基于数字水印的网络控制系统回放攻击检测

颉新春[1,2], 费敏锐[1†], 杜大军[1]

(1. 上海大学 机电工程与自动化学院, 上海 200444; 2. 内蒙古科技大学 信息工程学院, 内蒙古 包头 014010)

摘要: 随着网络技术在控制系统中的广泛应用, 如何处理数据回放及数据篡改等网络攻击行为是广大科技工作者必须考虑和解决的一个问题. 在构建包含控制网络和检测网络的控制系统基础上, 本文给出了一种网络控制系统的数学模型, 并详细阐述了其传感器节点到控制器节点的数据传输方法. 考虑到控制器节点遭受回放攻击的实际情况, 通过数据帧中添加时间戳和水印数据, 并采用拟合优度检验等相关策略, 由此提出一种新的异常行为检测方法. 仿真结果表明, 该方法在一定程度上能够有效并快速地检测出针对控制器的回放攻击行为, 具有应用前景和推广参考价值.

关键词: 传感器水印; 网络控制系统; 拟合优度检验; 网络安全

引用格式: 颉新春, 费敏锐, 杜大军. 基于数字水印的网络控制系统回放攻击检测. 控制理论与应用, 2020, 37(9): 2047 – 2053

DOI: 10.7641/CTA.2020.90635

Replay attack measurement of networked control system based on digital watermark

JIE Xin-chun[1,2], FEI Min-rui[1†], DU Da-jun[1]

(1. School of Mechanical Engineering and Automation, Shanghai University, Shanghai 200444, China;
2. School of Information and Engineering, Inner Mongolia University of Science and Technology, Baotou Inner Mongolia 014010, China)

Abstract: With the wide application of networks in control systems, the process of data replay attacks and keeping the stability of networked control systems is a problem that the researchers must consider. On the basis of establishing a control system including a control network and a measurement network, a data transmission method between sensor nodes and controller nodes is proposed. The mathematical model of the networked control system is given. Considering the presence of the replay attacks to the controller nodes, a real-time anomaly detection method is presented on the basis of adding the time stamp in the data frame. The strategy of adding digital watermarking and chi-square goodness of fit test are described in detail. The simulation results show that the detection method of goodness of fit test can detect the replay attack actions in real time. The method has certain application prospects and popularization values.

Key words: sensor watermarking; networked control system; goodness of fit tests; network security

Citation: JIE Xinchun, FEI Minrui, DU Dajun. Replay attack measurement of networked control system based on digital watermark. *Control Theory & Applications*, 2020, 37(9): 2047 – 2053

1 引言

随着芯片制造技术、计算机网络及通讯技术的发展, 工业应用中的控制设备、仪表和检测装置都向智能化、网络化方向发展. 闭环控制系统的控制功能由系统中作为唯一计算单元的控制器完成逐步发展为具有高速智能计算功能并包含某种标准通信协议的传感器、执行器共同完成. 基于工业现场总线、企业局域网及Internet为基础构建分布式测控系统正逐步成为当前控制系统设计的主流. 应用于工业过程控制系统的网络及现场总线技术在提高企业的管理和运行效率的同时也增加了整个控制系统被黑客攻击的可能性[1–2]. 发生在 2010 年攻击伊朗核电站著名的“震网”病毒和发生在2000年的澳大利亚昆士兰污水厂的控制系统攻击事件说明控制系统的安全控制问题是所有生产企业不可回避的问题[3–4]. “震网”病毒通过回放传感器的历史输出数据达到了破坏实际对象的目的[5], 后者通过攻击windows操作系统, 有针对性的操控某种工业组态软件和世界上广泛使用的可

收稿日期: 2019−08−01; 录用日期: 2020−03−26.
†通信作者. E-mail: mrfei@staff.shu.edu.cn; Tel.: +86 21-56382143.
本文责任编委: 陈积明.
国家自然科学基金项目(61633016)资助.
Supported by the National Natural Science Foundation of China (61633016).

编程控制器(programmable logic controller, PLC)来破坏了被控物理对象的工作参数[6]. 目前针对网络控制系统的攻击行为包括DoS攻击、数据注入攻击、传感器回放攻击等[7–8]. DoS攻击是一种占用网络带宽的一种攻击行为, 攻击者通过向数据接收方不断发送伪装成合法的数据请求帧以占用接收方CPU资源, 降低数据接收端接收正常合法数据的效率, 从而降低了系统控制性能[9]. 数据注入攻击是在破解收发双方通讯协议的基础上, 攻击者将获取的合法数据叠加一个攻击数据后发往接收端以欺骗接收方达到破坏控制系统的目的[10]. 回放攻击是攻击者在读取一定时间范围内合法的传感器输出数据后作为攻击数据, 在攻击时间段将这些数据再伪装成传感器输出以欺骗控制器, 同时在执行器端施加信号恶意操控被控对象从而达到毁坏控制系统的目的[11–13]. 目前针对控制系统的攻击检测主要方法有特征检测和异常检测两种方法[14]. 特征检测是指在控制系统遭受黑客攻击后获取攻击行为的特征数据, 在此基础上由专业的反病毒公司研制防御软件来阻止该攻击行为. 异常检测是在获取控制系统正常运行特征数据的基础上, 在系统运行时不断重新获取新的特征数据并与原有正常特征数据对比, 通过比对结果判断系统是否遭受了异常攻击. 特征检测法由于具有攻击行为难于解析及防护滞后性的特点, 在实际应用中很难达到长期维护系统正常运行的目的. 针对异常检测方法的研究, 文献[15]通过传感器的期望输出与实际输出比较差序列是否发生变化判断是否遭受了攻击行为, 该方法采用非参数累积和(cumulative sum, CUSUM)方法判断运行中的传感器数据是否偏离了期望值. 文献[16]提出一种消息观测机制(message observation mechanism, MOM), 有效解决了无线传感器网络中的DoS攻击检测问题. 文献[17]在假定系统存在状态干扰和输出干扰均为正态分布情况下采用χ^2检验法检验控制系统是否存在数据注入攻击行为. 目前χ^2 检验法已成为最常用的异常行为检测方法之一. 当网络控制系统发生回放攻击行为时, 攻击者会发送系统正常工作时的一段传感器历史数据给控制器. 由于这些历史数据的统计特性与正常情况下没有区别, 所以χ^2检验法对这种回放攻击行为的检测没有效果. 针对网络控制系统回放攻击的研究, 文献[11]采用向控制器输出信号中添加服从一定分布规律的随机序列的方法, 在降低系统控制性能和指标的前提下提出了一种回放攻击检测策略. 在文献[18]中作者通过添加水印滤波器和均衡滤波器, 基于约定的加密数据不断切换滤波器参数以达到检测回放攻击的目的. 为了保持控制特性不变化, 该方法需在切换滤波器参数时不断初始化水印滤波器和均衡滤波器的状态. 在有较强实时性要求的应用条件下该方法很难满足要求. 本文重点研究网络控制系统中针对传感器回放攻击行为的实时检测问题, 目的在于寻找一种不影响控制系统特性的实时在线检测方法. 在回放攻击行为发生后检测器在尽可能短的时间内产生报警给控制器, 以便控制器能快速切换控制策略对被控对象形成有效保护.

2 网络控制系统的结构及传感器回放攻击

2.1 网络控制系统的结构及控制方式

网络化控制系统通常由一个多输入–多输出(multiple input multiple output, MIMO)对象、控制器、从控制器到执行机构传输命令的控制网络以及一个从传感器到控制器传输检测信息的检测网络构成. 考虑到系统的安全性, 需设计异常行为检测器. 一个网络化控制系统结构如图1所示.

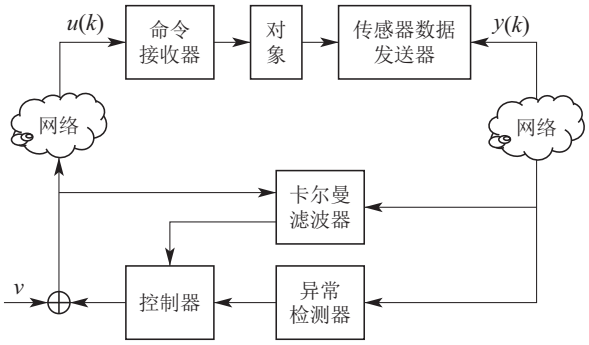

图 1 网络化控制系统结构

Fig. 1 Structure of the networked control system

一个实际MIMO对象的离散数学模型可以表述为

$$x(k+1) = Ax(k) + Bu(k) + \zeta(k), \tag{1}$$

$$y(k) = Cx(k) + \eta(k), \tag{2}$$

这里: A为$n \times n$常系数矩阵; B为$n \times p$常系数输入矩阵; C为$q \times n$输出矩阵; $u(k)$为p维输入向量, $y(k)$为q维传感器的实际输出向量; $\zeta(k)$和$\eta(k)$分别为n维状态干扰向量和q维输出干扰向量, 它们都是服从正态分布的随机过程. 假定实际对象为可稳定且状态能观测的, 系统控制方式为时间驱动方式. 当传感器数据发送器内的定时器计时结束时, 一个数据帧由传感器发送装置向控制器发出并引发控制信息的产生. 该计时时间即为系统采样时间T. 假定在该采样时间内传感器到控制器的数据传输、状态估计、控制算法及控制器到实际对象的数据传输都能够完成. 控制器采用卡尔曼滤波器获得系统的状态估计, 其迭代方法如下[19]:

1) 任意给定状态初值向量和误差协方差矩阵P_{k-1}, 计算如下先验状态估计向量:

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_{k-1}. \tag{3}$$

2) 计算先验误差协方差矩阵

$$P_k^- = AP_{k-1}A^{\mathrm{T}} + Q, \tag{4}$$

这里Q为控制通道干扰向量的协方差矩阵.

3) 计算卡尔曼增益矩阵

$$K_k = P_k^- C^{\mathrm{T}}(CP_k^- C^{\mathrm{T}} + R)^{-1}, \tag{5}$$

这里: C为观测矩阵, R为量测噪声协方差矩阵.

4) 计算状态向量估计值, 获得最小方差估计

$$\hat{x}_k = \hat{x}_k^- + K_k(y_k - C\hat{x}_k^-), \tag{6}$$

这里y_k为观测向量.

5) 更新误差协方差, 为下一步迭代做准备

$$P_k = (I - K_k C)P_k^-. \tag{7}$$

选取最优性能指标J如下:

$$J = \sum_{k=0}^{\infty}[(x^{\mathrm{T}}(k))Q_1(x^{\mathrm{T}}(k))^{\mathrm{T}} + u^{\mathrm{T}}(k)R_1u(k)], \tag{8}$$

其中: Q_1为$n \times n$正半定常数矩阵, R_1为$m \times m$正定常数矩阵. 笔者知道在给定常线性系统条件下, 使该目标函数取最小值的最优控制序列为

$$u^*(k) = v - F\hat{x}(k), \tag{9}$$

式中:

$$F = \tilde{R}^{-1}B^{\mathrm{T}}PA, \tag{10}$$

$$\tilde{R}^{-1} = R_1 + B^{\mathrm{T}}PB, \tag{11}$$

其中P为如下黎卡提方程的对称正定解[20]:

$$P = Q_1 + A^{\mathrm{T}}PA - A^{\mathrm{T}}PB(R_1 + B^{\mathrm{T}}PB)^{-1}B^{\mathrm{T}}PA. \tag{12}$$

系统正常运行时v应设定为p维常数向量, 该向量的选取应与实际对象的被控参数y_{p}相对应. v的设定按照以下几步来进行:

1) 基于理想的$y_{\mathrm{p}}(\infty)$获取一个较为理想的$u(\infty)$, 二者关系为

$$y_{\mathrm{p}}(\infty) = C(I - A)^{-1}Bu(\infty). \tag{13}$$

2) 基于这个$u(\infty)$获取对应的$x(\infty)$, 即

$$x(\infty) = (I - A)^{-1}Bu(\infty). \tag{14}$$

3) 计算状态反馈信号$w(\infty)$

$$w(\infty) = Fx(\infty). \tag{15}$$

4) 计算反馈后的输入信号$v(\infty)$

$$v(\infty) = w(\infty) + u(\infty). \tag{16}$$

2.2 传感器回放攻击场景

针对网络控制系统的回放攻击和其他网络攻击方式一样, 也是为了破坏控制系统的稳定性或对被控对象形成直接的损毁. 攻击者首先读取某段时间内系统正常工作条件下的传感器输出数据作为攻击数据. 在适当条件下需要对系统展开攻击时由攻击节点模拟传感器数据发送装置将这些攻击数据发送给控制器.

图2中攻击者从k_1时刻开始读取并记录所有传感器的输出值, 直到k_T时刻结束. 记录周期T为采样时间的整数倍($T = n \times T_s$). 攻击节点从k_1时刻开始发动回放数据, 即将记录下的k_1到k_T的数据周期性的发送给控制器. 假定攻击者已经破解了检测通道的加密方式和数据通讯格式, 并可以对其中对应的内容作任意修改. 定义矩阵

$$\Phi = (A + BF)(I - KC). \tag{17}$$

由文献[11]可知, 在回放攻击情况下, 若Φ不稳定, χ^2检测器产生的残差将会随时间增加趋于无穷大. 若Φ稳定, 由于攻击者回放的是合法的历史数据, χ^2检测器将变为无效, 不能产生有效的报警. 此时若在控制网络注入一个非法数据将有可能对实际对象产生破坏作用.

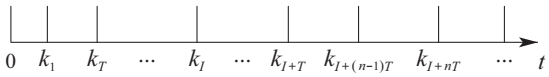

图 2 记录及回放攻击时间

Fig. 2 Time of record and replay attack

3 检测通道的数据传输模式及水印策略

在控制对象的输出端, 数据发送器将所有传感器的输出数据采样后通过检测通道将数据打包后发送给控制器. 本文定义的数据帧如图3所示.

St	DA	SA	T	Sensors	L	Data	CRC	End

图 3 检测网络数据传输格式

Fig. 3 Data transmission format of measurement network

该数据帧中所有字节均由ASCII码表示, 其各个标识段的含义见表1.

表 1 数据帧中各个标识字段的含义

Table 1 Meaning of different segment in data frame

标识段	字节数	含义
St	1	帧起始标识字节
DA	1	节点目标地址
SA	1	节点源地址
T	15	时间戳
Sensors	2	传感器数量
L	2	每个传感器数据的长度
Data	Sensors$*L$	传感器数据
CRC	4	CRC校验值
End	1	帧结束标识字节

本文假定攻击者有能力完全掌握该数据通讯格式, 并试图通过攻击节点伪造出一个数据帧. 在采用回放数据并修改时间戳的情况下, 企图通过执行器攻击以达到破坏控制对象的目的. 文献[18]指出通过在传输数据中添加相应的水印信息可以有效检测回放攻击

行为. 数字水印是利用数字作品中普遍存在的冗余数据与随机性, 向数字作品中加入不易察觉但可以判定和区分的秘密信息, 从而起到保护数字作品版权或完整性的一种技术[21–22]. 被嵌入的水印可以是一段字符、标识、ID序列号等. 水印信息通常是不可见或不可察的, 它与原始数据(如图像、音频、视频数据)紧密结合并隐藏其中, 并成为不可分离的一个整体. 数据发送端在数据帧中加入水印数据的目的在于使接收端能够利用这些数据判断出其他数据是否由约定节点发出或在传输过程中是否发生了篡改行为. 对于攻击者来说, 即使完全得到了准确的水印数据, 也非常难于得到蕴含在这些数据当中的水印信息. 一般来说, 水印信息用与一定数量的约定参数来表示, 这些参数可将水印信息确定下来, 即

$$\Delta = \varphi(\theta_1, \theta_2, \cdots, \theta_n), \tag{18}$$

这里: $\theta_i(i = 1 \sim n)$表示水印参数, Δ表示水印信息. 对于发送端来说, 在已知水印信息的情况下采用水印产生算法产生一定数量的水印数据. 而在接收端, 水印信息提取算法基于一定数量的水印数据可回算出蕴含在其中的水印信息. 如果采用水印信息提取算法得出的水印信息与发送端不同, 可认为发生了数据攻击或数据篡改行为. 水印数据和水印信息的产生过程如图4所示.

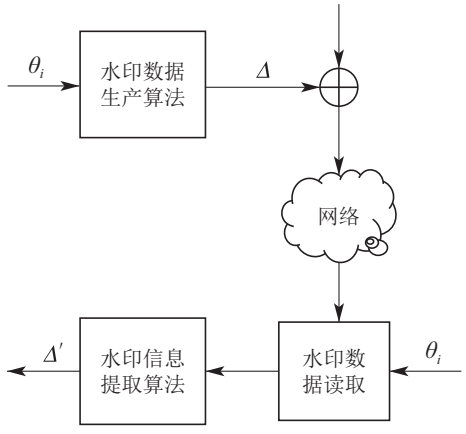

图 4 水印信息的产生与提取过程

Fig. 4 Generation and extraction procedure of watermark information

4 基于拟合优度检验的回放攻击检测

4.1 水印参数的设置与水印信息的生成

将时间戳标识字段中的某个时间值用$T_{\rm m}$来描述, 用4个参数$\theta_1, \theta_2, \theta_3$和$\theta_4$将$T_{\rm m}$表述为另外两个函数值, 即

$$T_{\rm y1} = f(\theta_1, T_{\rm m}, \theta_2), \tag{19}$$

$$T_{\rm y2} = g(\theta_3, T_{\rm m}, \theta_4), \tag{20}$$

这里f和g为加密函数. 笔者认为这4个参数只有监测网络的发送端和接收端知道, 而攻击者不知道该参数的具体数值. 以$T_{\rm y1}$为均值, $T_{\rm y2}$为标准差, 随机产生一个服从正态分布的数据, 即

$$w = \text{normrnd}(T_{y_1}, T_{y_2}). \tag{21}$$

当发送数据时, 可将该值作为水印数据隐藏在数据帧中的某个传感器数据字段中. 如果需要产生n个水印数据时, 则需要$4n$个参数. 这里提供两种水印数据隐藏方法. 第一种方法是将水印数据虚构成同等数量的传感器信息, 即将“sensors”字段中的数据加倍, 并在后续的“data”字段中添加该水印数据作为虚构的传感器数据. 另一种方法是固定好w的整数部分和小数部分位数后, 将w的有效信息隐藏到传感器数据的相应位置中. 两种方法的水印数据嵌入字段描述如图5和图6所示.

Sensors	Length of sensor data	Data1	Data2

Sensors: 真实传感器数量+虚构传感器数量

Length of sensor data: 每个传感器数据长度

Data1: 所有真实传感器数据

Data2: 所有虚构传感器数据

图 5 虚构传感器数据作为水印数据

Fig. 5 Watermark data based on fictitious sensors

Sensors	Length of sensor data	Data

Sensors: 真实传感器数量

Length of sensor data: 添加了水印数据的传感器数据长度

Data: 添加了水印的传感器数据

图 6 传感器信息嵌入水印数据

Fig. 6 Watermark data based on real sensors

4.2 基于水印信息的回放攻击行为检测

存在n组水印数据的情况下, 由于每组水印数据是基于时间戳产生的服从正态分布的随机数据, 在某段时间内(比如1 s内)这些水印数据可视为一个正态分布的随机过程. 假定在第i个传感器中隐藏的水印数据为

$$w_i \sim N(u_i, {\sigma_i}^2). \tag{22}$$

这里该变量的均值u_i和标准差σ_i由当前时间戳和隐藏的水印参数决定. n个水印信息的平方和统计量

$$Z = \sum_{i=1}^{n} (\frac{w_i - u_i}{\sigma_i})^2 \tag{23}$$

服从自由度为n的卡方分布. 即

$$Z \sim \chi^2(n). \tag{24}$$

判断系统工作是否正常可以通过分析该统计量是否服从χ^2分布来决定. 在一定时间范围内, 通过记录一

定数量的水印数据, 计算出统计量, 即通过一定数量的$Z(k)$可以获得K Pearson统计量[23]. K. Pearson统计量描述为

$$\chi^2 = \sum_{i=1}^{r} \frac{(n_i - np_i)^2}{np_i}, \tag{25}$$

这里将获得一定时间内的$Z(k)$划分为r个部分来分析. 这r个部分分别用$A_1, A_2, \cdots, A_r$来表示. 在正常情况下, 每部分的理论概率为已知的, 即出现的频数是固定不变的. 假设在一段时间内一共完成n次采样, 变量$Z(k)$分别落入$A_1, A_2, \cdots, A_r$各个区域的次数用$n_i(i = 1 \sim r)$表示, 且满足

$$n = n_1 + n_2 + \cdots + n_r. \tag{26}$$

当样本容量足够大且系统正常运行时, Pearson统计量服从自由度为$r-1$的χ^2分布, 即

$$\chi^2 \sim \chi^2(r-1). \tag{27}$$

当攻击节点对控制器数据接收端进行回放攻击时, 由于水印数据与时间戳有关, 故Pearson统计量不再服从χ^2分布. 通过拟合优度检验法可以检验系统是否发生了攻击行为. 选定一个显著性水平α, 作如下假设:

$$H_0 : p(A_i) = p_i,\ i = 1, 2, \cdots, r, \tag{28}$$

其中

$$\sum_{i=1}^{r} p(A_i) = 1. \tag{29}$$

该假设成立表示系统运行正常, 否则表示存在数据注入攻击行为. 对应的拒绝域为

$$W = \{\chi^2 \geqslant c\}. \tag{30}$$

判别门限c由显著性水平α决定, 即

$$c = \chi^2_\alpha(r-1). \tag{31}$$

控制器接收端的攻击检测算法如图7所示. 该算法实际上通过时间戳来对回放攻击做检测. 如果攻击者有能力篡改回放数据的时间戳, 那么回放数据所携带的水印信息必然与当前时间戳不符, 从而引发异常报警器报警. 如果攻击者企图篡改传感器数据和水印数据, 由于不知道通讯双方约定好的水印参数和加密函数, 很难伪造出满足要求的水印数据.

5　数据仿真及结果

一个线性离散系统的系统矩阵和输入、输出矩阵描述为

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.99 & 0 \\ 0 & 0 & 0.99 \end{pmatrix}, \tag{32}$$

$$B = \begin{pmatrix} 0.05 & 0 & 0 \\ 0 & 0.04 & 0 \\ 0 & 0 & 0.04 \end{pmatrix}, \tag{33}$$

$$C = \begin{pmatrix} 0.5 & 0.4 & 0.3 \\ 0.3 & 0.2 & 0.1 \\ 0.2 & 0.4 & 0.6 \end{pmatrix}. \tag{34}$$

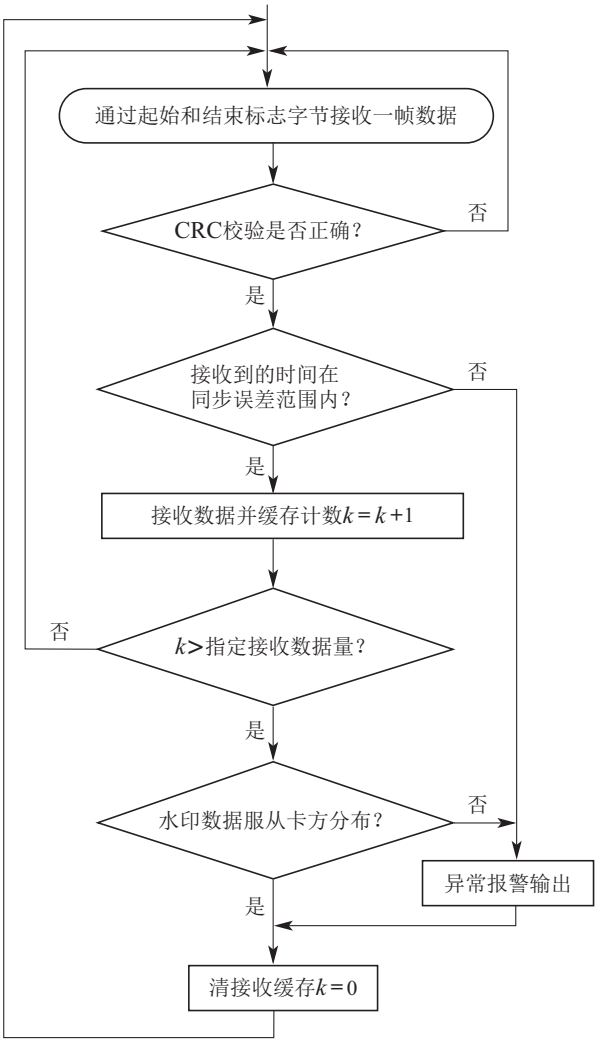

图 7　回放攻击检测算法

Fig. 7　Measurement algorithm of replay attack

从系统矩阵及其输入、输出矩阵可知, 该系统为可稳定且可观测的. 假定实际对象的状态干扰向量$\eta(k)$和输出干扰向量$\xi(k)$都服从均值为0的正态分布, 各个干扰信号之间相互独立. 其协方差阵分别表示为

$$Q_{\rm p} = \mathrm{dig}\{0.01^2, 0.02^2, 0.03^2\}, \tag{35}$$

$$R_{\rm p} = \mathrm{dig}\{0.012^2, 0.018^2, 0.024^2\}. \tag{36}$$

在没有回放攻击情况下, 该系统在短时间内就达到稳定状态. 系统的3个传感器数据需要通过检测网络进行数据传输, 这里构建了3个虚拟传感器用于传输水印数据. 随机水印数据的产生以时间戳的分钟信息$T_{\min}$作为标准驱动源. 均值u和标准差σ的产生采用如

下两个加密函数:

$$u_i = \theta_1 \times T_{\min}^2 + \theta_2, \tag{37}$$

$$\sigma_i = \theta_3 \times T_{\min} + \theta_4. \tag{38}$$

通讯双方约定的水印参数的设置见表2. 设定系统采样时间为0.1 s, 即每隔0.1 s由传感器数据发送器驱动一次数据传输并触发控制器完成一次控制算法. 异常检测器每获取100个数据后对水印数据$w_i(i = 1 \sim 3)$进行分析和检验. 正常情况下在一定时间范围内统计量

$$Z = \sum_{i=1}^{3}(\frac{w_i - u_i}{\sigma_i})^2 \tag{39}$$

服从自由度为3的χ^2分布, 即

$$Z \sim \chi^2(3). \tag{40}$$

按照卡方分布表将统计量Z的取值概率划分为14($r = 14$)个区域. 统计量Z在各个区域的取值概率见表3.

表 2 水印参数的设置

Table 2 Setting of different watermark parameter

水印数据	θ_1	θ_2	θ_3	θ_4
1	0.005	5	0.05	0.2
2	0.002	3	0.02	0.1
3	0.001	7	0.08	0.6

表 3 χ^2分布表中区域概率取值

Table 3 Probability values of chi-square distribution

区域	Z取值范围	理论概率
A_1	$Z \leqslant 0.072$	0.005
A_2	$0.072 < Z \leqslant 0.115$	0.005
A_3	$0.115 < Z \leqslant 0.216$	0.015
A_4	$0.216 < Z \leqslant 0.352$	0.025
A_5	$0.352 < Z \leqslant 0.584$	0.05
A_6	$0.584 < Z \leqslant 1.213$	0.15
A_7	$1.213 < Z \leqslant 2.37$	0.25
A_8	$2.37 < Z \leqslant 4.108$	0.25
A_9	$4.108 < Z \leqslant 6.251$	0.15
A_{10}	$6.251 < Z \leqslant 7.815$	0.05
A_{11}	$7.815 < Z \leqslant 9.348$	0.025
A_{12}	$9.348 < Z \leqslant 11.354$	0.015
A_{13}	$11.354 < Z \leqslant 12.838$	0.005
A_{14}	$12.838 < Z$	0.005

如果每隔10 s进行一次数据分析, 即选取100个数据作为样本进行数据统计($n = 100$). 采样到n个数据后分别统计落入各个区域(A_i)的个数n_i, 然后计算统计量

$$\chi^2 = \sum_{i=1}^{14}\frac{(n_i - np_i)^2}{np_i}. \tag{41}$$

设置置信度水平$\alpha = 0.005$, 则拒绝域门限值为

$$c = \chi_\alpha^2(r-1) = \chi_{0.005}^2(13) = 29.819. \tag{42}$$

通过拒绝域

$$W = \{\chi^2 \geqslant 29.819\}. \tag{43}$$

判断是否存在回放攻击行为. 在没有回放攻击情况下, 检验值K. Pearson统计量的输出值如图8所示. 从图8可以看出, 在100次的检测中几乎所有的检验值都小于报警门限值, Pearson统计量没有超出拒绝域. 攻击者从第2~5.2 s开始记录30组数据作为攻击数据. 从20 s开始周期性地向控制器发动回放攻击. 在攻击者篡改了回放数据时间戳情况下, K. Pearson统计量的输出如图9所示. 从图9可以看出, 所有的检验值都超过了报警门限值29.819, 没有发生误报现象, 正确率为100%. 攻击者从第2~16 s开始记录140组数据作为攻击数据. 从20 s开始周期性地向控制器发动回放攻击. K. Pearson统计量的输出如图10所示. 从图10可以看出, 所有的检验值也都超过了报警门限值29.819, 也没有发生误报现象, 正确率始终保持为100%.

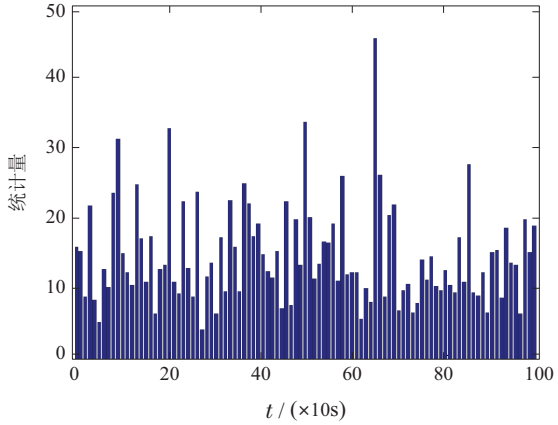

图 8 正常情况下的K. Pearson统计量

Fig. 8 Pearson statistics under normal conditions

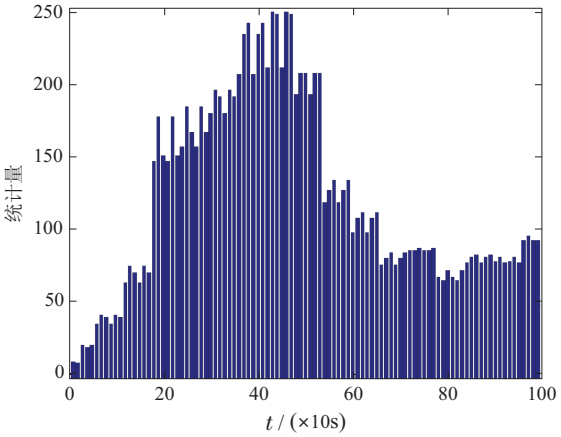

图 9 记录周期为3 s的K. Pearson统计量

Fig. 9 Pearson statistics when record time is 3 s

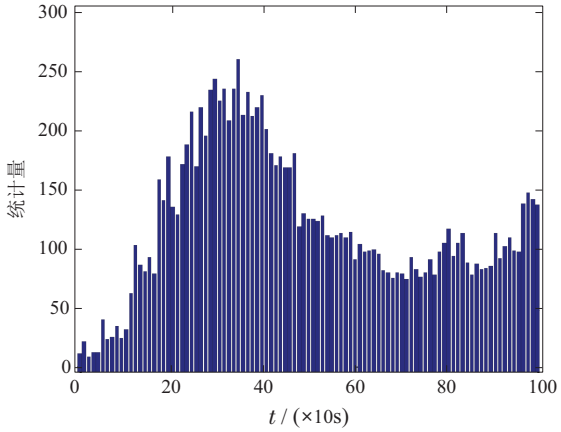

图 10　记录周期为14 s的K. Pearson统计量

Fig. 10　Pearson statistics when record time is 14 s

6　结语

在考虑网络控制系统中存在回放攻击及数据篡改等网络攻击行为情况下, 本文提出的通过在数据帧中添加时间戳和水印数据, 在接收方利用χ^2拟合优度检验对非法水印数据进行判别方法, 目的在于异常检测器能够快速地向控制器给出报警信号以切换控制策略. 通过基于加密函数及不同水印参数随机产生水印数据的方法, 大大增加了攻击者的破解难度, 在一定程度上起到了保护控制系统的作用. 仿真结果表明该保护及检测方法能够满足一般控制系统的实时性要求. 对于报警实时性要求较高的网络控制系统, 可采用提高时间戳精度或减少采样样本的策略以提高报警实时性能. 考虑到网络控制系统中数字水印数据的篡改行为, 未来的研究重点在于水印数据的隐藏及篡改定位问题. 对于一些为了达到破坏网络控制系统的蓄意性隐秘攻击和数据篡改行为, 还需做专门研究并针对性地给出应对方案.

参考文献:

[1] ALVARO A, CARDENAS, AMINS, SASTRY S. Research challenges for the security of control systems. *The 3rd USENIX Workshop on Hot Topics in Security*. San Jose, CA: USENIX, 2008.

[2] DING D, HAN Q L, XIANG Y, et al. A survey on security control and attack detection for industrial cyber-physical systems. *Neurocomputing*, 2017, 275: 1674 – 1683.

[3] CHEN T M, ABU-NIMEH S. Lessons from stuxnet. *Computer*, 2011, 44(4): 91 – 93.

[4] ALVARO A C, AMIN S, LIN Z S, et al. Attacks against process control systems: Risk assessment, detection, and response. *Proceedings of the 6th ACM Symposium on Information, Computer & Communications Security*. Hong Kong, China: [s.l.], 2011, 3: 355 – 366.

[5] MO Y, WEERAKKODY S, SINOPOLI B. Physical authentication of control systems: Designing watermarked control inputs to detect counterfeit sensor outputs. *IEEE Control Systems*, 2015, 35(1): 93 – 109.

[6] NTALAMPIRAS S. Automatic identication of integrity attacks in cyber physical systems. *Expert Systems with Applications*, 2016, 58(C): 164 – 173.

[7] MAHMOUD M S, HAMDAN M M, BAROUDI U A. Modeling and control of cyber-physical systems subject to cyber attacks: a survey of recent advances and challenges. *Neurocomputing*, 2019, 338: 101 – 115.

[8] PAN L, ZHENG X, CHEN H X, et al. Cyber security attacks to modern vehicular systems. *Journal of Information Security and Applications*, 2017, 36: 90 – 100.

[9] SHI L. Analysis and design of secure cyber-physical systems. *Control Theory and Technology*, 2014, 12(4): 413 – 414.

[10] HUA L, WANG Z, HAN Q L, et al. State estimation under false data injection attacks: Security analysis and system protection. *Automatica*, 2018, 87: 176 – 183.

[11] MO Y L, SINOPOLI B. Secure control against replay attacks. *The 47th Annual Allerton Conference on Communication, Control & Computing*. Monticello: IEEE, 2009, 10: 911 – 918.

[12] WANG Xiaowu, LIU Ying. Direction-based replay attack defense mechanism. *Communications Technology*, 2019, 52(6): 1500 – 1503. (王效武, 刘英. 基于方向的重放攻击防御机制. 通信技术, 2019, 52(6): 1500 – 1503.)

[13] SABINA H, ROTONDO D, ESCOBET T. Detection of replay attacks in cyber-physical systems using a frequency-based signature. *Journal of the Franklin Institute*, 2019, 356(5): 2798 – 2824.

[14] MARKAM V, DUBRY L S M. A general study of associations rule mining in intrusion detection system. *International Journal of Emerging Technology and Advanced Engineering*, 2012, 2(1): 347 – 356.

[15] RAJASEGARAR S, LECKIE C, PALANISWAMI M. Anomaly detection in wireless sensor networks. *IEEE Wireless Communications*, 2008, 15(4): 34 – 40.

[16] ZHANG Y Y, LI X Z, LIU Y A. The detection and defence of DoS attack for wireless sensor network. *Journal of China Universities of Posts & Telecommunications*, 2012, 19: 52 – 56.

[17] PRAKASH J, PATWARDHAU S C, NARASIMHAN S. A supervisory approach to fault-tolerant control of linear multivariable systems. *Industrial & Engineering Chemistry Research*, 2002, 41(9): 2270 – 2281.

[18] FERRARI RICCARDO M G, TEIXEIRA ANDRE M H. Detection and isolation of replay attacks through sensor watermarking. *IFAC-Papers Online*, 2017, 50(1): 7363 – 7368.

[19] MO Y, GARONE E, CASAVOLA A, et al. False data injection attacks against state estimation in wireless sensor networks. *The 49th IEEE Conference on Decision and Control*. Atlanta: IEEE, 2010, 11: 5967 – 5972.

[20] GONG Deen. *Introduction to Discrete Control System Theory*. Beijing: China Railway Press, 2004.
(龚德恩. 离散控制系统理论引论. 北京: 中国铁道工业出版社, 2004.)

[21] ZHANG Zhiming, WANG Lei, XU Naiping. A survey of digital watermarking applied for information hiding techniques. *Computer Engineering and Applications*, 2002, 23: 46 – 49.
(张志明, 王磊, 徐乃平. 信息隐藏技术中的数字水印研究. 计算机工程与应用, 2002, 23: 46 – 49.)

[22] TAN Hui. Digital watermark technology and its application. *China Computer & Communication*, 2018, 13: 221 – 222, 225.
(谭慧. 数字水印技术及其应用. 信息与电脑(理论版), 2018, 13: 221 – 222, 225.)

[23] MAO Shisong, ZHOU Jixiang. *Probability Theory and Mathematical Statistics*. Beijing: China Statistic Press, 2004.
(茆诗松, 周纪芗. 概率论与数理统计. 北京: 中国统计工业出版社, 2004.)

作者简介:

颉新春　博士研究生, 目前主要研究方向为网络化测控系统及安全控制, E-mail: jjjxxxccc@163.com;

费敏锐　博士, 教授, 博士生导师, 主要研究方向为智能化网络控制理论、系统和仿真以及其关键技术在电力、冶金、电器、实验装备自动化中的应用, E-mail: mrfei@staff.shu.edu.cn;

杜大军　博士, 教授, 主要研究方向为网络化先进控制及应用, E-mail: ddj@i.shu.edu.cn.