控制与决策
Control and Decision

基于伪周期控制信号编码的重放攻击检测

张正道, 王瑶瑶, 谢林柏

引用本文:
张正道,王瑶瑶,谢林柏. 基于伪周期控制信号编码的重放攻击检测[J]. 控制与决策, 2023, 38(10): 2962−2968.

在线阅读 View online: https://doi.org/10.13195/j.kzyjc.2021.1967

基于伪周期控制信号编码的重放攻击检测

张正道[1,2†], 王瑶瑶[1], 谢林柏[1,2]

(1. 江南大学 物联网工程学院，江苏 无锡 214122；
2. 江南大学 物联网应用技术教育部工程研究中心，江苏 无锡 214122)

摘 要: 现有基于控制信号编码的信息物理系统重放攻击检测方法主要是通过判断检测函数值的大小是否超出给定阈值来实现检测,这类方法普遍存在检测率与系统控制性能损失之间的矛盾. 鉴于此,提出一种基于伪周期控制信号编码的检测方法. 首先,在控制信号中加入预先设计的伪周期随机编码信号,并构造与之对应的伪周期测量值补偿信号,验证当系统矩阵稳定时,补偿信号的周期性;然后,对接收到的测量值信号,利用不同补偿信号进行补偿,获得检测函数最小时对应的补偿信号在周期中的位置;最后,通过比较该补偿信号与实际控制量水印信号在周期中的位置实现对重放攻击的检测. 仿真实验结果表明,采用所提出方法,只需比较采用不同补偿信号下检测函数值的相对大小,便能够在有效地检测重放攻击的同时,降低控制编码信号的方差并减小系统控制性能损失.

关键词: 信息物理系统；重放攻击；攻击检测；控制信号编码；伪周期性

中图分类号: TP273 **文献标志码:** A

DOI: 10.13195/j.kzyjc.2021.1967

引用格式: 张正道,王瑶瑶,谢林柏. 基于伪周期控制信号编码的重放攻击检测[J]. 控制与决策, 2023, 38(10): 2962-2968.

Replay attack detection method based on pseudo periodic control signal coding

ZHANG Zheng-dao[1,2†], *WANG Yao-yao*[1], *XIE Lin-bo*[1,2]

(1. College of Internet of Things Engineering，Jiangnan University，Wuxi 214122，China；2. Engineering Research Center of Things Applied Technology，Ministry of Education，Jiangnan University，Wuxi 214122，China)

Abstract: The existing replay attack detection methods of cyber-physical system based on control signal coding mainly realize the detection by judging whether the value of the detection function exceeds a predetermined threshold, which results in a conflict between the detection rate and the loss of system control performance. Therefore, this paper proposes a detection method based on pseudo periodic control signal coding. Firstly, the pre-designed pseudo periodic random coding signal is added to the control signal, and the corresponding pseudo periodic measurement compensation signal is constructed. The periodicity of the compensation signal is proved when the system matrix is stable. Then, the received measurement is compensated by different compensation signals to obtain the position of the compensation signal corresponding to the minimum value of the detection function in the period. By comparing the position of the compensation signal and the actual watermark signal in the time, the replay attack is detected. The simulation results show that the proposed method only needs to compare the relative values of the detection function under different compensation signals, so as to effectively detect the replay attack, decrease the variance of the control coding signal and reduce the performance loss of the system control.

Keywords: cyber-physical system；replay attack；attack detection；control signal coding；pseudo periodicity

0 引 言

信息物理系统(cyber-physical system, CPS)是通过融合传感器、通信装置和计算单元等设备,实现物理设施与信息过程交互的系统,已广泛应用于智能制造、交通运输等关键领域[1]. 同时,CPS开放和易访问的特性使其更易受恶意实体的攻击[2]. 近年来,频发的CPS攻击事件造成了巨大的生命和财产损失,引起了学术界对CPS攻击检测的广泛关注.

针对CPS攻击类型主要有拒绝服务攻击、虚假数据注入攻击和重放攻击等[3]. 重放攻击通过窃听并

收稿日期: 2021-11-12；**录用日期:** 2022-05-23.
基金项目: 国家重点研发计划课题项目(2018YFB1701903).
责任编委: 左志强.
†通讯作者. E-mail: wxzzd@jiangnan.edu.cn.

记录系统历史数据,并将其回放至通信链路上替换当前时刻的实际数据,进而诱导控制器发出错误的控制信号,破坏系统正常运行. 与网络延迟现象中数据滞后到达控制器但保持值不变不同,重放攻击改变了当前的真实数据,诱导控制器作出错误响应,故处理延迟现象的方案不适用于重放攻击的检测. 重放攻击不要求攻击者具有被攻击系统的先验知识,且重放的是系统正常运行情况下的数据,攻击易于实施且具有较强的隐匿性,因此重放攻击的检测是一个值得研究的问题.

现有的重放攻击检测方法主要有时间戳技术、测量信号编码和向控制信号添加物理水印等. 时间戳方法需保证信息发送端与接收端的时钟保持严格同步,否则会出现误判,且攻击者可能会拦截并修改添加的时间戳[4]. 文献[5]利用已知均值和方差的高斯噪声信号对测量值进行编码,并在接收端进行相应的解码. 此方案要求在传感器端和控制器端产生2个相同的高斯噪声序列,有一定的实现难度. 文献[6]将控制量进行相关变换然后作为编码信号注入测量值,通过变换矩阵实现在重放攻击下破坏闭环系统的稳定性. 但是,在控制器端的网络传输出现丢包或延迟时,该方法会造成误报. 将CPS建模为离散线性时不变系统,文献[7]提出向最优控制信号中添加额外的水印信号,使得重放攻击前后残差信号出现明显变化. 此方法会牺牲一定的系统控制性能,且水印信号协方差的大小与性能损失满足正相关关系[8]. 文献[9]采用基于数据的自适应算法,在存在干扰、随机噪声和重放攻击的情况下设计最佳博弈策略以获得最优的测量值水印信号协方差,实现性能损失与检测性能间的平衡. 文献[10]进一步提出向控制信号中周期性地加入噪声信号,并通过马尔可夫方法选取合适的周期长度使得检测率与性能损失达到平衡. 在系统参数需要在线识别的前提下,文献[11]提出了一种基于数据驱动的最优物理水印设计方法. 文献[12]考虑网络传输过程的丢包情况,对控制信号注入物理水印使得重放攻击前后残差数据分布不同,应用K-L散度实现重放攻击检测. 文献[13]设计了一种乘法水印方案,其在传感器侧添加水印生成器并在控制器侧加入均衡滤波器,根据预先设定的加密数据不断切换滤波器参数,此方案不损失控制性能,但提高了系统操作的复杂性. 文献[14]提出在控制信号中注入水印信号的同时在测量值中添加补偿信号,使得性能损失仅与当前水印信号值有关,并给出水印信号协方差、检测率与检测阈值间的最优化问题. 针对不连续重放攻击,文献[15]提出了周期性地向控制信号添加水印信号方法,并在性能损失约束下得到近似检测性能的最优水印调度策略.

从现有研究中可发现,水印信号的大小决定了重放攻击的检测率,同时决定了系统的控制性能损失. 其原因主要为,现有检测方法均需要在攻击发生后检测函数的绝对大小大于系统正常工作时的最大检测函数值. 为此,本文提出一种伪周期水印方法,通过比较对应周期内不同时刻补偿信号下检测函数的相对大小判断有无攻击发生. 首先,向控制量中注入伪周期随机水印信号,并构造相应的补偿信号,验证了当状态空间方程中状态转移矩阵稳定时,补偿信号具有周期性;然后,对接收到的测量值信号,利用补偿信号进行补偿,获得检测函数最小时对应的补偿信号;最后,通过比较该补偿信号与实际水印信号在周期中的位置实现对重放攻击的检测.

1　系统模型和重放攻击描述

1.1　系统模型

将CPS系统表示为如下线性时不变系统:

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + \omega_k,\\ y_k &= Cx_k + \nu_k. \end{aligned} \tag{1}$$

其中: $x_k \in R^{n_x}$ 为状态向量; $u_k \in R^{n_u}$ 为控制输入; $y_k \in R^{n_y}$ 为输出向量; A、B、C 为具有相应维度的已知系统矩阵;过程噪声 $\omega_k \sim N(0, W)$ 与测量噪声 $\nu_k \sim N(0, V)$ 相互独立,系统初始状态为0. 假设 (A, B) 和 $(A, W^{1/2})$ 完全可控, (A, C) 完全可观. 应用如下Kalman滤波器得到系统的状态估计 $\widehat{x}_{k|k}$:

$$\begin{aligned} &\widehat{x}_{k+1|k} = A\widehat{x}_{k|k} + Bu_k,\ P_{k+1|k} = AP_kA^{\mathrm{T}} + W,\\ &L_k = P_{k|k-1}C^{\mathrm{T}}(CP_{k|k-1}C^{\mathrm{T}} + V)^{-1},\\ &\widehat{x}_{k|k} = \widehat{x}_{k|k-1} + L_k(y_k - C\widehat{x}_{k|k-1}),\\ &P_k = P_{k|k-1} - L_kCP_{k|k-1}. \end{aligned} \tag{2}$$

其中Kalman滤波增益 L_k、预测误差协方差矩阵 P_k 经过动态迭代将分别收敛至常值 L 和 P,即

$$\begin{aligned} L &\triangleq PC^{\mathrm{T}}(CPC^{\mathrm{T}} + V)^{-1},\\ P &\triangleq \lim_{k\to\infty} P_{k|k-1}. \end{aligned} \tag{3}$$

设计LQG控制器以确保系统稳定,定义控制性能指标函数为

$$\Lambda = \min \lim_{k\to\infty} E({x_k}^{\mathrm{T}}Qx_k + {u_k}^{\mathrm{T}}Ru_k). \tag{4}$$

其中: Q、R 为相应变量的权重矩阵,且 $Q \geqslant 0, R \geqslant 0$.

在满足性能指标最小的情况下,得到最优控制信号 u_k^* 以及最优性能指标 Λ^* 分别为

$$u_k^* = -(B^{\mathrm{T}}SB+R)^{-1}B^{\mathrm{T}}SA\hat{x}_{k|k} = F\hat{x}_{k|k},$$
$$\Lambda^* = \mathrm{tr}(SW) + \mathrm{tr}[(A^{\mathrm{T}}SA+Q-S)(P-LCP)]. \tag{5}$$

其中:最优反馈矩阵$F=-(B^{\mathrm{T}}SB+R)^{-1}B^{\mathrm{T}}SA$,$S$为方程$S=A^{\mathrm{T}}SA-A^{\mathrm{T}}SB(B^{\mathrm{T}}SB+R)^{-1}B^{\mathrm{T}}SA+Q$的解.

1.2 重放攻击模型

本文考虑测量信号被重放的情形,且假设:1)攻击者可窃听、修改传感器传输的真实测量数据;2)攻击者可记录足够多的历史测量数据,记为y_{k_1};3)攻击者可在任意时刻k用记录的历史测量数据替换当前时刻的真实测量信号并传输给控制器,即$y_k=y_{k_1}$. 记重放数据延迟时刻为$\Delta k=k-k_1$.

文献[7]指出,若系统矩阵$\Gamma=(A+BF)(I_{n_x}-LC)$是稳定的,则重放攻击可以绕过$\chi^2$检测. 文献[8]分析了当系统矩阵$A$稳定时,重放攻击下闭环系统仍然是稳定的. 故本文仅研究系统矩阵Γ和A稳定情况下,隐匿重放攻击的检测方案.

2 重放攻击检测

与文献[15]方法不同,所提出方法向控制信号中持续注入周期性伪随机信号作为控制量水印. 通过构造辅助系统,得到周期性的补偿信号序列$Y^s=[Y_0^s,Y_1^s,\dots,Y_{T-1}^s]$,$Y_i^s(i\in[0,T-1])$为补偿信号,$T$为水印信号的周期. 对接收到的测量值$y_k^*$分别加入不同补偿信号,计算各补偿信号下检测函数的值并获取检测函数最小值对应的补偿信号在周期中的位置. 若该补偿信号在周期中的位置与实际控制量水印在周期中的位置不一致,则表明系统发生了重放攻击. 所提出方法的检测结构如图1所示.

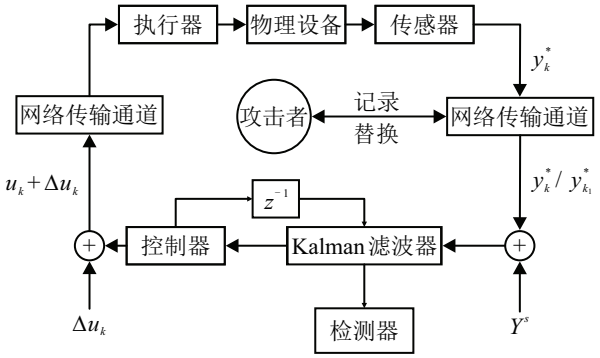


图1 本文检测方法的结构

2.1 周期水印信号构建

令水印信号的周期为正整数T,记周期水印信号的集合为$\Delta u=[\Delta u_0,\Delta u_1,\dots,\Delta u_{T-1}]$,其中$\Delta u_i$ $(i=0,1,\dots,T-1)$独立抽取自均值为0、协方差矩阵为Σ_u的高斯白噪声信号.

设系统在t_0时刻已稳定运行,并从该时刻开始向控制量中持续注入信号,则在k时刻,注入的控制量水印信号Δu_k可表示为

$$\Delta u_k = \Delta u_i,\ i=(k-t_0)\bmod T, \tag{6}$$

其中$(k-t_0)\bmod T$为$(k-t_0)$除以T的余数,且$0\leqslant(k-t_0)\bmod T\leqslant T-1$.

注入周期水印信号后,系统状态可表示为

$$x_{k+1}^* = Ax_k^* + B(u_k^*+\Delta u_k)+\omega_k,$$
$$y_k^* = Cx_k^* + v_k. \tag{7}$$

2.2 周期补偿信号构建

构造如下辅助系统以获得加入水印后$(k>t_0)$的测量值补偿信号y_k^s:

$$x_{k+1}^s = Ax_k^s + B\Delta u_k,$$
$$y_k^s = -Cx_k^s, \tag{8}$$

其中辅助状态x_k^s的初始值为0,即$x_{t_0}^s=0$.

定理1 当矩阵A稳定时,对于周期为T的水印信号,存在正整数n_0使得$A^{(n_0-1)T}\doteq 0$,则在加入n_0个周期的水印信号后,$k\geqslant t_0+n_0T$,补偿信号与水印信号具有相同的周期,且在同一周期内补偿信号与测量值中水印信号相互对应,即$\Delta u_{i-1}\leftrightarrow Y_i^s, i=(k-t_0)\bmod T$.

证明 矩阵C为已知的常值矩阵,故若要证明补偿信号y_k^s在$k\geqslant t_0+n_0T$后为周期的,则只需证明补偿状态x_k^s在$k\geqslant t_0+n_0T$后为周期的即可.

当$k\in[t_0,t_0+T-1]$时,补偿信号可表示为

$$x_k^s = Ax_{k-1}^s + B\Delta u_{k-1} = \sum_{i=0}^{k-t_0} A^iB\Delta u_{k-t_0-i}. \tag{9}$$

当$k=t_0+nT+1, i\in[0,T-1]$时,补偿状态信号经迭代可表示为

$$x_k^s = x_{t_0+nT+i}^s = Ax_{t_0+nT+i-1}^s + B\Delta u_{t_0+nT+i-1} =$$
$$Ax_{t_0+nT+i-1}^s + B\Delta u_{i-1} =$$
$$A^{(n-1)T+i}x_{t_0+T}^s + x_{t_0+(n-1)T+i}^s. \tag{10}$$

矩阵A为稳定的,因此存在正整数n_0使得$A^{(n_0-1)T}\doteq 0$,n_0的取值仅与矩阵A有关. 则当$n\geqslant n_0$时,有$x_{nT+t_0+i}^s=x_{(n-1)T+t_0+i}^s$,即

$$x_k^s = x_{k-T}^s. \tag{11}$$

由式(11)可知,在运行n_0周期后,补偿信号的状态值为周期的,且同一周期内水印信号相互独立,故同一周期内补偿信号状态值互不相同. 补偿信号$y_k^s=-Cx_k^s$,而矩阵$-C$为常值矩阵,故补偿信号y_k^s为周期的且在同一周期内互不相同.

由式(8)可知,在$k=t_0+n_0T+i(i\in[0,T-1])$时,补偿信号值$Y_i^s$与此时测量值中水印信号值$\Delta u_{k-1}=$

$\Delta u_{i-1}, i=(k-t_0) \bmod T$,具有对应关系. 水印信号序列是从同一个正态分布中独立抽取的,因此当$i\neq j$时,有$P(\Delta u_i=\Delta u_j)=0$. 补偿信号值$Y_i^s$与水印信号序列$\Delta u_{i-1}$一一对应. □

2.3　检测可行性

在k时刻,对接收到的测量值y_k^*加入补偿信号$Y_m^s, m\in[0,T-1]$,则补偿后的测量值为$y_{k,m}^e=y_k^*+Y_m^s$. 令此时Kalman滤波器得到的预测状态估计为$\hat{x}_{k|k-1}^e$,估计的残差为$r_{k,m}^e=y_{k,m}^e-C\hat{x}_{k|k-1}^e$. 基于残差的卡方检测函数为$g_{k,m}=(r_{k,m}^e)^{\mathrm{T}}\Sigma_r^{-1}r_{k,m}^e$,其中$\Sigma_r=CPC^{\mathrm{T}}+V$.

引理1[7]　对于配备Kalman滤波器和LQG控制器的线性时不变系统,如式(1)所示,其残差$r_k=y_k-C\hat{x}_{k|k-1}$服从均值为0、协方差为$\Sigma_r=CPC^{\mathrm{T}}+V$的高斯分布.

定理2　在k时刻,$k=t_0+nT+i(n\geqslant n_0, i\in[0,T-1])$,且此时控制量水印信号为$\Delta u_{i-1}$,则注入不同补偿信号$Y_m^s(m\in[0,T-1])$后,卡方检测函数满足$g_{k,i}<g_{k,j}, j\in[0,T-1]$且$j\neq i$. 即注入补偿信号$Y_m^s$满足$m=i$时,检测函数值$g_{k,m}$最小.

证明　当$m=i$时,由定理1可知此时补偿信号为$Y_m^s=Y_i^s$,则加入补偿信号后的测量值$y_{k,m}^e$为

$$\begin{aligned}&y_{k,m}^e=y_k^*+Y_m^s=y_k^*+Y_{t_0+nT+i}^s=\\&Cx_k^*+\nu_k-Cx_k^s=\\&C\Big[A^kx_0+\sum_{i=0}^{k-1}A^i(Bu_{k-1-i}+\omega_{k-1-i})\Big]+\nu_k=y_k.\end{aligned}\tag{12}$$

此时,加入补偿信号Y_m^s后的测量值$y_{k,m}^e$与原始系统测量值y_k相同,补偿信号可抵消对应水印信号对测量值造成的影响.

记补偿后Kalman滤波器的预测状态估计为$\hat{x}_{k|k-1}^e$,原始系统状态预测估计值为$\hat{x}_{k|k-1}$. 在0初始时刻,有$\hat{x}_{0|-1}^e=\hat{x}_{0|-1}$. 由式(12),得到

$$\begin{aligned}&\hat{x}_{k|k-1}^e=A\hat{x}_{k-1|k-1}^e+Bu_{k-1}=\\&(A+BF)\hat{x}_{k-1|k-1}^e=\\&\Gamma\hat{x}_{k-1|k-2}^e+(A+BF)Ly_{k-1,m-1}^e=\hat{x}_{k|k-1},\end{aligned}\tag{13}$$

即在$m=i$的条件下,补偿后状态预测估计值$\hat{x}_{k|k-1}^e$与未注入水印的$\hat{x}_{k|k-1}$相同. 因此,残差$r_{k,m}^e$满足

$$r_{k,m}^e=y_{k,m}^e-C\hat{x}_{k|k-1}^e=y_k-C\hat{x}_{k|k-1}=r_k.\tag{14}$$

根据引理1,$r_{k,m}^e\sim N(0,\Sigma_r)$,故补偿后基于残差的检测函数值为

$$g_{k,m}=(r_{k,m}^e)^{\mathrm{T}}{\Sigma_r}^{-1}r_{k,m}^e={r_k}^{\mathrm{T}}{\Sigma_r}^{-1}r_k=g_k,$$

即补偿后基于残差的卡方检测函数值与未注入控制量水印时的检测函数值g_k相同.

若$k-1$时注入水印信号$\Delta u_{k-1}=\Delta u_{i-1}$,而在$k$时刻,注入的补偿为$Y_j^s$,且有$j\neq i$,则Kalman滤波器端接收到的测量值为

$$y_{k,j}^e=y_k^*+Y_j^s=y_k+Y_j^s-Y_i^s.\tag{15}$$

此时残差$r_{k,j}^e$为

$$r_{k,j}^e=y_{k,j}^e-C\hat{x}_{k|k-1}^e=r_k+Y_j^s-Y_i^s.\tag{16}$$

基于残差的卡方检测信号为

$$\begin{aligned}&g_{k,j}=(r_{k,j}^e)^{\mathrm{T}}{\Sigma_r}^{-1}r_{k,j}^e=\\&(r_k+Y_j^s-Y_i^s)^{\mathrm{T}}{\Sigma_r}^{-1}(r_k+Y_j^s-Y_i^s)=\\&g_k+|\Sigma_r|^{-1}\|Y_j^s-Y_i^s\|_2^2,\end{aligned}\tag{17}$$

$j\neq i$,且补偿信号在同一周期内为互不相等的,故$g_{k,j}>g_k=g_{k,i}$. □

由定理2,设计如下重放攻击检测方法. 从t_0时刻开始,持续向控制量注入水印信号,在$k-1$时刻,该水印信号为$\Delta u_{k-1}=\Delta u_{i-1}, i=(k-t_0) \bmod T$. 在$k$时刻,$k=t_0+nT+i, n\geqslant n_0, i\in[0,T-1]$,记加入不同补偿信号后卡方检测函数的最小值满足

$$a_k=\min_m\ (g_{k,m}),\ m=0,1,\ldots,T-1.\tag{18}$$

定义判别函数为

$$J_k=a_k-i.\tag{19}$$

若$J_k=0$,则认为系统正常运行;若$J_k\neq0$,则认为系统受到重放攻击.

定理3　对于式(1)所示的线性定常系统,当重放延迟时刻不为整数倍的水印信号周期时,采用所提出方法可实现对重放攻击的检测.

证明　当系统正常运行时,由定理2可知,在k时刻,$k=t_0+nT+i$,使得残差检测函数值最小的补偿信号为Y_m^s,且有$m=i$,由式(18)有$a_k=m$,则所提出方法判别函数为$J_k=a_k-i=m-i=0$.

若$k=t_0+nT+i$时刻发生重放攻击,则其重放的数据为k_1时刻的正常数据$y_{k_1}^*, k_1=t_0+n_1T+m_1$. 重放的测量信号对应的控制量水印为$\Delta u_{k_1-1}=\Delta u_{m_1-1}, m_1=(k_1-t_0) \bmod T$. 系统在$k-1$时刻正常,即$y_{k-1}^e=y_{k-1}$,则有$\hat{x}_{k|k-1}^e=\hat{x}_{k|k-1}$. 由定理2可得,为使得检测函数取得小值,则注入的补偿信号为$y_{k_1}^s=Y_{m_1}^s$,此时有$a_k=m_1$. 当$i\neq m_1$时,重放数据延迟时刻$\Delta k=k-k_1=(n-n_1)T+i-m_1$不是$T$的整数倍,此时判别函数$J_k=a_k-i=m_1-i\neq0$,因此可认为此时发生重放攻击. □

注1　为实现破坏系统稳定运行的目的,重放攻击者在展开攻击前需要记录并存储足够时间长度的

系统信号,而控制水印可在系统开始运行时便加入控制量中,且n_0只需使得$A^{(n_0-1)T}$近似为0即可,故本文假设在n_0周期内不发生重放攻击.

2.4 性能损失

构造增广矩阵$\tilde{x}_k=[x_k \quad e_k \quad \zeta_k]^{\mathrm{T}}$,其中$e_k\triangleq x_k-\hat{x}_{k|k}$,$\zeta_k\triangleq x_k^*-x_k$,则有

$$\tilde{x}_{k+1}=\begin{bmatrix}A+BF & -BF & 0\\ 0 & A-LCA & 0\\ 0 & 0 & A\end{bmatrix}\tilde{x}_k+\begin{bmatrix}\omega_k\\ (I_n-LC)w_k-Lv_{k+1}\\ B\Delta u_k\end{bmatrix}=\Theta\tilde{x}_k+\varphi_k. \tag{20}$$

且$\tilde{x}_k$的协方差满足

$$\mathrm{cov}(\tilde{x})=\Theta\mathrm{cov}(\tilde{x})\Theta^{\mathrm{T}}+E(\varphi\varphi^{\mathrm{T}}). \tag{21}$$

因此本文控制性能指标为

$$\begin{aligned}&\Lambda=\\ &\lim_{H\to\infty}\frac{1}{H}\sum_{k=1}^{H-1}E\{(u_k+\Delta u_k)^{\mathrm{T}}R(u_k+\Delta u_k)+\\ &x_k^{*\mathrm{T}}Qx_k^*\}=\\ &\lim_{H\to\infty}\frac{1}{H}\sum_{k=1}^{H-1}E\{\tilde{x}_k\Theta\tilde{x}_k^{\mathrm{T}}+\Delta u_k^{\mathrm{T}}R\Delta u_k\}=\\ &\mathrm{tr}(\mathrm{cov}(\tilde{x}_k)\Phi+\Sigma_uR),\end{aligned} \tag{22}$$

其中$\Phi=\begin{bmatrix}Q+F^{\mathrm{T}}RF & Q & -F^{\mathrm{T}}RF\\ -F^{\mathrm{T}}RF & 0 & F^{\mathrm{T}}RF\\ Q & Q & 0\end{bmatrix}$.

所提出方法系统性能损失为$\Delta\Lambda=\Lambda-\Lambda^*$,其中系统最优控制性能指标$\Lambda^*$由式(5)计算得到.

3 仿真实验

以工作于平衡状态的直流电机系统[14]为例. 设采样时间为0.1 s,考虑存在过程噪声w_k和量测噪声v_k,可得到离散化后直流电机的模型为

$$\begin{aligned}&x_{k+1}=\\ &\begin{bmatrix}-0.0019 & -0.0039\\ 0.4692 & 0.9731\end{bmatrix}x_k+\begin{bmatrix}0.2439\\ 1.61\end{bmatrix}u_k+w_k,\\ &y_k=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}x_k+v_k.\end{aligned}$$

其中:$w_k\sim N(0,W)$,$v_k\sim N(0,V)$,协方差矩阵$W=V=\mathrm{diag}(0.001,0.001)$. LQG控制器中半正定矩阵$Q=I_2$,$R=1$. 得到Kalman滤波器增益矩阵为$L=\begin{bmatrix}0.5 & -0.0005\\ -0.005 & 0.6305\end{bmatrix}$,控制反馈矩阵为

$$F=[-0.2190 \quad -0.4542].$$

取$t_0=1.1$ s,且设注入的伪周期水印信号周期$T=10$,注入的水印信号为均值为0、方差为0.4353的高斯噪声信号.此时$A^{20T}\doteq 0$. 加入水印信号20个水印信号周期后的水印信号和补偿信号分别如图2和图3所示. 由图2和图3可见,补偿信号与水印信号对应呈周期性,且在同一周期内互不相同.

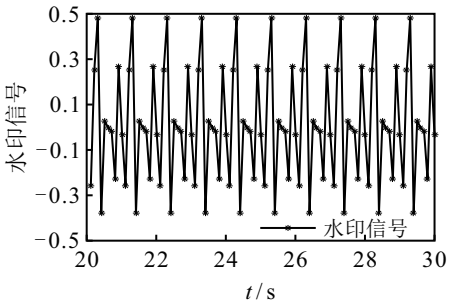


图 2 $T=10$伪周期水印信号曲线

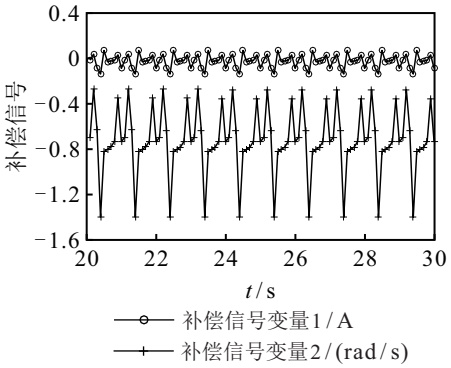


图 3 $T=10$构造的补偿信号曲线

若系统不存在过程噪声和量测噪声干扰,当注入的水印信号方差$\Sigma_u=0.05$,系统正常运行情况下的检测结果如图4所示. 由图4可见,判别函数值均为0,不存在误报. 系统存在噪声干扰时,分别取水印信号方差$\Sigma_u=0.05$和$\Sigma_u=0.4353$,系统正常运行的检测结果如图5所示. 系统的状态噪声和量测噪声将造成误检和漏检. 随着水印信号方差Σ_u增大,产生的误报减少.

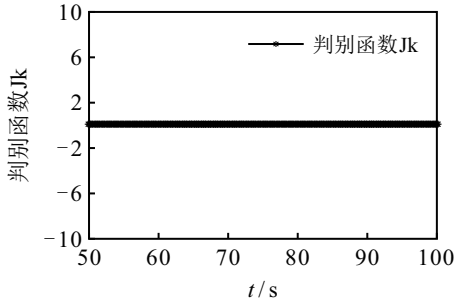
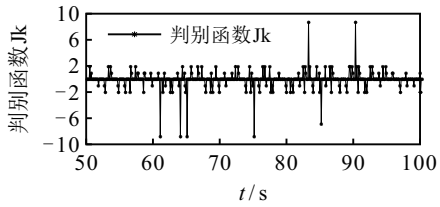
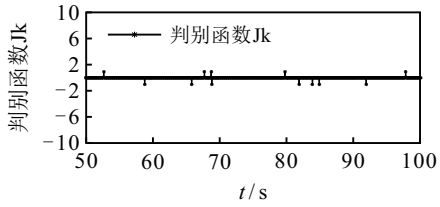


图 4 无噪声时正常情况检测



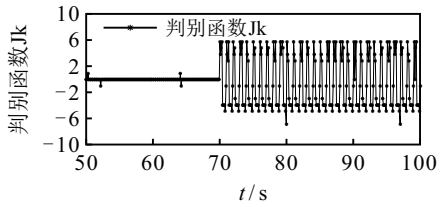
(a) 水印方差为0.05



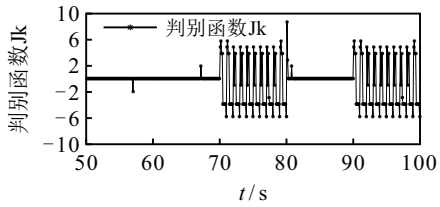
(b) 水印方差为0.435 3

图 5　有噪声时不同水印方差下正常情况检测

假设存在两种重放攻击场景: 1) 攻击者在70.1～100 s间连续重放所记录的25.5～55.4 s的测量值数据; 2) 攻击者分别在70.1～80 s、90.1～100 s两个时间段内, 连续重放所记录的25.5～35.4 s的测量值数据. 改变伪周期信号的周期大小可发现, 随着周期T减小, 检测率和误警率均有小幅减小, 而当$T>12$后检测率变化不明显但误警率有较大增加, 综合考虑后本文选择伪周期水印信号周期$T=10$. 此时两种攻击场景下的检测结果如图6所示. 图6(a)中: 判别函数值在50.1～70 s间仅有极少数时刻不为0, 在70.1～100 s间大多不为0, 可知在70.1～100 s时间段内的绝大多数时刻所提出方法均检测出了重放攻击, 仅有少数时刻存在漏检和误检. 图6(b)中: 判别函数值在70.1～80 s、90.1～100 s内大多不为0, 而在50.1～70 s以及80.1～90 s内大部分均为0, 表明对正常情况误检时刻较少, 且可有效检测出两次重放攻击. 故所提出方法对连续重放攻击以及非连续重放攻击均有效. 重复5 000次实验, 计算两类场景下的检测率分别为94.17%、90.39%, 误检率分别为4.3%、6.59%.



(a) 连续攻击场景



(b) 非连续攻击场景

图 6　$T=10$时两种重放攻击场景下检测结果

为进一步表明所提出方法的有效性, 在$T=10$时, 对比所提出方法与文献[8]、文献[14]方法的水印信号方差、检测率、误报率、控制性能损失4个指标, 结果如表1所示. 在注入方差为0.435 3的相同水印信号时, 所提出方法与文献[14]所引起的性能损失相同且小于文献[8]方法. 此时所提出方法在两类场景下的检测率分别为94.17%和90.39%, 比文献[14]方法分别提高了10.42%和7.89%, 比文献[8]方法分别提高了24.08%和21.86%. 且此时两种攻击下的误检率分别为4.3%、6.59%, 比文献[8]分别降低了6.7%和4.5%, 比文献[14]分别降低了2.07%和1.92%. 当3种方法的检测率均为85%左右时, 所提出方法的性能损失为4.835 7, 为文献[14]方法的20.77%, 为文献[8]方法的6.96%; 当3种方法的检测率均为93%左右时, 所提出方法的性能损失为11.554 5, 为文献[14]方法的39.65%, 为文献[8]方法的6.30%. 故相较现有结果, 所提出方法造成的性能损失较小. 其原因为文献[8]和文献[14]是通过判断残差的检测函数值与系统正常状态时的最大检测函数值之间的绝对大小进行攻击的检测, 加入的控制水印大小需保证受攻击后的检测函数值大于系统正常状态时的最大检测函数值, 而本文通过比较不同补偿信号下的检测函数值的相对大小实现攻击的判定, 不需要检测函数值大于系统正常状态时的最大检测函数值. 因此, 所提出方法造成的控制性能损失也更小.

表 1　重放攻击的检测率、误报率和性能损失结果

	水印信号方差	性能损失	连续攻击检测率/%	连续攻击误报率/%	非连续攻击检测率/%	非连续攻击误报率/%
本文方法	0.435 3	23.277 8	94.17	4.3	90.39	6.59
	0.33	11.554 5	93.14	8.24	92.48	8.6
	0.1	4.835 7	85.52	20.31	81.01	21
文献[8]方法	2.85	183.318 8	93.16	3.43	89.98	3.71
	1.08	69.470 1	84.67	6.54	82.33	7.38
	0.435 3	28.002 1	70.09	11	68.53	11.09
文献[14]方法	0.545 0	29.143 3	93.12	5.31	90.85	6.75
	0.435 3	23.277 8	83.75	6.37	82.5	8.51

为对比所提出方法与文献[15]方法的性能,采用文献[15]中的系统模型. 设置系统在66.1 s时刻开始重放第60.5 s的数据,重放时长为$T_{s_1} = T_0 + X_1$;在70.1 s开始重放第60.5 s的数据,重放时长为$T_{s_2} = T_0 + X_2$;在75.1 s开始重放第60.5 s的数据,重放时长为$T_{s_3} = T_0 + X_3$,其余时间系统正常运行,其中X_1、X_2、X_3根据概率密度函数$f(X) = 0.2\mathrm{e}^{-0.2X}$随机生成. 对于文献[15]中的方法,考虑$q/p = 0.5$和$q/p = 0.8$两种水印注入方式,选取周期水印序列分别为$\theta_1 = [1\,110\,010\,010]$和$\theta_2 = [1\,110\,111\,101]$. 所提出方法设置水印信号长度$T = 10$. 重复实验5 000次. 文献[15]方法在两种水印注入方式下的性能损失分别为25 308、40 404,重放攻击的检测率分别为77.43 %和85.08 %. 改变所提出方法注入水印的方差大小使得两种方法的性能损失相同,此时所提出方法的检测率分别为80.3 %和87.85 %,均高于文献[15]提出的方法.

4 结 论

针对CPS系统中重放攻击的检测问题,本文提出一种伪周期水印方法,验证了该方法对于重放攻击的可检测性. 与现有水印方法依据检测函数的绝对大小检测攻击不同,所提出方法通过比较不同补偿信号下卡方检测函数的相对大小关系实现攻击检测. 因此可在大幅降低控制量水印信号的方差大小的前提下保证较高的攻击检测率,使得在牺牲较小系统性能的情况下有效地检测到重放攻击的发生.

参考文献(References)

[1] 彭大天, 董建敏, 蔡忠闽, 等. 假数据注入攻击下信息物理融合系统的稳定性研究[J]. 自动化学报, 2019, 45(1): 196-205.
(Peng D T, Dong J M, Cai Z M, et al. On the stability of cyber-physical systems under false data injection attacks[J]. Acta Automatica Sinica, 2019, 45(1): 196-205.)

[2] Ding D R, Han Q L, Xiang Y, et al. A survey on security control and attack detection for industrial cyber-physical systems[J]. Neurocomputing, 2018, 275: 1674-1683.

[3] 孙子文, 张炎棋. 工业信息物理系统的攻击建模研究[J]. 控制与决策, 2019, 34(11): 2323-2329.
(Sun Z W, Zhang Y Q. Research on attack modeling of industrial cyber physical systems[J]. Control and Decision, 2019, 34(11): 2323-2329.)

[4] Farha F, Ning H S. Enhanced timestamp scheme for mitigating replay attacks in secure ZigBee networks[C]. IEEE International Conference on Smart Internet of Things. Tianjin, 2019: 469-473.

[5] Ye D, Zhang T Y, Guo G. Stochastic coding detection scheme in cyber-physical systems against replay attack[J]. Information Sciences, 2019, 481: 432-444.

[6] Guo H B, Pang Z H, Sun J, et al. An output-coding-based detection scheme against replay attacks in cyber-physical systems[J]. IEEE Transactions on Circuits and Systems II: Express Briefs, 2021, 68(10): 3306-3310.

[7] Mo Y L, Sinopoli B. Secure control against replay attacks[C]. The 47th Annual Allerton Conference on Communication, Control, and Computing. Monticello, 2009: 911-918.

[8] Chabukswar R, Mo Y L, Sinopoli B. Detecting integrity attacks on SCADA systems[J]. IEEE Transactions on Control Systems Technology, 2014, 22(4): 1396-1407.

[9] Zhai L J, Vamvoudakis K G. A data-based private learning framework for enhanced security against replay attacks in cyber-physical systems[J]. International Journal of Robust and Nonlinear Control, 2021, 31(6): 1817-1833.

[10] Tran T T, Shin O S, Lee J H. Detection of replay attacks in smart grid systems[C]. International Conference on Computing, Management and Telecommunications. Ho Chi Minh City, 2013: 298-302.

[11] Liu H X, Mo Y L, Yan J Q, et al. An online approach to physical watermark design[J]. IEEE Transactions on Automatic Control, 2020, 65(9): 3895-3902.

[12] Zaman A, Safarinejadian B, Birk W. Security analysis and fault detection against stealthy replay attacks[J]. International Journal of Control, 2020, 65(9): 1-14.

[13] Ferrari R M G, Teixeira A M H. Detection and isolation of replay attacks through sensor watermarking[J]. IFAC-PapersOnLine, 2017, 50(1): 7363-7368.

[14] 张正道, 杨佳佳, 谢林柏. 基于辅助信息补偿和控制信号编码的重放攻击检测方法[J]. 自动化学报, DOI: 10.16383/j.aas.c210092.
(Zhang Z D, Yang J J, Xie L B. Auxiliary information compensation based control signal-coding scheme for replay attack detection [J]. Acta Automatica Sinica, DOI: 10.16383/j.aas.c210092.)

[15] Fang C R, Qi Y F, Cheng P, et al. Optimal periodic watermarking schedule for replay attack detection in cyber-physical systems[J]. Automatica, 2020, 112: 108698.

作者简介

张正道(1976−), 男, 副教授, 博士, 从事信息物理系统安全性、系统状态监测与故障诊断等研究, Email: wxzzd@jiangnan.edu.cn;

王瑶瑶(1997−), 女, 硕士生, 从事信息物理系统攻击检测的研究, Email: 2939807290@qq.com;

谢林柏(1973−), 男, 教授, 博士生导师, 从事过程建模与控制、智能检测与系统安全性等研究, E-mail: Xie_linbo@jiangnan.edu.cn.