全区域为椭球 $\mathcal{E}_s(\Phi)=\{\boldsymbol{x}_k\in\mathbf{R}^n|{\boldsymbol{x}_k}^{\mathrm{T}}\Phi\boldsymbol{x}_k\le 1\}$, 其中 $\Phi\succeq 0$ 是描述椭球形状的矩阵. 下面将设计矩阵 E、F 和 G, 使 $\mathcal{R}(\boldsymbol{x}_k)\subset\mathcal{E}_s(\Phi)$, $\forall k\ge 0$.

2 主要结论

若要保证 $\boldsymbol{x}_k$ 始终在安全集 $\mathcal{E}_s(\Phi)$ 内, 需要使 $\mathcal{R}(\boldsymbol{\zeta}_k)$ 是系统的一个不变可达集 (即若 $\boldsymbol{\zeta}_k\in\mathcal{R}(\boldsymbol{\zeta}_k)$, 则 $\boldsymbol{\zeta}_j\in\mathcal{R}(\boldsymbol{\zeta}_k)$, $\forall j\ge k$), 并且保证 $\mathcal{R}(\boldsymbol{\zeta}_k)$ 在 $\boldsymbol{x}_k$ 超平面的投影 $\mathcal{R}(\boldsymbol{x}_k)$ 是安全集 $\mathcal{E}_s(\Phi)$ 的子集. 本文考虑系统的一种椭球形的不变可达集, 给出其与控制器式 (5) 和式 (6) 的关系, 并进一步设计控制器参数, 保证系统安全.

2.1 椭球形不变可达集

首先, 介绍用于计算系统椭球形不变可达集的引理.

引理 1. 假设某系统在 k 时刻的状态为 $\boldsymbol{\rho}_k\in\mathbf{R}^n$ 并受到噪声 $\boldsymbol{w}_{i,k}\in\mathbf{R}^{n_i}$ 的影响, 其中 $\boldsymbol{w}_{i,k}^{\mathrm{T}}W_i\boldsymbol{w}_{i,k}\le 1$, $W_i\succ 0$ 是描述噪声范围的矩阵, $i=1,\cdots,\theta$. 给定一个常数 $\alpha\in(0,1)$, 如果存在常数 $\alpha_{i,k}\in(0,1)$, $i=1,\cdots,\theta$, 满足 $\sum_{i=1}^{\theta}\alpha_{i,k}\le 1-\alpha$ 和矩阵 $\Gamma\succeq 0$, 使不等式:

$$\begin{aligned}&\boldsymbol{\rho}_{k+1}^{\mathrm{T}}\Gamma\boldsymbol{\rho}_{k-1}-\alpha\boldsymbol{\rho}_k^{\mathrm{T}}\Gamma\boldsymbol{\rho}_k-\\&\sum_{i-1}^{\theta}\alpha_{i,k}\boldsymbol{w}_{i,k}^{\mathrm{T}}W_i\boldsymbol{w}_{i,k}\le 0\end{aligned}\tag{12}$$

成立, 那么椭球 $\{\boldsymbol{\rho}_k|\boldsymbol{\rho}_k^{\mathrm{T}}\Gamma\boldsymbol{\rho}_k\le 1\}$ 是该系统的 1 个不变可达集.

证明. 取李雅普诺夫函数为 $V_k=\boldsymbol{\rho}_k^{\mathrm{T}}\Gamma\boldsymbol{\rho}_k$, 由式 (12), 可得:

$$V_{k+1}\le\alpha\boldsymbol{\rho}_k^{\mathrm{T}}\Gamma\boldsymbol{\rho}_k+1-\alpha$$

因此, 有:

$$V_{k+1}-V_k\le(1-\boldsymbol{\rho}_k^{\mathrm{T}}\Gamma\boldsymbol{\rho}_k)(1-\alpha)$$

当 $\boldsymbol{\rho}_k^{\mathrm{T}}\Gamma\boldsymbol{\rho}_k>1$ 时, 有 $V_{k+1}-V_k<0$. 这意味着当系统状态在椭球 $\boldsymbol{\rho}_k^{\mathrm{T}}\Gamma\boldsymbol{\rho}_k\le 1$ 外时, 会逐渐收敛到椭球 $\boldsymbol{\rho}_k^{\mathrm{T}}\Gamma\boldsymbol{\rho}_k\le 1$ 内, 所以椭球 $\boldsymbol{\rho}_k^{\mathrm{T}}\Gamma\boldsymbol{\rho}_k\le 1$ 是一个不变可达集. □

式 (9) 噪声 $\boldsymbol{w}_k$ 和 $\boldsymbol{v}_{k+1}$ 的取值范围分别为椭球 $\boldsymbol{w}_k^{\mathrm{T}}(\Sigma_w^{-1}/b)\boldsymbol{w}_k\le 1$ 和 $\boldsymbol{v}_{k+1}^{\mathrm{T}}(\Sigma_v^{-1}/c)\boldsymbol{v}_{k+1}\le 1$, 而残差所有可能的取值组成集合为:

$$\begin{aligned}\mathcal{E}_r=\{\boldsymbol{r}_{k+1}|&(\boldsymbol{r}_{k+1}-\boldsymbol{\eta}_{k+1})^{\mathrm{T}}\Sigma_{k+1}(\boldsymbol{r}_{k+1}-\boldsymbol{\eta}_{k+1})\le a,\\&D(\bar{\boldsymbol{r}}_{k+1}||\boldsymbol{r}_{k+1})\le\delta\}\end{aligned}$$

在利用引理 1 计算系统式 (9) 的不变可达集前, 需要先计算 $\mathcal{E}_r$ 的外椭球近似集. 设残差 $\bar{\boldsymbol{r}}_k$ 的协方差矩阵 Q 的特征值分解为:

$$Q=\bar{U}^{\mathrm{T}}\mathrm{diag}\{\mu_1,\mu_2,\cdots,\mu_m\}\bar{U}$$

式中, $\bar{U}$ 是正交矩阵, $\mu_1,\mu_2,\cdots,\mu_m$ 为矩阵 Q 的特征值, 且 $\mu_i>0$, $i=1,\cdots,m$.

令

$$\begin{aligned}\tilde{U}&=\mathrm{diag}\{\sqrt{\mu}_1,\sqrt{\mu}_2,\cdots,\sqrt{\mu}_m\}\\\iota^*&=\min\{\iota|\iota-\ln\iota-m\le 2\delta\}\end{aligned}$$

并定义以下优化问题.

优化问题 1.

$$\begin{aligned}&\max_{\Sigma_r}\mathrm{Tr}(\Sigma_r)\\&\text{s.t.}\\&\begin{bmatrix}\iota^*I_m-\tilde{U}\bar{U}\Sigma_r\bar{U}^{\mathrm{T}}\tilde{U} & -\tilde{U}\bar{U}\Sigma_r\bar{U}^{\mathrm{T}}\tilde{U}\\-\tilde{U}\bar{U}\Sigma_r\bar{U}^{\mathrm{T}}\tilde{U} & \iota^*I_m-\tilde{U}\bar{U}\Sigma_r\bar{U}^{\mathrm{T}}\tilde{U}\end{bmatrix}\succeq 0\end{aligned}\tag{13}$$

那么, 基于优化问题 1, 可以给出 $\mathcal{E}_r$ 的一个椭球形近似.

引理 2. 假设优化问题 1 的最优解为 $\Sigma_r=\Sigma_r^*$, 那么:

$$\mathcal{E}_r\subset\left\{\boldsymbol{r}_{k+1}^{\mathrm{T}}\frac{\Sigma_r^*}{a+2\delta}\boldsymbol{r}_{k+1}\le 1\right\}\tag{14}$$

证明. 由于 $\boldsymbol{r}_k$ 和 $\bar{\boldsymbol{r}}_k$ 是服从高斯分布的, 根据定义 1, 有:

$$\begin{aligned}D(\bar{\boldsymbol{r}}_k||\boldsymbol{r}_k)=\frac{1}{2}[&\mathrm{Tr}(\Sigma_kQ)-\ln(|Q||\Sigma_k|)-m+\\&\boldsymbol{\eta}_k^{\mathrm{T}}\Sigma_k\boldsymbol{\eta}_k]\end{aligned}\tag{15}$$

因此, Σ_{k+1} 满足:

$$\mathrm{Tr}(\Sigma_{k+1}Q)-\ln(|Q||\Sigma_{k+1}|)-m\le 2\delta\tag{16}$$

且

$$\boldsymbol{\eta}_{k+1}^{\mathrm{T}}\Sigma_{k+1}\boldsymbol{\eta}_{k+1} \le 2\delta + m + \ln(|Q||\Sigma_{k+1}|) - \mathrm{Tr}(\Sigma_{k+1}Q) \tag{17}$$

因此

$$\mathcal{E}_r = \bigcup_{\Sigma_{k+1}\in\{\Sigma_{k+1}|\text{式 (16)}\}} \mathcal{E}_r(\Sigma_{k+1}) \tag{18}$$

式中

$$\mathcal{E}_r(\Sigma_{k+1}) = \{\boldsymbol{r}_{k+1}|(\boldsymbol{r}_{k+1}-\boldsymbol{\eta}_{k+1})^{\mathrm{T}} \times \Sigma_{k+1}(\boldsymbol{r}_{k+1}-\boldsymbol{\eta}_{k+1}) \le a, \text{式 (17)}\}$$

根据式 (17) 和 $(\boldsymbol{r}_{k+1}-\boldsymbol{\eta}_{k+1})^{\mathrm{T}}\Sigma_{k+1}(\boldsymbol{r}_{k+1}-\boldsymbol{\eta}_{k+1}) \le a$, 如果存在矩阵 $\Sigma_r \succ 0$, 满足 $\boldsymbol{r}_{k+1}^{\mathrm{T}}\Sigma_r\boldsymbol{r}_{k+1} \le (\boldsymbol{r}_{k+1}-\boldsymbol{\eta}_{k+1})^{\mathrm{T}}\Sigma_{k+1}(\boldsymbol{r}_{k+1}-\boldsymbol{\eta}_{k+1}) + \boldsymbol{\eta}_{k+1}^{\mathrm{T}}\Sigma_{k+1}\boldsymbol{\eta}_{k+1}$, 即:

$$\begin{bmatrix}\boldsymbol{\eta}_{k+1}\\ \boldsymbol{r}_{k+1}-\boldsymbol{\eta}_{k+1}\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}\Sigma_{k+1}-\Sigma_r & -\Sigma_r\\ -\Sigma_r & \Sigma_{k+1}-\Sigma_r\end{bmatrix}\times \begin{bmatrix}\boldsymbol{\eta}_{k+1}\\ \boldsymbol{r}_{k+1}-\boldsymbol{\eta}_{k+1}\end{bmatrix} \ge 0 \tag{19}$$

那么, $\boldsymbol{r}_{k+1}^{\mathrm{T}}\Sigma_r\boldsymbol{r}_{k+1} \le d$, 其中:

$$d = a + 2\delta + m + \ln(|Q||\Sigma_{k+1}|) - \mathrm{Tr}(\Sigma_{k+1}Q)$$

注意到 $\tilde{U}\bar{U}$ 是满秩的, 因此不等式 (19) 等价于线性矩阵不等式:

$$\begin{bmatrix}\tilde{U}\bar{U} & \mathbf{0}\\ \mathbf{0} & \tilde{U}\bar{U}\end{bmatrix}\begin{bmatrix}\Sigma_{k+1}-\Sigma_r & -\Sigma_r\\ -\Sigma_r & \Sigma_{k+1}-\Sigma_r\end{bmatrix}\times \begin{bmatrix}\bar{U}^{\mathrm{T}}\tilde{U} & \mathbf{0}\\ \mathbf{0} & \bar{U}^{\mathrm{T}}\tilde{U}\end{bmatrix} \succeq 0 \tag{20}$$

因此, 若式 (20) 满足, 则有:

$$\mathcal{E}_r(\Sigma_{k+1}) \subset \{\boldsymbol{r}_{k+1}|\boldsymbol{r}_{k+1}^{\mathrm{T}}\Sigma_r\boldsymbol{r}_{k+1} \le d\} \tag{21}$$

为了使式 (20) 对所有 $\Sigma_{k+1} \in \{\Sigma_{k+1}|\text{式 (16)}\}$ 均成立, 需要使用 $\tilde{U}\bar{U}\Sigma_{k+1}\bar{U}^{\mathrm{T}}\tilde{U}$ 的下界替换式 (20) 中的 $\tilde{U}\bar{U}\Sigma_{k+1}\bar{U}^{\mathrm{T}}\tilde{U}$.

注意到:

$$\mathrm{Tr}(\Sigma_{k+1}Q) - \ln(|Q||\Sigma_{k+1}|) = \mathrm{Tr}(\tilde{U}\bar{U}\Sigma_{k+1}\bar{U}^{\mathrm{T}}\tilde{U}) - \ln|\tilde{U}\bar{U}\Sigma_{k+1}\bar{U}^{\mathrm{T}}\tilde{U}| = \sum_{i=1}^{m}(\iota_i - \ln(\iota_i)) \tag{22}$$

式中, ι_i, $i = 1, \cdots, m$, 为 $\tilde{U}\bar{U}\Sigma_{k+1}\bar{U}^{\mathrm{T}}\tilde{U}$ 的特征值.

由于 $\iota_i - \ln(\iota_i) \ge 0$, 所以根据式 (16) 和式 (22), 有:

$$\iota_i - \ln(\iota_i) - m \le 2\delta$$

因此, 对于所有满足不等式 (16) 的 Σ_{k+1}, 矩阵 $\tilde{U}\bar{U}\Sigma_{k+1}\bar{U}^{\mathrm{T}}\tilde{U}$ 的特征值始终满足 $\iota_i \ge \iota^*$, 所以 $\tilde{U}\bar{U}\Sigma_{k+1}\bar{U}^{\mathrm{T}}\tilde{U} \succeq \iota^* I_m$. 故如果式 (13) 成立, 那么式 (20) 对于所有 Σ_{k+1} 总成立. 然后, 为使 $\boldsymbol{r}_{k+1}^{\mathrm{T}}\Sigma_r\boldsymbol{r}_{k+1} \le d$ 尽可能贴合 $\mathcal{E}_r(\Sigma_{k+1})$, 需要求解优化问题 1 得到最优 Σ_r 的取值 Σ_r^*.

由于椭球 $\boldsymbol{r}_{k+1}^{\mathrm{T}}\Sigma_r^*\boldsymbol{r}_{k+1} \le d$ 的体积与 d 正相关, 即对于 $0 \le d_{small} < d_{large}$, 有 $\{\boldsymbol{r}_{k+1}|\boldsymbol{r}_{k+1}^{\mathrm{T}}\Sigma_r^*\boldsymbol{r}_{k+1} \le d_{small}\} \subset \{\boldsymbol{r}_{k+1}|\boldsymbol{r}_{k+1}^{\mathrm{T}}\Sigma_r^*\boldsymbol{r}_{k+1} \le d_{large}\}$. 所以, 根据式 (18) 和式 (21), 有:

$$\mathcal{E}_r \subset \bigcup_{\Sigma_{k+1}\in\{\Sigma_{k+1}|(16)\}} \{\boldsymbol{r}_{k+1}|\boldsymbol{r}_{k+1}^{\mathrm{T}}\Sigma_r^*\boldsymbol{r}_{k+1} \le d\} = \left\{\boldsymbol{r}_{k+1}|\boldsymbol{r}_{k+1}^{\mathrm{T}}\Sigma_r^*\boldsymbol{r}_{k+1} \le \max_{\Sigma_{k+1}\in\{\Sigma_{k+1}|(16)\}} d\right\} = \{\boldsymbol{r}_{k+1}|\boldsymbol{r}_{k+1}^{\mathrm{T}}\Sigma_r^*\boldsymbol{r}_{k+1} \le a + 2\delta\}$$

□

此外, 还需要定义如下优化问题:

优化问题 2.

$$\min_{\mathcal{P},\,\alpha_1,\,\alpha_2,\,\alpha_3} -\ln|\mathcal{P}| \tag{23}$$

$$\text{s.t.}\quad \alpha_1,\,\alpha_2,\,\alpha_3 \in (0,\,1) \tag{24}$$

$$\alpha_1 + \alpha_2 + \alpha_3 \le 1 - \alpha \tag{25}$$

$$\bar{\mathcal{Q}} = \begin{bmatrix}\alpha\mathcal{P} & \mathbf{0} & \mathcal{A}^{\mathrm{T}}\mathcal{P}\\ \mathbf{0} & \mathcal{W} & \mathcal{B}^{\mathrm{T}}\mathcal{P}\\ \mathcal{P}\mathcal{A} & \mathcal{P}\mathcal{B} & \mathcal{P}\end{bmatrix} \succeq 0 \tag{26}$$

式中

$$\mathcal{B} = [\mathcal{B}_w \quad \mathcal{B}_v \quad \mathcal{B}_r]$$

$$\mathcal{W} = \mathrm{diag}\left\{\frac{\alpha_1\Sigma_w^{-1}}{b},\,\frac{\alpha_2\Sigma_v^{-1}}{c},\,\frac{\alpha_3\Sigma_r^*}{a+2\delta}\right\}$$

常数 a、b 和 c 在可达集的定义式 (10) 中给出, $\alpha \in (0,\,1)$ 是一个预先给定的常数.

当矩阵 E、F 和 G 已知时, 基于引理 1、引理 2 和优化问题 2, 可以给出闭环系统式 (9) 的一个不变可达集.

定理 1. 考虑系统式 (9), 给定常数 $\alpha \in (0,\,1)$, 假设存在常数 α_i, $i = 1,\,2,\,3$ 和矩阵 $\mathcal{P} \succ 0$ 满足优化问题 2 中的约束式 (24) ~ (26), 那么:

$$\mathcal{E}_{\zeta_k}(\mathcal{P}) = \left\{\boldsymbol{\zeta}_k\middle|\boldsymbol{\zeta}_k^{\mathrm{T}}\mathcal{P}\boldsymbol{\zeta}_k \le 1\right\}$$

是系统式 (9) 的一个椭球形不变可达集. 假设优化问题 2 的解为 $\alpha_i = \alpha_i^*$, $i = 1,\,2,\,3$, $\mathcal{P} = \mathcal{P}^*$, 那么:

$$\mathcal{E}_{\zeta_k}(\mathcal{P}^*) = \left\{\boldsymbol{\zeta}_k\middle|\boldsymbol{\zeta}_k^{\mathrm{T}}\mathcal{P}^*\boldsymbol{\zeta}_k \le 1\right\}$$

在所有参数 $\mathcal{P}$ 满足约束式 (24) ~ (26) 的椭球 $\mathcal{E}_{\zeta_k}(\mathcal{P})$ 中体积最小.

证明. 在引理 1 中, 令 $\theta = 3$ 并选择向量 $\boldsymbol{\rho}_k$ 和 $\boldsymbol{w}_{i,\,k}$, $i = 1,\,2,\,3$ 为 $\boldsymbol{\zeta}_k$、$\boldsymbol{w}_k$、$\boldsymbol{v}_{k+1}$、$\boldsymbol{r}_{k+1}$. 此外, 选

择矩阵 Γ 为 $\mathcal{P}$, 其中 $\mathcal{P}\succ 0$ 为待设计矩阵. 根据式 (9) 和式 (12), 有:

$$\bar{\zeta}_k^{\mathrm{T}}\mathcal{Q}\bar{\zeta}_k\geq 0 \tag{27}$$

式中, $\bar{\zeta}_k^{\mathrm{T}}=[\zeta_k^{\mathrm{T}}\ \ \boldsymbol{w}_k^{\mathrm{T}}\ \ \boldsymbol{v}_{k+1}^{\mathrm{T}}\ \ \boldsymbol{r}_{k+1}^{\mathrm{T}}]$, 且:

$$\mathcal{Q}=\begin{bmatrix}\alpha\mathcal{P}-\mathcal{A}^{\mathrm{T}}\mathcal{P}\mathcal{A} & -\mathcal{A}^{\mathrm{T}}\mathcal{P}\mathcal{B}\\ -\mathcal{B}^{\mathrm{T}}\mathcal{P}\mathcal{A} & \mathcal{W}-\mathcal{B}^{\mathrm{T}}\mathcal{P}\mathcal{B}\end{bmatrix}$$

不等式 (27) 成立当且仅当 $\mathcal{Q}\succeq 0$. 根据舒尔补引理, $\mathcal{Q}\succeq 0$ 等价于线性矩阵不等式 (26). 所以当约束式 (24) ~ (26) 满足时, 椭球 $\mathcal{E}_{\zeta_k}(\mathcal{P})$ 是系统的一个不变可达集 $\mathcal{R}(\zeta_k)$.

进一步, 寻找体积最小的椭球 $\mathcal{E}_{\zeta_k}(\mathcal{P})$. 根据文献 [26] 可知, $|\mathcal{P}|^{-\frac{1}{2}}$ 和椭球 $\mathcal{E}_{\zeta_k}(\mathcal{P})$ 的体积成正比. 由于在同样的约束下, 使用目标函数 $\ln|\mathcal{P}|^{-1}$ 和 $|\mathcal{P}|^{-\frac{1}{2}}$ 的优化问题的解相同, 因此优化问题 2 可采取凸函数 $\ln|\mathcal{P}|^{-1}$ 作为目标函数[31]. □

注 3. 文献 [23, 26−27] 分别基于李雅普诺夫函数, 给出了与引理 1 类似的结论, 并用该结论构造优化问题, 以计算可达集. 需要指出的是, 文献 [23, 26−27] 求解的是可达集外椭球近似集, 而定理 1 求解的是椭球形不变可达集.

注 4. 定理 1 给出一类系统的椭球形不变可达集 $\mathcal{E}_{\zeta_k}(\mathcal{P})$. 其体积越大, 状态 ζ_k 可能取值的范围越大, 状态 $\boldsymbol{x}_k=[I_n\ \ \boldsymbol{0}\ \ \boldsymbol{0}]\zeta_k$ 越有可能超出安全区域 $\mathcal{E}_s(\Phi)$, 因此在定理 1 中, 需要寻找体积最小的 $\mathcal{E}_{\zeta_k}(\mathcal{P})$.

2.2 控制器参数设计

对于给定的控制器参数 E、F 和 G, 可以通过定理 1 计算出系统的椭球形不变可达集 $\mathcal{E}_{\zeta_k}(\mathcal{P}^*)$, 并将其投影到 $\boldsymbol{x}_k$ 超平面. 进一步, 通过分析投影是否在安全集 $\mathcal{E}_s(\Phi)$, 来判断系统的安全性.

当系统不安全时, 需要重新设计控制器参数 E、F 和 G, 此时矩阵 $\mathcal{A}$ 是变量, 因此式 (26) 不再是线性矩阵不等式, 优化问题 2 是非凸的. 本文设计一个可逆的线性变换, 构造出新的变量来替换原有的变量 $\mathcal{P}$ 和 $\mathcal{A}$, 使优化问题 2 成为对新变量的凸优化问题.

令:

$$\mathcal{P}=\begin{bmatrix}X & U & \boldsymbol{0}\\ U^{\mathrm{T}} & \tilde{X} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0} & S\end{bmatrix}$$

式中, $X\succ 0$, $\tilde{X}\succ 0$, U、$S\succ 0$ 均为 n 阶方阵.

进一步, 定义矩阵:

$$\mathcal{X}=\begin{bmatrix}X & U\\ U^{\mathrm{T}} & \tilde{X}\end{bmatrix},\quad \mathcal{X}^{-1}=\begin{bmatrix}Y & V\\ V^{\mathrm{T}} & \tilde{Y}\end{bmatrix}$$

$$\Gamma=\begin{bmatrix}Y & I_n\\ V^{\mathrm{T}} & \boldsymbol{0}\end{bmatrix}$$

容易证明 $YX+VU^{\mathrm{T}}=I_n$ 且 $YU+V\tilde{X}=\boldsymbol{0}$.

定义可逆线性变换矩阵:

$$\Pi_1=\mathrm{diag}\{\Gamma,\ I_n\}$$

$$\Pi_2=\mathrm{diag}\{\Pi_1,\ I_{n+2m},\ \Pi_1\}$$

定义如下以 Z_1、Z_2、Z_3、X、Y、S 和 α_i, $i=1,\ 2,\ 3$ 为变量的凸优化问题.

优化问题 3.

$$\begin{aligned}&\min\ \ \mathrm{Tr}(Y)\\ &\text{s.t.}\\ &\text{式(24), 式(25)}\\ &\Phi^{-1}-Y\succeq 0 &(28)\\ &\Pi_2^{\mathrm{T}}\bar{\mathcal{Q}}\Pi_2\succeq 0 &(29)\\ &\Gamma^{T}\chi\Gamma\succeq 0 &(30)\end{aligned}$$

式中

$$\Pi_2^{\mathrm{T}}\bar{\mathcal{Q}}\Pi_2=\begin{bmatrix}\alpha\Pi_1^{\mathrm{T}}\mathcal{P}\Pi_1 & \boldsymbol{0} & \Pi_1^{\mathrm{T}}\mathcal{A}^{\mathrm{T}}\mathcal{P}\Pi_1\\ \boldsymbol{0} & \mathcal{W} & \mathcal{B}^{\mathrm{T}}\mathcal{P}\Pi_1\\ \Pi_1^{\mathrm{T}}\mathcal{P}\mathcal{A}\Pi_1 & \Pi_1^{\mathrm{T}}\mathcal{P}\mathcal{B} & \Pi_1^{\mathrm{T}}\mathcal{P}\Pi_1\end{bmatrix}$$

$$\Pi_1^{\mathrm{T}}\mathcal{P}\Pi_1=\mathrm{diag}\{\Gamma^{\mathrm{T}}\chi\Gamma,\ S\}$$

$$\Gamma^{\mathrm{T}}\chi\Gamma=\begin{bmatrix}Y & I_n\\ I_n & X\end{bmatrix}$$

$$\Pi_1^{\mathrm{T}}\mathcal{P}\mathcal{B}=\begin{bmatrix}I_n-\hat{D}_1C\\ X-X\hat{D}_1C+Z_2C-Z_2\hat{D}_2C\\ S-S\hat{D}_1C\end{bmatrix}$$

$$\begin{bmatrix}-\hat{D}_1 & \hat{D}_1\\ -X\hat{D}_1CA+Z_2-Z_2\hat{D}_2 & X\hat{D}_1+Z_2\hat{D}_2\\ -S\hat{D}_1 & S\hat{D}_1-SK\end{bmatrix}$$

$$\Pi_1^{\mathrm{T}}\mathcal{P}\mathcal{A}\Pi_1=\begin{bmatrix}AY+BZ_3 & A\\ Z_1 & XA+Z_2CA\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}$$

$$\begin{bmatrix}-\hat{D}_1CA\\ -X\hat{D}_1CA-Z_2\hat{D}_2CA\\ SA-S\hat{D}_1CA\end{bmatrix}$$

$$\hat{D}_1=\tilde{D}\bar{D}_2^{\dagger}$$

$$\hat{D}_2=\bar{D}_2\bar{D}_2^{\dagger}$$

$$Z_1=XAY+UFCAY+XBGV^{\mathrm{T}}+UEV^{\mathrm{T}}+UFCBGV^{\mathrm{T}} \tag{31}$$

$$Z_2=UF \tag{32}$$

$$Z_3=GV^{\mathrm{T}} \tag{33}$$

可以得到以下定理.

定理 2. 给定参数 $0<\alpha<1$, 求解优化问题 3, 并进一步求解方程 (31) ~ (33) 和 $YX+VU^{\mathrm{T}}=I_n$, 可以获得安全控制器增益 E、F 和 G.

证明. 由于 Π_2 是非奇异变换, 因此式 (26) 等价于式 (29).

注意到:

$$X-[U\quad \mathbf{0}]\begin{bmatrix}\tilde{X} & \mathbf{0}\\ \mathbf{0} & S\end{bmatrix}^{-1}\begin{bmatrix}U^{\mathrm{T}}\\ \mathbf{0}\end{bmatrix}=Y^{-1}$$

因此, 椭球 $\mathcal{E}_{\zeta_k}(\mathcal{P})$ 在 $\boldsymbol{x}_k$ 超平面的投影可以描述为:

$$\mathcal{E}_{\boldsymbol{x}_k}(Y)=\{\boldsymbol{x}_k|\boldsymbol{x}_k^{\mathrm{T}}Y^{-1}\boldsymbol{x}_k\le 1\}$$

此外, 需要使 $\boldsymbol{x}_k$ 保持在安全集 $\mathcal{E}_s(\Phi)$ 内, 即:

$$\mathcal{E}_{\boldsymbol{x}_k}(Y)\subset\mathcal{E}_s(\Phi)$$

因此, 可得式 (28).

最后, 最小化椭球 $\mathcal{E}_{\boldsymbol{x}_k}(Y)$ 的体积 $|Y|^{\frac{1}{2}}$, 由于 $|Y|^{\frac{1}{2}}$ 不是凸函数, 因此考虑最小化 $|Y|^{\frac{1}{2}}$ 的上界. 假设 Y 的特征值为 ϕ_i, $i=1,\cdots,n$. 根据算术−几何平均不等式[32], 可得:

$$|Y|^{\frac{1}{2}}=\sqrt{\prod_{i=1}^{n}\phi_i}\le\sqrt{\left(\frac{1}{n}\sum_{i=1}^{n}\phi_i\right)^n}=\sqrt{\left(\frac{1}{n}\mathrm{Tr}(Y)\right)^n}$$

故可选用 $\mathrm{Tr}(Y)$ 作为目标函数.

根据式 (30) 和舒尔补引理, 有 $X-Y^{-1}\succ 0$, 那么有 $VU^{\mathrm{T}}=I_n-YX\prec 0$, 即矩阵 VU^{T} 是非奇异的. 因此, 如果式 (30) 成立, 总能找到满足 $YX+VU^{\mathrm{T}}=I_n$ 的非奇异矩阵 U 和 V. 那么, 在已知变量 Z_1、Z_2、Z_3、X、Y、S 的取值情况下, 利用满秩分解获得 U 和 V, 并根据式 (31) ~ (33), 可解出 E、F 和 G 的值. □

注 5. 优化问题 2 和优化问题 3 的解取决于事先给定的参数 $\alpha\in(0,1)$. 值得注意的是, 如果把 α 也作为变量, 那么优化问题 2 和优化问题 3 不再是凸的. 因此, 需要在区间 $(0,1)$ 使用网格搜索法, 以找到使目标函数最小的 α.

注 6. 由引理 1 可以看出, 当系统不受噪声影响时, 李雅普诺夫函数 $\boldsymbol{\rho}_k^{\mathrm{T}}\Gamma\boldsymbol{\rho}_k$ 满足 $\boldsymbol{\rho}_{k+1}^{\mathrm{T}}\Gamma\boldsymbol{\rho}_{k+1}-\alpha\boldsymbol{\rho}_k^{\mathrm{T}}\Gamma\boldsymbol{\rho}_k\le 0$, 因此基于引理 1 设计的控制器式 (5)、式 (6) 能够保证系统稳定.

注 7. 由定理 1 和定理 2 可以看出, 本文采用系统状态 $\boldsymbol{x}_k$、控制器状态 $\boldsymbol{z}_k$ 和滤波器的估计误差 $\boldsymbol{x}_k-\hat{\boldsymbol{x}}_k$ (不是滤波器的状态估计值 $\hat{\boldsymbol{x}}_k$) 来构造闭环系统的状态变量 $\boldsymbol{\zeta}_k$, 其原因有以下 2 点: 1) 在定理 1 计算出不变可达集 $\mathcal{E}_{\zeta_k}(\mathcal{P}^*)$ 后, 还可按照本文研究思路, 将其投影到 $\boldsymbol{x}_k-\hat{\boldsymbol{x}}_k$ 超平面, 以分析滤波器的估计性能; 2) 当采用 $\boldsymbol{x}_k$、$\boldsymbol{z}_k$ 和 $\hat{\boldsymbol{x}}_k$ 构造 $\boldsymbol{\zeta}_k$ 时, 闭环系统式 (9) 中各个矩阵的表达式会发生变化, 使用定理 2 设计的非奇异变换 Π_2 无法构造凸优化问题.

3　数值仿真

本文采用弹簧−质量−阻尼系统的例子, 来验证主要结论. 其物理模型为:

$$\begin{bmatrix}\dot{\boldsymbol{x}}_{position}\\ \dot{\boldsymbol{x}}_{velocity}\end{bmatrix}=\begin{bmatrix}0 & 1\\ -\dfrac{e}{M} & -\dfrac{d}{M}\end{bmatrix}\begin{bmatrix}\boldsymbol{x}_{position}\\ \boldsymbol{x}_{velocity}\end{bmatrix}+\begin{bmatrix}0\\ \dfrac{1}{M}\end{bmatrix}u$$

式中, $\boldsymbol{x}_{position}$ 表示位移, $\boldsymbol{x}_{velocity}$ 表示速度, $M=1\ \mathrm{kg}$ 表示质量块质量, $d=10\ (\mathrm{N\cdot s})/\mathrm{m}$ 表示阻尼器黏滞摩擦系数, $e=100\ \mathrm{N/m}$ 表示弹簧弹性系数, u 表示施加在系统上的力.

取采样间隔为 0.2 s, 并将该模型离散化, 则式 (1) 的矩阵 A 和 B 取值分别为 $\begin{bmatrix}0.150\,6 & 0.041\,9\\ -4.192\,8 & -0.268\,7\end{bmatrix}$ 和 $\begin{bmatrix}0.008\,5\\ 0.041\,9\end{bmatrix}$. 式 (1) 和式 (2) 中的其他参数取为: $C=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$, $D_1=\begin{bmatrix}1\\ 0\end{bmatrix}$, $D_2=\begin{bmatrix}0\\ 1\end{bmatrix}$, $\Sigma_w=\Sigma_v=0.005I_2$, 系统的初始状态 $\boldsymbol{x}_0$ 和状态估计值 $\hat{\boldsymbol{x}}_0$ 为 $\begin{bmatrix}0\\ 0\end{bmatrix}$.

首先, 验证定理 1 的结果. 考虑噪声和残差在对应椭球的概率分别为 95.00%、99.00%、99.75%, 即式 (10) 参数分别取 $a=b=c=5.99/9.21/11.98$. 此外, 参数 $\alpha=0.4$. 图 1 展示不同 a、b、c 值下, 参数 δ 和椭球形不变可达集体积 $|\mathcal{P}|^{-\frac{1}{2}}$ 的关系. 由图 1 可以看出, KL 散度的阈值 δ 越大, 椭球形不变可达集的体积越大, 即隐蔽攻击下系统的状态越有可能超出安全区域. 下文中均以 $\delta=0.1$ 为例进行研究. 图 2 为不同 a、b、c 值下, 参数 α 和椭球形不变可达集体积 $|\mathcal{P}|^{-\frac{1}{2}}$ 的关系. 由图 2 可以看出, 当参数 $\alpha=0.4$ 时, 椭球形不变可达集体积最小.

令 $\alpha=0.4$ 并求解优化问题 2. 图 3 为开环系统状态 (即 $\boldsymbol{u}_k=\mathbf{0}$ 时) 的椭球形不变可达集. 可以看出, a、b、c 取值越大, 椭球形不变可达集的体积越大.

然后, 验证定理 2 的结果, 图 4 为当 $a=b=c=5.99/9.21/11.98$ 时, 参数 α 和椭球形不变可达集在系统状态 $\boldsymbol{x}_k$ 超平面投影体积上界 $\mathrm{Tr}(Y)$ 的关系. 由图 4 可知, 当 $\alpha=0.7$ 时, $\mathrm{Tr}(Y)$ 最小.

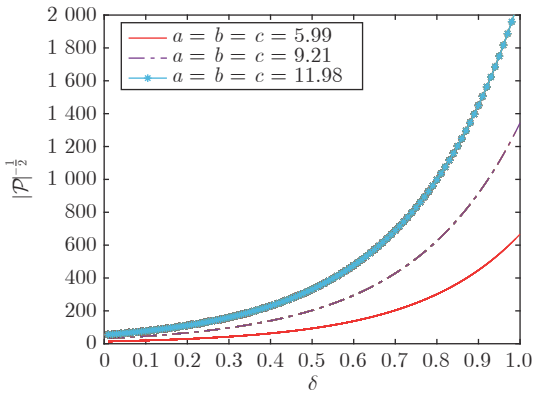

图 1 不同 a、b、c 值下, 参数 δ 和椭球形不变可达集体积的关系

Fig. 1 Relationship between δ and volume of ellipsoidal invariant reachable set with different a, b and c

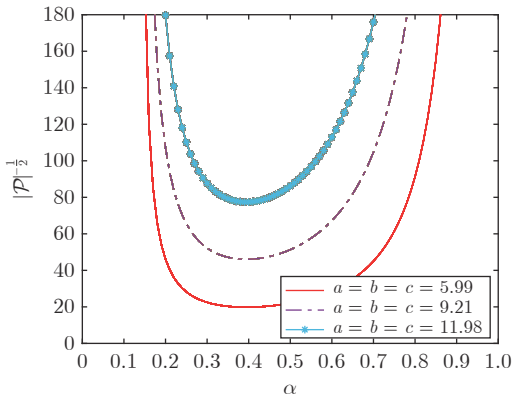

图 2 不同 a、b、c 值下, 参数 α 和椭球形不变可达集体积的关系

Fig. 2 Relationship between α and volume of ellipsoidal invariant reachable set with different a, b and c

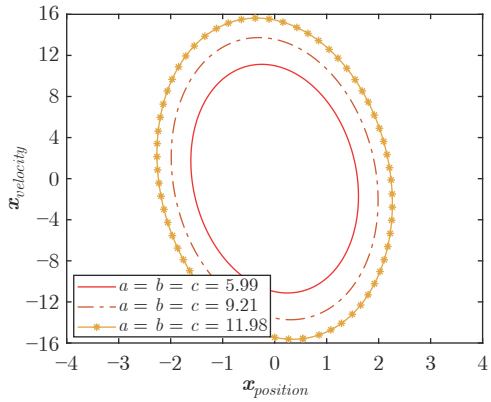

图 3 不同 a、b、c 值下, 系统状态的椭球形不变可达集

Fig. 3 Ellipsoidal invariant reachable sets of system′s state with different a, b and c

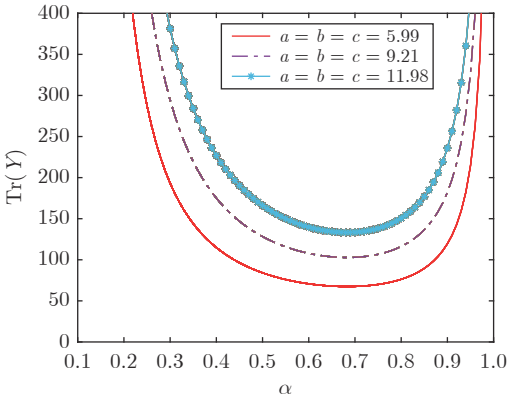

图 4 不同 a、b、c 值下, 参数 α 和 $\mathrm{Tr}(Y)$ 的关系

Fig. 4 Relationship between α and $\mathrm{Tr}(Y)$ with different a, b and c

最后, 以 $a = b = c = 9.21$ 为例验证控制器式 (5) 和式 (6) 的效果. 假设弹簧的速度不能超过 10 m/s. 因此, 对于安全集 $\mathcal{E}_s(\Phi)$, 矩阵 Φ 可以取 $\begin{bmatrix} 1/10^2 & 0 \\ 0 & 1/10^2 \end{bmatrix}$. 未使用控制器和使用控制器时的参数 α 分别取 0.4 和 0.7, 并求解优化问题 3 获得控制器式 (5) 和式 (6) 的参数. 由图 5 可以看出, 开环系统的可达集与危险状态相交, 而使用控制器式 (5) 和式 (6) 后, 弹簧的最大速度显著降低, 此时能保证系统状态在安全集范围内.

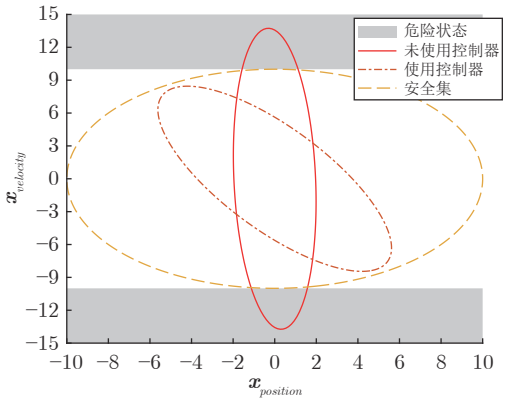

图 5 未使用控制器和使用控制器式 (5)、(6) 时, 系统状态的椭球形不变可达集

Fig. 5 Ellipsoidal invariant reachable set of system′s state without controller and with controller (5), (6)

4 结束语

本文对隐蔽攻击下信息物理系统的安全控制进行了研究. 采用 KL 散度作为攻击的隐蔽性指标, 并基于系统的随机特性定义可达集. 首先, 给出隐蔽攻击下检测器残差的一个外椭球近似集. 其次,

基于该近似集和噪声的范围构造出一个优化问题, 以描述控制器参数与系统椭球形不变可达集之间的关系. 进一步设计一种可逆线性变换, 将该优化问题转化成凸优化问题, 以求解控制器参数和相应的不变可达集. 最后, 以弹簧–质量–阻尼系统为例, 验证本文提出控制算法的有效性. 仿真结果表明, 本文设计的控制器能够显著降低弹簧的最大速度, 保证系统的状态在安全范围内. 未来可将本文的工作扩展到更为复杂的系统, 如非线性系统和云控制系统[33] 等.

References

1 Derler P, Lee E A, Vincentelli A S. Modeling cyber-physical systems. *Proceedings of the IEEE*, 2012, **100**(1): 13−28

2 Li Hong-Yang, Wei Mu-Heng, Huang Jie, Qiu Bo-Hua, Zhao Ye, Luo Wen-Cheng, et al. Review of information physical systems technology. *Acta Automatica Sinica*, 2019, **45**(1): 37−50
(李洪阳, 魏慕恒, 黄洁, 邱伯华, 赵晔, 骆文城, 等. 信息物理系统技术综述. 自动化学报, 2019, **45**(1): 37−50)

3 Kim S, Park K J, Lu C Y. A survey on network security for cyber-physical systems: From threats to resilient design. *IEEE Communications Surveys & Tutorials*, 2022, **24**(3): 1534−1573

4 Liu Ting, Tian Jue, Wang Jia-Zhou, Wu Hong-Yu, Sun Li-Min, Zhou Ya-Dong, et al. Research on integrated security threat and defense of information physical fusion system. *Acta Automatica Sinica*, 2019, **45**(1): 5−24
(刘烃, 田决, 王稼舟, 吴宏宇, 孙利民, 周亚东, 等. 信息物理融合系统综合安全威胁与防御研究. 自动化学报, 2019, **45**(1): 5−24)

5 Duo W L, Zhou M C, Abusorrah A, Valente J, Faisal M, Ruths J, et al. A survey of cyber attacks on cyber physical systems: Recent advances and challenges. *IEEE/CAA Journal of Automatica Sinica*, 2022, **9**(5): 784−800

6 Giraldo J, Urbina D, Cardenas A. A survey of physics-based attack detection in cyber-physical systems. *ACM Computing Surveys*, 2018, **51**(4): Article No. 76

7 Tan S, Guerrero J M, Xie P L, Han R K, Vasquez J C. Brief survey on attack detection methods for cyber-physical systems. *IEEE Systems Journal*, 2020, **14**(4): 5329−5339

8 Yang Chao-Qun, Zhang Heng. Multi-attack detection approach based on CBMeMBer filt. *Control Engineering of China*, 2022, **29**(6): 1033−1039
(杨超群, 张恒. 基于 CBMeMBer 滤波器的多攻击检测方法. 控制工程, 2022, **29**(6): 1033−1039)

9 Mu Chao, Wang Xin, Yang Ming, Zhang Heng, Chen Zhen-Ya, Wu Xiao-Ming. Hard-coded vulnerability detection approach for IoT device firmware. *Chinese Journal of Network and Information*, 2022, **8**(5): 98−110
(穆超, 王鑫, 杨明, 张恒, 陈振娅, 吴晓明. 面向物联网设备固件的硬编码漏洞检测方法. 网络与信息安全学报, 2022, **8**(5): 98−110)

10 Zhang Q R, Liu K, Teixeira A, Li Y Z, Chai S C, Xia Y Q. An online Kullback-Leibler divergence-based stealthy attack against cyber-physical systems. *IEEE Transactions on Automatic Control*, 2022, **68**(6): 3672−3679

11 Zhang Q R, Liu K, Xia Y Q, Ma A Y. Optimal stealthy deception attack against cyber-physical systems. *IEEE Transactions on Cybernetics*, 2020, **50**(9): 3963−3972

12 Zhang Q R, Liu K, Han D Y, Su G Z, Xia Y Q. Design of stealthy deception attacks with partial system knowledge. *IEEE Transactions on Automatic Control*, 2023, **68**(2): 1069−1076

13 Guo Z Y, Shi D W, Johansson K H, Shi L. Worst-case stealthy innovation-based linear attack on remote state estimation. *Automatica*, 2018, **89**: 117−124

14 Mo Y L, Sinopoli B. False data injection attacks in control systems. In: Proceedings of the 1st Workshop on Secure Control Systems. Stockholm, Sweden: IEEE, 2010. 1−6

15 Kwon C, Liu W Y, Hwang I. Analysis and design of stealthy cyber attacks on unmanned aerial systems. *Journal of Aerospace Information Systems*, 2014, **11**(8): 525−539

16 Sui T J, Mo Y L, Marelli D, Sun X M, Fu M Y. The vulnerability of cyber-physical system under stealthy attacks. *IEEE Transactions on Automatic Control*, 2020, **66**(2): 637−650

17 Sui T J, Sun X M. The vulnerability of distributed state estimator under stealthy attacks. *Automatica*, 2021, **133**: Article No. 109869

18 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

19 Xu W Y, Wang Z D, Hu L, Kurths J. State estimation under joint false data injection attacks: Dealing with constraints and insecurity. *IEEE Transactions on Automatic Control*, 2022, **62**(2): 6745−6753

20 Jovanov I, Pajic M. Relaxing integrity requirements for attack-resilient cyber-physical systems. *IEEE Transactions on Automatic Control*, 2019, **64**(12): 4843−4858

21 Zhang T Y, Ye D. False data injection attacks with complete stealthiness in cyber-physical systems: A self-generated approach. *Automatica*, 2020, **120**: Article No. 109117

22 Mo Y L, Sinopoli B. On the performance degradation of cyber-physical systems under stealthy integrity attacks. *IEEE Transactions on Automatic Control*, 2015, **61**(9): 2618−2624

23 Murguia C, Shames I, Ruths J. Security metrics and synthesis of secure control systems. *Automatica*, 2020, **115**: Article No. 108757

24 Kwon C, Hwang I. Reachability analysis for safety assurance of cyber-physical systems against cyber attacks. *IEEE Transactions on Automatic Control*, 2018, **63**(7): 2272−2279

25 Liu H, Niu B, Qin J H. Reachability analysis for linear discrete-time systems under stealthy cyber attacks. *IEEE Transactions on Automatic Control*, 2021, **66**(9): 4444−4451

26 Zhang Q R, Liu K, Pang Z H, Xia Y Q, Liu T. Reachability analysis of cyber-physical systems under stealthy attacks. *IEEE Transactions on Cybernetics*, 2022, **52**(6): 4926−4934

27 Hashemi N, Ruths J. Co-design for resilience and performance. *IEEE Transactions on Control of Network Systems*, 2022, **10**(3): 1387−1399

28 Bai C Z, Pasqualetti F, Gupta V. Data-injection attacks in stochastic control systems: Detectability and performance tradeoffs. *Automatica*, 2017, **82**: 251−260

29 Fang C R, Qi Y F, Chen J M, Tan R, Zheng W X. Stealthy actuator signal attacks in stochastic control systems: Performance and limitations. *IEEE Transactions on Automatic Control*, 2020, **65**(9): 3927−3934

30 Kullback S. *Information Theory and Statistics*. Chelmsford: Courier Corporation, 1997.

31 Boyd S, El Ghaoui L, Feron E, Balakrishnan V. *Linear Matrix Inequalities in System and Control Theory*. Philadelphia: Society for Industry and Applied Mathematics, 1994.

32 Diananda P H. A simple proof of the arithmetic mean geometric mean inequality. *The American Mathematical Monthly*, 1960, **67**(10): Article No. 1007

33 Xia Yuan-Qing. Cloud control systems and their challenges. *Acta Automatica Sinica*, 2016, **42**(1): 1−12
(夏元清. 云控制系统及其面临的挑战. 自动化学报, 2016, **42**(1): 1−12)

张淇瑞 中国矿业大学信息与控制工程学院副教授. 主要研究方向为信息物理系统的脆弱性分析, 安全控制和隐私保护.
E-mail: qiruizhang@cumt.edu.cn
(**ZHANG Qi-Rui** Associate professor at the School of Information and Control Engineering, China University of Mining and Technology. His research interest covers vulnerability analysis of cyber-physical systems, secure control, and privacy protection.)

孟思琪 中国矿业大学信息与控制工程学院硕士研究生. 主要研究方向为信息物理系统的安全控制.
E-mail: siqimeng@cumt.edu.cn
(**MENG Si-Qi** Master student at the School of Information and Control Engineering, China University of Mining and Technology. Her main research interest is secure control of cyber-physical systems.)

王兰豪 中国矿业大学国家煤加工与洁净化工程技术研究中心副教授. 主要研究方向为复杂工业过程的工艺参数检测、优化决策与智能控制. 本文通信作者.
E-mail: wanglanhao888@163.com
(**WANG Lan-Hao** Associate professor at National Engineering Research Center of Coal Preparation and Purification, China University of Mining and Technology. His research interest covers process parameter detection, optimal decision making, and intelligent control of complex industrial process. Corresponding author of this paper.)

刘 坤 北京理工大学自动化学院研究员. 主要研究方向为网络化控制理论与应用, 复杂网络系统安全.
E-mail: kunliubit@bit.edu.cn
(**LIU Kun** Professor at the School of Automation, Beijing Institute of Technology. His research interest covers theory and applications of networked control, and security of complex networked systems.)

代 伟 中国矿业大学信息与控制工程学院教授. 主要研究方向为复杂工业过程建模、运行优化与控制.
E-mail: weidai@cumt.edu.cn
(**DAI Wei** Professor at the School of Information and Control Engineering, China University of Mining and Technology. His research interest covers modeling, operational optimization and control for complex industrial process.)