(1) 两个测量不相交.

首先比较当前时刻(t)任意两个传感器的测量,如果两个传感器的测量不相交,则至少有一个传感器提供了故障测量,即它们在 t 时刻是弱不一致关系.

$$S_i^F(t)\cap S_j^F(t)=\varnothing \Rightarrow WI(S_i,S_j,t).$$

然后融合过去和当前的测量,将所有传感器的测量从 t–1 时刻映射到 t 时刻,在 t 时刻共 $2N$ 个测量.最后在 t 时刻比较任意两个测量值,注意,不包括同一个传感器不同时刻的两个测量.如果这两个测量不相交,那么这两个传感器之间存在弱不一致关系.

$$S_i^{F'}(t)\cap S_j^{F'}(t)=\varnothing \Rightarrow WI(S_i,S_j,t),\ i\neq j.$$

(2) 两个测量相交.

当两个测量相交时,主要是从不同的角度利用融合间隔和历史测量来判断故障.

第 1 步:首先计算 t 时刻的融合间隔,让每个传感器和融合间隔进行比较,如果两个传感器 S_i 和 S_j 均不与融合间隔相交,并且这两个传感器的测量相交,则 S_i 和 S_j 是弱不一致关系.

$$\begin{cases} S_i^F(t)\cap F_{N+1}^f(S,t)=\varnothing \\ S_j^F(t)\cap F_{N+1}^f(S,t)=\varnothing \\ S_i^F(t)\cap S_j^F(t)\neq\varnothing, i\neq j \end{cases} \Rightarrow WI(S_i,S_j,t).$$

第 2 步:结合历史测量进行判断.首先计算 t–1 时刻的融合间隔,将每个传感器在 t–1 时刻的测量与该融合间隔进行比较;然后把 t–1 时刻的测量映射到 t 时刻,再进行两两比较.当满足下列条件时,说明传感器 S_i 和 S_j 是弱不一致关系.

$$\begin{cases} S_i^F(t-1)\cap F_{N+1}^f(S,t-1)=\varnothing \\ S_j^F(t)\cap F_{N+1}^f(S,t)=\varnothing \\ S_i^{F'}(t)\cap S_j^F(t)\neq\varnothing, i\neq j \end{cases} \Rightarrow WI(S_i,S_j,t).$$

算法 1 给出了具体实现.第 1 行~第 8 行实现了在 t 时刻任何两个测量之间的比较,并将不一致性信息存储在弱不一致数组中.第 9 行将所有传感器的测量值从时间 t–1 时刻映射到 t 时刻.第 10 行调用融合算法计算融合间隔.第 11 行实现了当前时刻和历史的传感器测量之间的比较.第 12 行完成了单个传感器和融合间隔的比较,并将这些传感器的信息存储在不一致数组中.使用弱不一致检测方法最终会得到一个弱不一致数组,里面存放了在每一个时刻存在弱不一致关系的传感器的信息.这些信息将会在传感器的攻击检测和识别中用到.

Algorithm 1. Weak inconsistent detection algorithm.

Input: N measurements of the abstract sensor.

1: $w\leftarrow 0$;
2: **for** i=0 $\rightarrow$ N–2 **do**
3: **for** j=i+1 $\rightarrow$ N–1 **do**
4: **if** $S_i^F(t)\cap S_j^F(t)==\varnothing$ **then**
5: $weaks[w{+}{+}]\leftarrow(i{+}1,j{+}1,t)$
6: **end if**
7: **end for**
8: **end for**
9: $S_i^{F'}(t)\leftarrow S_i^F(t-1)$ (i=1,2,…,N);
10: fusion();
11: compareTwoSensorHistory ($S_i^{F'}(t), S_j^F(t)$);
12: SensorAndFusionCompare(sensors, $F_N^f(S,t)$);
13: return weaks;

3.1.2 强不一致检测

本小节通过采用上述的不一致概念来展示本文提出的传感器攻击检测方法.由定义 6 可知,如果两个传感

器之间存在强不一致关系,则说明两个传感器中至少有一个是非瞬态故障.因此,由定义 3 攻击的定义可知,如果存在时间 $t \leqslant T$,使得传感器 S_i 和 S_j 是强不一致的,则说明系统中存在被攻击的传感器.其中,T 是系统运行的总时间.

引理 2(攻击检测).

$$\sum_{t=1}^{t} SI'(S_i,S_j,t) \geqslant 1 \Rightarrow AD(S_i,S_j,t), t \geqslant 1,$$

其中,如果 $SI(S_i,S_j,t)$存在,则 $SI'(S_i,S_j,t)=1$;否则,$SI'(S_i,S_j,t)=0$.

证明: $\sum_{t'=1}^{t} SI'(S_i,S_j,t) \geqslant 1$ 说明传感器 S_i 和 S_j 在 t 时刻是强不一致关系,由强不一致的定义可知,S_i 和 S_j 至少有一个是非瞬态故障.根据定义 3 可知, $\neg TF(S_i,t) \vee \neg TF(S_j,t) \Rightarrow A(S_i,t) \vee A(S_j,t)$, 因此,可以断定系统中存在攻击.证毕. □

3.2 攻击识别

上述的攻击检测方法仅考虑检测系统中是否存在传感器被攻击,并没有考虑哪个传感器被攻击的问题.在本小节中,为了确定哪个传感器受到攻击,本文假设系统中至多有 $s(s<N-1)$个传感器受到攻击.本文提出的攻击识别方法是:累积强不一致信息,在强不一致对中,如果传感器 S_i 出现的次数超过 s,则称 S_i 被攻击了.

引理 3. 给定传感器 S_i 和时间 t,用 $degree(S_i,t)$表示在 t 时刻与 S_i 存在强不一致关系的传感器的数量.

$$degree(S_i,t) > s \Rightarrow A(S_i,t).$$

证明:假设有 $n(n>s)$个传感器和 S_i 相连,与 S_i 相连的传感器用 S_j 表示.由于 S_i 和 S_j 是强不一致关系,如果 S_i 没有被攻击,那么 S_j 必须被攻击.此时,共有 n 个传感器受到攻击,这与最多有 s 个传感器被攻击的假设相矛盾.证毕. □

3.3 瞬态故障模型参数选择

本文提出的 BPI 方法需要精确的瞬态故障模型.现在的制造商一般会提供传感器的瞬态故障规范.尽管制造商有时候会提供算法所需要的全部参数,但对应不同的应用场景和需求,这些参数可能并不是最佳的(例如,在被高建筑物包围的环境中使用 GPS).此外,针对不同的传感器攻击算法,有些参数制造商可能无法提供,比如本文提出的 BPI 算法的参数 f_i.那么,就有必要基于经验数据来开发瞬态故障模型.因此,本节提出一种通过构建 ROC 曲线来选择瞬态故障模型参数的新方法,用于抽象传感器的攻击检测.

受试者工作特征曲线为 ROC 曲线(receiver operating characteristic curve).通常,该曲线的横轴是假阳性概率(false positive rate),纵轴是真阳性概率(true positive rate),其目的主要是用来选择最佳的界限值.将多个实验得到的 ROC 曲线绘制到同一个坐标中,就能很直观地看出,ROC 曲线的最左上角的点表示错误率最少且准确率最高,意味着该点对应的阈值是最好的.

本文提出的选择瞬态故障模型参数方法的思想是:从实际的实验平台上获取大量的实验数据(传感器的值),把这些数据应用到攻击检测算法中(BPI 方法),从而得到不同场景下传感器的攻击识别率和误报率.然后把各个场景得到的 ROC 曲线绘制到同一个坐标系中.找到识别率最高、误报率最小的点,即图中最左上角的点,该点对应的参数就是最佳的参数.然而,由于每个参数对应的值可能有无数多个,不可能把每一个参数的每个值都进行实验来得到其对应的识别率和误报率.因此,在进行实验之前,需要筛选出这些参数可能的取值范围,然后再取这些范围内的点进行实验,最终确定参数的取值.在第 1.3 节中提出的瞬态故障模型有 3 个参数.参数 δ_i 使用制造商提供的值,另外两个参数 f_i 和 w_i 使用本文提出的基于构建 ROC 曲线选择参数的方法进行确定.其基本步骤如下.

(1) 初步筛选参数,选择若干个值作为 f_i 的候选值.根据受到攻击的传感器数量,本文考虑 3 种攻击情况:只有一个传感器受到攻击,另外两个传感器没有故障;一个传感器受到攻击,另外两个传感器仅有一个传感器存在瞬时故障;一个传感器受到攻击,另外两个传感器均存在瞬时故障.通过对无攻击无故障的传感器数据添加上面 3

种攻击和故障,会得到一些训练数据.此外,分别用 f_{i-1} 和 f_{i+1} 代替 3 种类型的攻击中的 f_i 来形成一系列新的训练集.通过实验得到不同 f_i 对应的识别率和误报率.最后,本文选择几个误报率较小但识别率还能接受的 f_i 作为候选值.这样做虽然不能保证攻击检测率最大,但可以保证误报率最小.

(2) 根据第 1.4 节中设计的攻击模型,为原始数据(无攻击、无故障的传感器数据)添加 3 种类型的攻击.针对(1)中选取的每个候选值进行实验,获得对应窗口的误报率和识别率后,构建 ROC 曲线,然后确定最终的参数值.本文选取参数的具体细节将在第 4.1 节中加以详细介绍.

3.4 传感器攻击检测实例

本小节主要是想通过一个简单的例子来进一步解释上述传感器攻击检测和识别算法.为了便于解释,本文使用一个不一致图 $G(V,E)$来描述该算法.在不一致图中,每个传感器对应一个顶点,两个顶点之间的关系代表它们之间的不一致关系.如果该关系是弱不一致关系,则该图表示弱不一致图,用 *WI_Graph* 表示;如果该关系是强不一致关系,则该图表示强不一致图,用 *SI_Graph* 表示.该不一致图的定义如下.

定义 7(不一致图).

$$V=\{S_1,S_2,\ldots,S_N\},$$

$$E=\{(i,j)|WI(S_i,S_j,\mathrm{t})\ \text{or}\ SI(S_i,S_j,t)\}.$$

本文设计了如图 4 所示的实例,图中有 4 个传感器,窗口 W=5,垂直虚线表示真实值(未知).假设系统中传感器 S_1 和 S_4 受到攻击(最多有 2 个传感器被攻击).注意,不要求被攻击的传感器在每一轮次中均受损,但需要满足在给定窗口中是非瞬态故障,每个传感器的瞬态故障模型见表 1.

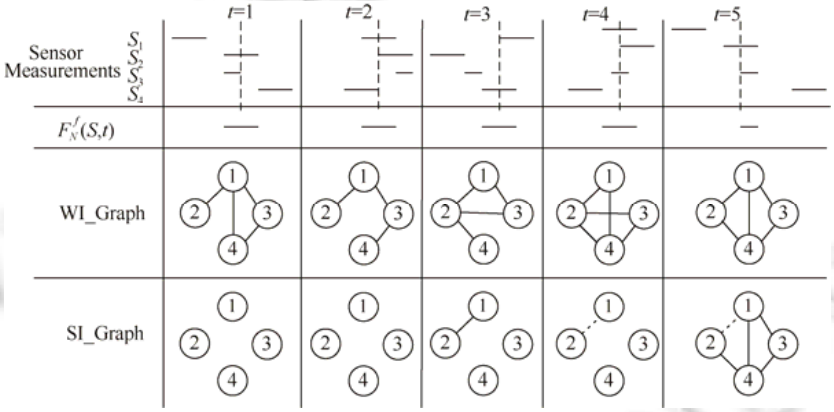

Fig.4 BPI algorithm example

图 4 BPI 算法实例

Table 1 The parameters of the transient fault model in the example of Fig.4

表 1 图 4 实例中对应的瞬态故障模型的参数

传感器	δ_i	f_i	w_i
S_1	1	1	5
S_2	1	1	3
S_3	0.5	2	5
S_4	1	1	5

根据第 3.1.1 节介绍的弱不一致检测算法,得到如 *WI_Graph* 所示的弱不一致关系图.从图中可以看出,系统在每一轮次中都检测到了弱不一致对.在 t=1 时,传感器 S_1 和 S_2 之间存在弱不一致关系,这意味着这两个传感器至少有一个提供了错误的测量.到 t=3 为止,S_1 和 S_2 之间共出现了 3 次弱不一致关系.根据第 3.1.2 节介绍的强不一致检测方法可知,S_1 和 S_2 之间存在强不一致关系,这意味着系统中存在攻击,其他类似.该实例中所有传感器的强不一致关系如图 4 中的 *SI_Graph* 所示.*SI_Graph* 中的虚线代表在之前的检测中已经出现过该强不一致对.虽然,在 t=3 时,检测到系统中存在攻击,但此时无法判断哪个传感器受到了攻击.直到 t=5 时,发现传感器 S_1 和 S_4

的度均为 3,大于 2,根据第 3.2 节中给出的攻击识别算法可以识别出在 t=5 时 S_1 和 S_4 均受到了攻击.

4 实验评估

本节从 EV3 机器人平台上获取实际的实验数据来评估 BPI 算法的性能.首先介绍实验的基本设置,之后介绍瞬态故障模型的参数选择问题,然后评估几种算法的攻击检测和识别性能,最后分析误报的原因.

4.1 实验设置

本文的实验平台选择 LEGO EV3 地面车辆,如图 5 所示.EV3 是 2013 年 LEGO 公司开发的第三代 MINDSTORMS 机器人.它可以安装多个传感器,包括超声波、电机(内嵌角度传感器)、陀螺仪、颜色传感器等.根据需求,本文使用 2 个大型电机和 1 个超声波传感器来测量 EV3 的速度.这 3 个传感器均可提供 10Hz 的测量.

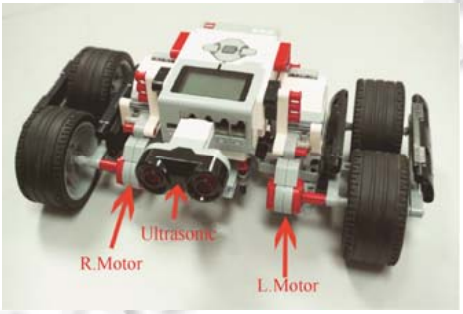

Fig.5 EV3 robot

图 5 EV3 机器人

为了获得鲁棒的瞬态故障模型参数,本文根据第 3.3 节提出的参数选择方法选择最终的实验参数.具体是,根据制造商提供的测量误差设置传感器的瞬态故障模型的参数 δ_i.对某个确定的窗口 w_i,通过真实的训练数据来确定 f_i 的值.本文首先以 0.7m/s 的恒定速度驱动 EV3 机器人直线运动,每个传感器收集 400 个测量数据.利用第 3.3 节提出的参数选择方法进行实验,为了模拟真实的攻击情形,窗口 w_i 中每个传感器提供的故障数是随机的.根据实验结果,本文建立了如图 6 所示的 ROC 曲线来确定 f_i 的值.ROC 曲线的 x 轴表示误报率(误报的数量/识别的总数量),y 轴表示识别率((识别的总数–误报的总数)/测试总数).为了避免混淆,图 6 中不包括 w_{400} 的情况.从图中可以看出,左上角的数据点具有更高的识别率、更低的误报率,该点对应的 f_i 是最佳阈值.表 2 总结了可变窗口大小的故障模型参数,其中,w_i 表示窗口大小为 i 的检测器.

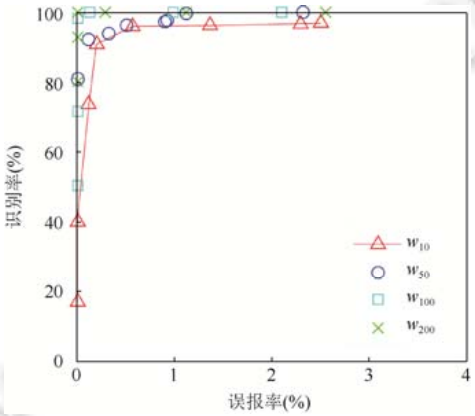

Fig.6 ROC curve: Identification rate and false alarm rate under three types of attacks

图 6 ROC 曲线:3 种类型攻击下的识别率和误报率

Table 2 Fault models for the sensors on EV3

表 2 EV3 上瞬态故障模型的参数

Detector	Ultrasonic		L.Motor		R.Motor	
	δ	f	δ	f	δ	f
W_{10}	0.2	1	0.037	1	0.037	2
W_{100}	0.2	20	0.037	16	0.037	18
W_{200}	0.2	31	0.037	20	0.037	24
W_{300}	0.2	40	0.037	35	0.037	30
W_{400}	0.2	64	0.037	45	0.037	50

4.2 检测性能分析

为了评估本文在第 3 节中介绍的 BPI 算法,本节首先在平坦的地面上以恒定的速度驱动 EV3 来收集每个传感器的数据.本文收集了 20 次来自 EV3 设备上的未被攻击的传感器数据,其中,所有传感器以 10Hz 平均采样 40s,因此,每个传感器有 400 个测量值.

本文假设在固定窗口中,攻击者对 3 个抽象传感器中的 1 个进行攻击,具体哪一个传感器被攻击是未知的.同时,对于被攻击的传感器,本文添加第 1.4 节中设计的 3 种攻击类型.注意,被攻击的传感器不要求在每一轮中都提供故障的测量,每个窗口中传感器提供故障测量的数量是随机的,但需要保证在给定窗口中是非瞬态故障.此外,在所有的检测数据中,其他传感器可能存在瞬态故障,也可能不存在瞬态故障.表 3 给出了不同窗口和攻击场景下 3 种攻击检测和识别方法的检测率.其中,KF 方法是一种基于卡尔曼滤波器的传感器攻击检测方法.从表中可以看出,本文提出的 BPI 算法能够比已有的算法检测到更多的攻击.对于 3 种攻击,BPI 方法在任何窗口中都比 PI 和 KF 方法更加鲁棒,并且随着窗口大小的增加,BPI 方法逐渐达到稳态检测率.对于偏差攻击,KF 方法与 BPI 的检测性能相近,BPI 检测器的平均检测率比 PI 检测器大约高 25%,比 KF 方法高约 2%.然而,对于随机攻击,KF 方法的检测性能明显低于 BPI 方法,小窗口中 BPI 检测器比 PI 检测器平均约高 53%,在大窗口中约高 35%,平均比 KF 方法高 14%.特别地,对于隐身攻击来说,BPI 方法的优势是显而易见的.现有方法的检测率几乎是 0,而 BPI 方法能够检测到攻击,其检测率平均能达到 90%以上.PI 方法不能检测到隐形攻击,是因为隐身攻击的特点就是尽可能地最大化融合间隔并使当前时刻任意两个传感器之间的间隔尽可能地相交.而 PI 方法恰好是基于两个传感器之间的间隔不相交来判断故障的,所以,基于 PI 的方法无法检测到这种攻击.基于 KF 的方法在大的窗口中偶尔可以检测到攻击,但检测率仅有 0.01%左右.并且,其误报率很高,这将在下一小节中详细加以介绍.

Table 3 Detection rate

表 3 检测率

(a) 偏差估计

检测率(%)	W=10	W=100	W=200	W=300	W=400
PI	54.6	63	79.82	83.5	88.97
KF	99.7	99.48	94.16	96.21	98.1
BPI	99.68	99.71	99.86	99.95	99.99

(b) 随机攻击

检测率(%)	W=10	W=100	W=200	W=300	W=400
PI	31.53	38.97	45.89	55.79	71.73
KF	86.22	69.67	65.54	87.01	94.58
BPI	86.59	92	97.46	99.18	99.89

(c) 隐身攻击

检测率(%)	W=10	W=100	W=200	W=300	W=400
PI	0	0	0	0	0
KF	0	0	0.01	0.02	0.021 4
BPI	89.36	93.23	93.69	96.45	98.59

4.3 识别性能分析

本文提出的 BPI 算法的识别性能与检测性能几乎相同,但是识别比检测需要花费更长的时间.表 4 显示了

不同窗口和攻击情形下 3 种传感器攻击检测方法的识别率.这些结果表明,对于偏差攻击和随机攻击,除了 KF 方法外,其他两种方法的识别率通常随窗口大小而有所改善.从表 4(a)可以看出,对于偏差攻击,BPI 方法的识别率大约平均比 PI 方法高 30%左右,仅比基于卡尔曼的方法高 3.3%左右.对于随机攻击来说,BPI 方法大约平均比 PI 方法高 49%左右,比基于 KF 的方法约高 14%左右.从表 4(b)可以看出,基于 KF 的方法对于随机攻击的识别性能明显降低.这是由于,随机攻击为被攻击的传感器在被攻击的时刻随机添加一个 0~0.7 的一个偏差值,当这个偏差值较小时,由于测量值和估计值的偏差比较小,就会导致 KF 方法无法检测到故障.特别地,从表 4(c)中可以看出,对于隐身攻击,基于 PI 的方法的识别率为 0,基于 KF 的方法的识别率在大的窗口中偶尔能够检测到 1 次攻击,仅 0.01%左右,并且其误报非常高.然而,本文提出的 BPI 方法平均可达到 93%.

Table 4 Identification rate

表 4 识别率

(a) 偏差估计

识别率(%)	W=10	W=100	W=200	W=300	W=400
PI	53	59.75	76.09	79.1	80.67
KF	99.7	99.48	94.16	96.21	98.1
BPI	99.48	99.64	99.68	99.78	99.99

(b) 随机攻击

识别率(%)	W=10	W=100	W=200	W=300	W=400
PI	29.25	38.92	41.65	50.9	68.37
KF	86.22	69.67	65.54	87.01	94.58
BPI	86.54	92.1	96.8	98.41	99.17

(c) 隐身攻击

识别率(%)	W=10	W=100	W=200	W=300	W=400
PI	0	0	0	0	0
KF	0	0	0.01	0.02	0.021 4
BPI	86.33	90.87	95.54	96	97.11

4.4 误报分析

表 5 总结了不同窗口大小中 3 种方法的误报率.实验结果表明,3 种方法都存在误报,这可能是由于这些窗口中存在瞬态故障的原因.此外,我们注意到,BPI 方法的误报率略高于 PI 方法.主要原因是瞬态故障的存在不能保证每轮被损坏的传感器数量不超过 f,即无法保证融合间隔一定包含真实值.此外,由于使用传感器历史测量,在某些攻击情形中可能由于从根据历史测量预测的当前测量不够准确,导致误报增加.然而,从表中可以看出,KF 方法的误报率非常高.其出现误报的原因主要是由于瞬态故障的存在,导致 KF 方法产生残留污染和残余淹没,从而造成误报.此外,需要注意的是,对于偏差攻击和随机攻击,KF 方法的误报率与 BPI 相近,但对于隐身攻击,其误报率能达到 50%以上.

Table 5 False rate

表 5 误报率

误报率(%)	W=10	W=100	W=200	W=300	W=400
PI	0.75	0.14	0	0	0
KF	0.41	7.39	94.18	51.71	45.87
BPI	1.05	0.26	1.31	0.82	1.05

5 总 结

本文研究了在存在瞬态故障时 CPS 的安全问题.首先在 Marzullo 提出的经典融合算法的基础上,通过融入历史测量提出了一种改进的融合算法,该算法可以得到更精确的融合间隔,具有更强的鲁棒性.此外,本文结合历史测量和融合间隔提出了一种新颖的传感器攻击检测和识别方法,用于具有测量相同物理变量的多个传感器的 CPS.并且,提出了一种基于构建 ROC 曲线的方法来选择瞬态故障模型.最后在 EV3 平台上获得实际的实验数据,验证了算法的性能,并与现有的基于卡尔曼滤波器的方法和基于 PI 的方法进行了比较.实验结果表明,

该算法在各种攻击场景下均优于现有的算法.基于本文的评估,未来的工作包括尝试使用形式化验证方法来验证所提出算法的正确性,并将该算法部署在实际系统上进行工业实践.

References:

[1] Miao F, Zhu Q, Pajic M, Pappas GJ. Coding schemes for securing cyber-physical systems against stealthy data injection attacks. IEEE Trans. on Control of Network Systems, 2017,4(1):106–117.

[2] Kim KD, Kumar PR. Cyber-physical systems: A perspective at the centennial. Proc. of the IEEE, 2012,100:1287–1308.

[3] Kong LL. Analysis of deception models and detection algorithms on CPS control layer [MS. Thesis]. Shanghai: East China University of Science and Technology, 2015 (in Chinese with English abstract).

[4] Jiang Y, Song H, Wang R, Gu M, Sun J, Sha L. Data-centered runtime verification of wireless medical cyber-physical system. IEEE Trans. on Industrial Informatics, 2017,13(4):1900–1909.

[5] Jiang Y, Zhang H, Song X, Jiao X, Hung WNN, Gu M, Sun J. Bayesian-network-based reliability analysis of plc systems. IEEE Trans. on Industrial Electronics, 2013,60(11):5325–5336.

[6] Yang K, Wang R, Jiang Y, Luo C, Guan Y, Li X, Shi Z. Enhanced resilient sensor attack detection using fusion interval and measurement history. In: Proc. of the 2018 Int'l Conf. on Hardware/Software Codesign and System Synthesis (CODES+ISSS). 2018. 1–3. [doi: 10.1109/CODESISSS.2018.8525941]

[7] Cardenas AA, Amin S, Sastry S. Secure control: Towards survivable cyber-physical systems. In: Proc. of the Int'l Conf. on Distributed Computing Systems Workshops. IEEE, 2008. 495–500.

[8] Checkoway S, Mccoy D, Anderson D, Kantor B, Shacham H, Savage S, Koscher K, Czeskis A, Roesner F, Kohno T. Comprehensive experimental analyses of automotive attack surfaces. In: Proc. of the Usenix Conf. on Security. 2012. 6.

[9] Koscher K, Czeskis A, Roesner F, *et al*. Experimental security analysis of a modern automobile. IEEE Journal of Selected Topics in Quantum Electronics, 2010,41(3):447–462.

[10] Slay J, Miller M. Lessons learned from the maroochy water breach. In: Proc. of the Int'l Conf. on Critical Infrastructure Protection. 2007. 73–82. [doi: 10.1007/978- 0-387-75462-8_6]

[11] Farwell JP, Rohozinski R. Stuxnet and the future of cyber war. Survival, 2011,53(1):23–40.

[12] Xiao L, Boyd S, Lall S. A scheme for robust distributed sensor fusion based on average consensus. In: Proc. of the Int'l Symp. on Information Processing in Sensor Networks. IEEE, 2005. 9.

[13] Olfati-Saber R, Shamma JS. Consensus filters for sensor networks and distributed sensor fusion. In: Proc. of the IEEE Conf. and the European Control Conf. on Decision and Control, CDC-ECC 2005. 2006. 698–6703.

[14] Yang K, Wang R, Jiang Y, Song H, Luo C, Guan Y, Li X, Shi Z. Sensor attack detection using history based pairwise inconsistency. Future Generation Computer Systems, 2018,86:392–402.

[15] Marzullo K. Tolerating failures of continuous-valued sensors. ACM Trans. on Computer Systems, 1990,8(4):284–304.

[16] Ivanov R, Pajic M, Lee I. Attack-resilient sensor fusion for safety-critical cyber-physical systems. ACM Trans. on Embedded Computing Systems, 2016,15(1):1–24.

[17] Ivanov R, Pajic M, Lee I. Resilient multidimensional sensor fusion using measurement history. In: Proc. of the Int'l Conf. on High Confidence Networked Systems. 2014. 1–10.

[18] Kalman RE. A new approach to linear filtering and prediction problems. Journal of Basic Engineering Transactions, 1960, 82(Series D):35–45.

[19] Kwon C, Hwang I. Security analysis for cyber-physical systems against stealthy deception attacks. In: Proc. of the American Control Conf. IEEE, 2013. 3344–3349.

[20] Jayasimha DN. Fault tolerance in a multisensory environment. In: Proc. of the 13th Symp. on Reliable Distributed Systems, SRDS'94. 1994. 2–11.

[21] Park J, Ivanov R, Weimer J, *et al*. Sensor attack detection in the presence of transient faults. In: Proc. of the 6th ACM/IEEE Int'l Conf. on Cyber-physical Systems. ACM, 2015. 1–10.

[22] Willsky AS. A survey of design methods for failure detection in dynamic systems. Automatica, 1975,12(6):601–611.

[23] Shoukry Y, Martin P, Tabuada P, Srivastava M. Non-invasive spoofing attacks for anti-lock braking systems. In: Proc. of the Int'l Conf. on Cryptographic Hardware and Embedded Systems. Springer-Verlag, 2013. 55–72.

附中文参考文献:

[3] 孔令霖.CPS 控制层欺骗攻击模型与检测算法的研究[硕士学位论文].上海:华东理工大学,2015.

杨康(1992－),女,山东菏泽人,硕士,主要研究领域为 CPS 的安全性,形式化验证.

李晓娟(1968－),女,博士,教授,CCF 专业会员,主要研究领域为系统形式建模与验证,机器人系统软件安全,计算机网络协议分析.

王瑞(1981－),女,博士,副教授,CCF 专业会员,主要研究领域为形式化方法.

施智平(1974－),男,博士,教授,CCF 高级会员,主要研究领域为形式化,人工智能.

关永(1966－),男,博士,教授,博士生导师,CCF 专业会员,主要研究领域为形式化验证,系统可靠性,嵌入式系统.

Xiaoyu Song(1963－),男,博士,教授,博士生导师,主要研究领域为形式化方法.