控制与决策
Control and Decision

云环境下工业信息物理系统现场层安全策略决策方法

朱美潘, 杨健晖, 李欣格, 杜鑫, 周纯杰

引用本文:
朱美潘,杨健晖,李欣格,杜鑫,周纯杰. 云环境下工业信息物理系统现场层安全策略决策方法[J]. 控制与决策, 2024, 39(1): 281–290.

在线阅读 View online: https://doi.org/10.13195/j.kzyjc.2022.0149

云环境下工业信息物理系统现场层安全策略决策方法

朱美潘[1], 杨健晖[1], 李欣格[1,2], 杜 鑫[1], 周纯杰[1,2†]

(1. 华中科技大学 人工智能与自动化学院，武汉 430070；2. 华中科技大学 网络空间安全学院，武汉 430070)

摘 要: 云环境下工业信息物理系统架构的转变使得工业现场设备更加暴露于网络攻击下,对工业现场层提出更高的安全需求. 随着系统结构愈渐复杂,网络攻击更加智能,系统难以准确获取安全状态,传统的基于状态的安全决策方法将不能实现有效防护,对此提出一种工业信息物理系统现场层安全策略决策方法. 首先,根据功能结构划分现场区域,分析潜在的攻击目标、攻击事件与系统防御策略间的关联性,构建攻击防御树;然后,从攻击和防护属性的视角,利用模糊层次分析法量化防御策略收益;接着,结合部分攻击状态构建部分可观的马尔可夫决策过程模型,通过求解模型得到最优安全策略;最后,以简化的田纳西-伊斯曼过程控制系统为对象验证所提出方法能够有效地决策出最优安全策略.

关键词: 工业信息物理系统；现场层；安全策略决策；部分可观的马尔可夫决策过程

中图分类号: TP273 **文献标志码:** A

DOI: 10.13195/j.kzyjc.2022.0149 **开放科学(资源服务)标识码(OSID):**

引用格式: 朱美潘, 杨健晖, 李欣格, 等. 云环境下工业信息物理系统现场层安全策略决策方法[J]. 控制与决策, 2024, 39(1): 281-290.

A security decision-making approach for field layer of cloud-integrated industrial cyber-physical systems

ZHU Mei-pan[1], *YANG Jian-hui*[1], *LI Xin-ge*[1,2], *DU Xin*[1], *ZHOU Chun-jie*[1,2†]

(1. School of Artificial Intelligence and Automation，Huazhong University of Sicience and Technology，Wuhan 430070，China；2. School of Cyber Science and Engineering，Huazhong University of Sicience and Technology，Wuhan 430070，China)

Abstract: The transformation of the cloud-integrated industrial cyber-physical systems' architecture makes industrial field equipment more exposed to the cyber-attacks, which puts forward higher security requirements for the industrial field layer. As the structure becomes more complex and cyber-attacks become more intelligent, it is difficult to accurately obtain the security state, and the traditional state-based security decision-making method will not achieve effective protection. This paper proposes a security decision-making approach for the field layer of industrial cyber-physical systems. First, the field area is divided according to the functional structure, and then the attack defense tree is constructed by analyzing the potential correlation between attack goals, attack events and defense strategies. Then, from the perspective of attack and defense attributes, the fuzzy analytic hierarchy process is used to quantify the payoff of the defense strategy. Combined with part of the attack state to construct the partially observable Markov decision process model, and the optimal security strategy is obtained by solving the model. Finally, a simplified Tennessee-Eastman process control system is used to verify that the proposed method can effectively decide the optimal security stategy.

Keywords: industrial cyber-physical systems；field layer；security decision-making；partially observable Markov decision process

0 引 言

工业信息物理系统(industrial cyber-physical systems, ICPS)是工业生产基础设施的核心部分,也是工业向数字化、智能化和网络化发展的重要支撑[1],广泛应用于生产制造、化工、智能电网等领域[2]. 在“中国制造2025”和“工业互联网”的背景下,以云计算为基础的新一代信息技术正大力推动智能制造等工业行业的信息化、智能化、网络化发展,云计算

收稿日期: 2022-01-21；**录用日期:** 2022-09-03.

基金项目: 国家自然科学基金项目(61873103, 62127808, 61433006).

责任编委: 张文安.

†通讯作者. E-mail: cjiezhou@hust.edu.cn.

为工业转型升级、产业创新发展提供了重要支撑[3],通过构建安全、稳定、知识共享以及高度适应且可扩展的云端能力是ICPS的发展趋势之一. 然而,随着云计算的引入,ICPS逐渐由封闭的传统架构向“工业云平台-企业私有云-通信网络-现场控制层-物理设备层”的架构演变. 相较于传统架构,这种开放式的架构使得工业现场的仪表、装置、物理对象等设施更易受到网络攻击的危害,云平台的资源共享易带来数据泄露、丢失的风险,泛在连接加大了高级持续性威胁(APT)渗透进现场系统的可能性[4-5]. 已有很多学者聚焦于工业云平台的安全防护[6],而大多针对ICPS的网络攻击目的是渗透进系统内部破坏现场运行过程或损害物理设备[7],一旦工业现场遭受恶意攻击,可能造成设备损害、财产损失、人员伤亡等重大危害. 随着网络攻击技术逐渐提高,传统的隔离、访问控制等安全防护技术无法应对新的安全威胁形式,急需研究以ICPS现场层安全为核心的信息安全防护方法来抵御智能复杂的网络攻击.

安全策略决策作为信息安全防护体系的关键环节,负责在检测到网络入侵后根据系统安全状态制定合适的安全策略,从而缓解、消除信息攻击对ICPS造成的影响[8]. 现有的ICPS安全策略决策方法的研究大多是以系统安全状态完全已知为前提[9-11],如Fessi等[9]从防御方角度考虑提出了一种多属性的网络安全决策模型,采用遗传算法搜寻可选的安全动作;Hao等[11]结合攻击方与防御方提出了一种基于自适应马尔可夫策略的最优安全策略制定方法,用于防范智能电网中的虚假数据注入. 然而,云环境的ICPS作为新型工业控制系统的典型对象,不仅具有规模庞大、工况复杂等特征,且面临网络攻击更加智能和隐蔽的安全挑战,系统恐难以感知攻击行为. 另外,现场运行设备间的强耦合性,加大了攻击目标识别和明确的难度,这使得现场层的安全策略决策具有很大的不确定性,基于状态的安全策略决策方法很难适用. 有学者提出了系统安全状态不确定下的最优防御策略选取方法,如Hu等[12]考虑了计算机网络中防御者面临的未知可能性和漏洞成功地利用未知影响,利用强化学习来选择最优防御策略,但是该方法以传统的计算机网络为对象开展防护研究,并不适用于ICPS; Miehling等[13]将如何最优地干扰攻击者的进程视为一个部分可观察的马尔可夫决策过程,利用不完美信息来指定最佳防御行动,但是该方法未能充分考虑现场设备计算资源的局限性难以支撑计算复杂的安全决策方法. 对此,有研究提出将云计算应用于ICPS的安全防护体系,利用云端的资源优势实现全局最优防护,然而,云端的安全决策只适用于周期性的、非实时的决策,且云边的通信延时和通信安全问题还未得到很好的解决,不能满足ICPS的可用性需求和强实时性需求. 因此,如何在现场层通过可感知到的不完全安全信息快速作出最优安全策略决策是目前的研究重点.

综合以上分析,针对云环境下的ICPS现场层面临严峻的安全威胁及其安全防护需具有自适应决策和实时响应能力的需求,本文提出一种部署于边缘计算节点的现场层区域性安全策略决策方法,结合边缘计算的快速响应优势实现实时的区域性最优安全策略决策. 首先,按照系统功能结构划分现场区域,在边缘节点中对每个现场区域进行安全策略决策,通过缩小决策规模来减小边缘节点的计算量,再结合现场区域的功能结构分析潜在的攻击目标、攻击事件与系统防御策略间的关联性,构建攻击防御树,直观地展示现场区域可能的攻击路径和防御措施,并在攻击防御树的攻击层和防御层分别引入攻击属性和防御属性,利用模糊层次分析法实现防御策略收益量化,通过综合考虑攻防双方因素合理评估防御策略的防护能力;然后,结合系统结构、安全知识、历史数据、组件信息以及防御收益等对现场区域构建部分可观的马尔可夫决策过程(partially observable Markov decision process, POMDP)模型来刻画现场层安全策略决策过程中的不确定性,利用实时观测到的攻击信息求解模型得到最优策略,实现系统安全状态部分已知下的最优安全策略决策;最后,以简化的田纳西过程控制系统为案例阐述决策方法执行过程,并通过与马尔可夫决策过程模型对比验证所提出决策方法的可行性和有效性.

1 云环境下ICPS现场层安全策略决策框架

中国工业互联网产业联盟发布的《工业云安全防护参考方案》[3]给出了云环境下的ICPS体系架构,如图1所示. 其中:工业云提供开放的工业增值服务,企业私有云支撑系统的运营管理,通信网络实现云与现场的连接以及柔性灵活组网,现场控制层控制物理生产过程,物理设备层提供底层数据基础支撑,边缘计算节点实现边缘数据分析和应用. 与传统架构相比,新型架构在外部网络中部署了工业云平台实现生产方式和服务模式的柔性、弹性和灵活性,在内部企业层部署了私有云以实现生产过程和设备管理优化,边缘计算节点为企业提供就近的边缘智能服务.

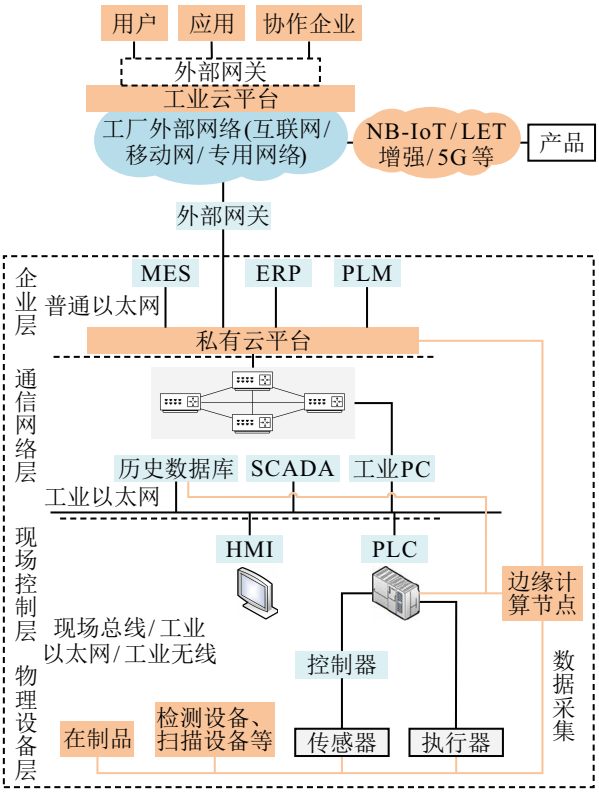

图 1　云环境下ICPS体系架构

云计算的引入在带来生产便利的同时,使得工业现场的设备、物理对象等设施也暴露于网络攻击下,导致ICPS面临更加严峻的安全威胁,如[3]: 1) 云计算的泛在连接增大了数据泄露、丢失的风险; 2) 工业云中接入了许多的智能设备,这可能会导致非法接入、非法窃听和非法控制等问题; 3) 云平台与现场的通信连接易遭受分布式拒绝服务攻击等. 然而,工业云是对ICPS信息层的扩充发展,ICPS的现场运行系统作为关键基础设施的核心依然是网络攻击的首要目标,大部分外部网络攻击通过信息层渗透至物理层中,破坏工业现场的正常运行或损坏现场设备. 基于云计算的现场层防护方法虽然可利用云端的资源优势和全局视角对大量复杂的工业数据进行处理和分析,但是,云边通信时延和通信安全依然是目前亟待解决的问题,需要有部署在现场运行系统的信息安全防护技术以实现及时可靠的信息安全防护.

随着边缘计算在ICPS中的应用,给边缘数据分析和应用运行带来了更好的支撑环境[14],能够弥补云计算大数据分析过程的时延性高、周期性长、网络耗能严重等缺陷. 对此,本文考虑到现场层强实时性的防护需求和计算资源局限的特点,提出一种部署于边缘计算节点的现场层区域性实时安全策略决策方法,其总体框架如图2所示. 根据系统的功能结构特征划分现场区域,在边缘计算节点中部署每个现场区域的安全策略决策模型,主要包括防御收益量化模块和安全策略决策模块,决策出的最优安全策略再与云端下发的全局安全策略相融合,得到最终的安全策略.

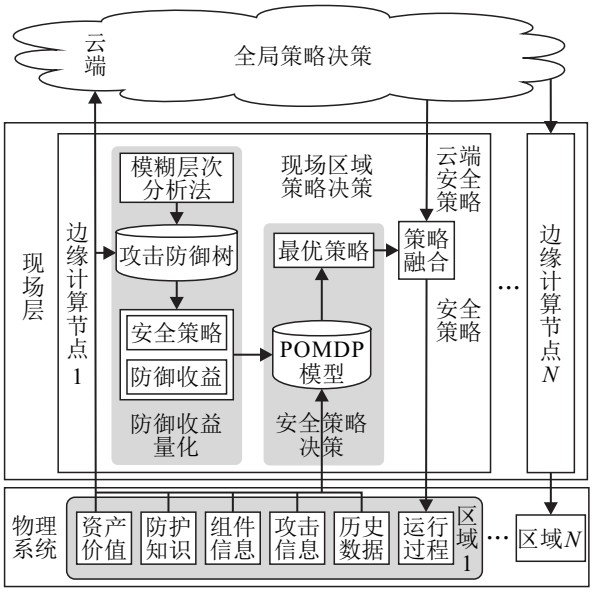

图 2　云环境下ICPS现场层安全策略决策框架

防御收益量化模块通过分析系统结构、攻击者攻击目标和手段以及防护知识来构建攻击防御树,从而对现场区域的攻击事件和防御策略建模; 然后,结合模糊层次分析法,在攻击层引入攻击属性和系统资产价值计算攻击节点收益; 再在防御层引入防御属性和攻击节点收益计算得到防御节点收益.

安全策略决策模块对现场区域构建POMDP模型. 首先,分析系统结构和防护知识设定状态空间和观测空间; 然后,通过分析历史运行数据总结经验得到观测函数; 接着,通过系统漏洞利用率设计安全状态转移概率; 最后,利用防御策略收益量化模块得到行动空间和收益函数,从而构建出POMDP决策模型. 根据检测到的实时攻击信息对决策模型求解,计算出最大防御策略收益值,从而得到最优防御策略.

云边的安全策略融合模块是在云端安全策略到达边缘后,对云边的安全策略进行融合后再执行. 云端的决策是在综合考虑了各边缘子系统的重要性后作出的全局最优决策,因此,云边安全策略融合应优先选择云端策略. 安全策略融合是从策略属性出发判断策略的关系(包括策略重合、策略冲突和策略不冲突3种关系),然后进行去重和去冲突的过程,具体如图3所示. 但是,当云端受到攻击造成云边通信断开时,直接执行边缘安全策略. 本文主要聚焦于边缘的安全策略决策方法,因此不对云边策略融合进行详细阐述.

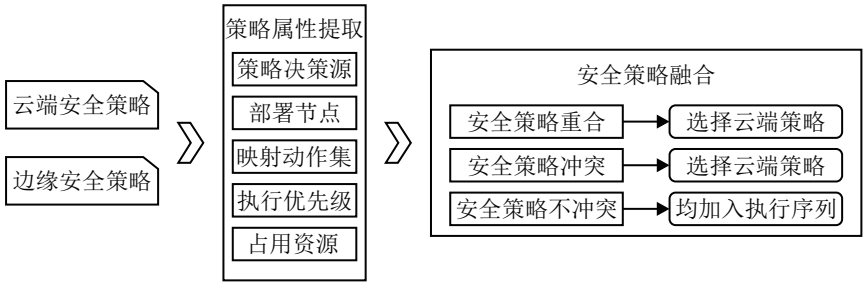

图 3 云边安全策略融合

2 安全策略决策方法

现场层安全策略决策方法主要包括防御策略收益量化模块和安全策略决策模块,其中防御策略收益为安全策略决策的评估指标. 下面将对这两部分进行详细描述.

2.1 防御策略收益量化

防御策略的收益是对防御策略的防护能力的评价,需要通过综合考虑防御策略的消耗成本及其缓解的系统风险值来得到. 本文利用攻击防御树对ICPS现场区域系统中的攻击事件和防御事件进行关联建模,然后在攻击层和防御层引入模糊层次分析法来计算防御策略的收益值.

2.1.1 攻击防御树构建

攻击树模型是一种利用树形结构描述系统所面临的安全威胁和系统可能受到的多种攻击的模型[15],常被用于分析系统存在的攻击路径、脆弱性评估、风险评估等. 攻击树包含根结点、各级子结点和叶结点,其中:根结点为攻击目标,叶结点为攻击事件,子结点为实现父结点的步骤. 攻击树的各分支为实现攻击目标可采取的各种攻击路径. 攻击防御树(attack-defense tree, ADTree)对攻击树进行了扩展,用一个图形表示攻击者可能采取的攻击措施以及防御者可使用的防御措施[16],即该图形包含2种相反类型的节点:攻击节点和防御节点. 可对节点赋予属性参数,从而实现各节点的功能量化.

构建攻击防御树不仅需要明确现场控制系统结构信息,熟悉系统功能组件及其关联关系,还需要分析潜在的攻击目标和攻击手段,掌握系统完备的防护知识. 图4为攻击防御树模型,包括目标节点G、中间结点M、攻击节点A以及防御节点D. 目标节点为攻击者的最终攻击目标,攻击者一般以破坏系统作为攻击目标;中间节点为实现目标节点的要素,可选择能够到达目标节点的系统设备作为中间节点;攻击节点即攻击者可采取的攻击手段或威胁事件;防御节点为针对攻击手段或威胁事件可采取的防御策略. 通过构建攻击防御树可得到现场区域中的攻击事件和防御策略以及两者间的攻防对应关系,其中防御策略的集合构成了POMDP模型中的行为空间.

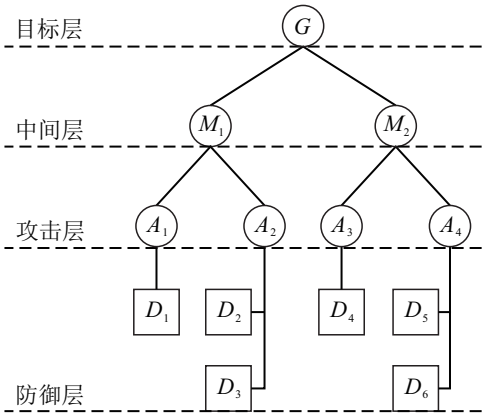

图 4 攻击防御树模型

2.1.2 模糊层次分析法模型

层次分析法是一种层次权重决策分析方法,其将定性与定量分析相结合,根据决策者的经验评价各因素的相对重要性,从而计算出各因素权重值. 模糊层次分析法(fuzzy analytic hierarchy process, FAHP)将模糊数学与层次分析法相结合,其将层次分析法中的标度值改为了模糊数标度,在一定程度上解决了判断矩阵一致性问题[17]. 在计算防御策略收益时需要对各攻击属性和防御属性进行重要性评价,可通过模糊层次分析法计算其各属性的权重,从而更合理地计算防御策略收益.

利用FAHP评价各因素相对重要性步骤如下.

1) 确定因素集,即评价指标$U=(u_1,u_2,\dots,u_n)$,指各攻击属性或防御属性.

2) 构造模糊判断矩阵$A=(a_{ij})_{n\times n}$,如表1所示. 根据各评价指标的相对重要性构造判断矩阵,使用[0, 1]标度表量化这种相对重要性,如表2所示.

表 1 模糊判断矩阵

A	u_1	u_2	$\cdots$	u_n
u_1	a_{11}	a_{12}	$\cdots$	a_{1n}
u_2	a_{21}	a_{22}	$\cdots$	a_{2n}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
u_n	a_{n1}	a_{n2}	$\cdots$	a_{nn}

表 2　[0, 1]相对重要性标度

标度	含义
0.5	u_i与u_j相比,两者同样重要
0.6	u_i比u_j稍微重要
0.7	u_i比u_j明显重要
0.8	u_i比u_j重要得多
0.9	u_i比u_j极端重要
0.1、0.2、0.3、0.4	若u_i与u_j相比得到a_{ij},则u_j与u_i相比得到$a_{ji}=1-a_{ij}$

3) 一致性校验. 当矩阵A满足以下条件时:

$$a_{ij}=0.5,\ i=1,2,\ldots,n; \tag{1}$$

$$a_{ij}=1-a_{ji},\ i,j=1,2,\ldots,n; \tag{2}$$

$$a_{ij}=a_{ik}-a_{jk}+0.5,\ i,j,k=1,2,\ldots,n. \tag{3}$$

则称为模糊一致矩阵.

若构造出的模糊判断矩阵不具有一致性,则可利用如下公式进行一致性转换:

$$a'_{ij}=\frac{1}{n}\sum_{k=1}^{n}(a_{ik}-a_{jk}+0.5). \tag{4}$$

4) 各因素权重计算,一般采用行和法,计算如下式所示:

$$w_i=\frac{1}{\sum\limits_{j=1}^{n}\dfrac{1}{a_{ij}}-n},\ i=1,2,\ldots,n. \tag{5}$$

2.1.3　防御策略收益计算

防御策略收益值是通过在构建的攻击防御树的攻击层和防御层分别引入FAHP计算得到. 在攻击层计算攻击节点收益,综合考虑攻击节点的攻击成本和攻击造成的严重程度,用攻击节点导致的风险值作为攻击节点收益;在防御层计算防御节点收益,综合考虑防御节点的防御成本和防御效果,防御效果为缓解的风险值,即攻击节点收益. 防御策略收益计算步骤如下.

1) 计算基于风险的攻击节点收益. ICPS的系统风险值取决于系统发生风险的概率和发生风险所产生的后果[18],因此风险值可由攻击事件发生的概率和造成的损失的乘积得到,即

$$\text{risk}(A_i)=p(A_i)\times q(A_i). \tag{6}$$

其中:$p(A_i)$为攻击事件A_i发生的概率,$q(A_i)$为A_i导致的资产损失.

选择攻击难度、攻击被发现的可能性和攻击后果的严重程度3个属性作为攻击节点的评价指标[19]. 攻击节点风险概率值如下式所示:

$$p(A_i)=w_{\text{diffi}}\times u_{\text{diffi}}+w_{\text{possi}}\times u_{\text{possi}}+w_{\text{sever}}\times u_{\text{sever}}. \tag{7}$$

其中:w为各攻击节点评价指标的权重,由式(5)计算得到;u为各评价指标的效用值,可用其等级评分的倒数得到. 根据现有研究,可得到等级评分标准,如表3所示.

表 3　攻击属性等级评分标准

攻击难度	等级	攻击被发现的可能性	等级	攻击后果的严重程度	等级
很难	5	很难	1	很严重	1
难	4	难	2	严重	2
中等	3	中等	3	中等	3
容易	2	容易	4	一般	4
很容易	1	很容易	5	不严重	5

2) 计算防御节点收益. 防御节点收益要考虑防御策略自身成本和缓解的系统风险,可用攻击节点的收益减去防御节点成本得到,计算如下式所示:

$$U(D_i)=\sum_{j=1}^{m}\varepsilon_j\times\text{risk}(A_j)-\text{cost}(D_i). \tag{8}$$

其中:ε_j为针对攻击事件A_j防御是否成功,防御成功为1,否则为0;$\text{risk}(A_j)$为攻击事件A_j的风险值;$\text{cost}(D_i)$为防御节点D_i的成本.

防御节点成本通过防御策略的消耗时间、占用资源和对系统的影响3个防御属性描述,计算如下式所示:

$$\text{cost}(D_i)=$$
$$w_t\times\text{time}+w_r\times\text{resource}+w_e\times\text{effect}, \tag{9}$$

其中w为各防御属性的权重,由式(5)计算得到. 各防御属性效用值用等级评分表示,如表4所示.

表 4　防御属性等级评分标准

消耗时间	等级	占用资源	等级	对系统的影响	等级
很长	5	很多	5	严重影响	5
长	4	多	4	影响	4
中等	3	中等	3	中等	3
短	2	少	2	一般影响	2
很短	1	很少	1	不影响	1

2.2　基于POMDP的安全策略决策

安全策略决策是在对系统的安全情况进行感知分析后制定安全防御策略,针对系统难以准确获取安全状态的现状,本文通过构建POMDP模型,利用观测到的不完美安全状态信息进行现场层的动态安全策略决策.

2.2.1　基于POMDP的安全策略决策

POMDP为部分可观测域中的优化行动提供了一般规划和决策框架,它非常适合存在普遍不确定性但是仍然需要进行决策的现实问题. 通常假设一个完整且正确的世界模型,具有随机状态转换、不完美

状态跟踪和奖励结构,找到一种能够最大化预期回报收益的行动策略[20]. POMDP可用一个7元组表示,其含义如下.

1) S为状态空间,$S=(s_1,s_2,\dots,s_m)$为系统所有可能状态的有限集合. 用状态描述当前的环境,系统中的每种可能的变化均分别对应空间中的某个状态. 在POMDP中状态是不能直接观察到的,Agent只能根据观测信息计算状态空间S上的信念分布.

本文所建POMDP模型的状态是指现场层的安全状态,对于安全状态的表示没有固定标准,在工业系统中通常用设备的被攻击状态表示[21],本文用现场设备是否被攻击作为系统的安全状态.

2) A为行动空间,$A=(a_1,a_2,\dots,a_n)$为Agent能够采取的行动集合,采取行动会影响系统的状态和观测,根据历史选择正确的行动是POMDP模型的核心问题. 本文的行动空间由防御者可采取的防御策略组成,即攻击防御树中防御节点.

3) Z为观测空间,$Z=(z_1,z_2,\dots,z_k)$为可得到的观测集合. Agent无法直接得到系统状态,但是可通过实时观测反映状态.

本文的观测空间由系统可检测到的异常信息(过程变量、操作变量异常)组成,由于现场设备遭到攻击时,被攻击的设备与其关联设备的运行数据均可能会出现异常,很难判断出具体的被攻击设备,但是可通过实时的异常信息更新系统安全状态的信念分布,从而计算最优防御.

4) T为转移概率函数,为状态到状态的转移概率,如下式所示:

$$T(s,a,s')=P(s'|s,a), \tag{10}$$

表示Agent在状态为s时,采取行动a后状态转移到s'的概率.

安全状态转移受到攻击者攻击行为和防御者防御行为的影响. 在现场控制系统中,外来攻击者大多是利用现场设备的漏洞发起攻击,因此在不采取防御策略时,安全状态的转移概率可用漏洞利用率表示. 漏洞利用率使用CVSS中的漏洞可利用性指标计算,该指标包括攻击途径(attack vector, AV)、攻击复杂度(attack complexity, AC)、权限要求(privilege required, PR)和用户交互(user interaction, UI),具体值可从开源的数据库中找到,漏洞利用率计算公式如下所示:

$$p(v)=\frac{1}{2}\times \mathrm{AV}\times \mathrm{AC}\times \mathrm{PR}\times \mathrm{UI}. \tag{11}$$

假设在t时刻系统安全状态为s_t,被攻击设备为e_i,$t+1$时刻采取的防御措施为a,系统安全状态为s_{t+1},被攻击设备为e_j,则状态转移概率$T(s_t,a,s_{t+1})$如下式所示:

$$T=\begin{cases}(1-\varepsilon(a))\times p(v_{ij}),\ s_{t+1}(e_i)=1,s_{t+1}(e_j)=1;\\ \varepsilon(a),\ s_t(e_i)=1,s_{t+1}(e_i)=0;\\ 0,\ \text{otherwise.}\end{cases} \tag{12}$$

其中$\varepsilon(a)$为防御策略的成功率,其值在[0, 1]区间. 当防御策略不能防御被攻击的设备e_i时,状态转移概率为攻击转到设备e_j的漏洞利用率$p(v_{ij})$乘以防御失败的概率;若防御策略恰好可防御被攻击的设备e_i时,则状态转移概率为防御成功的概率$\varepsilon(a)$.

5) O为观测概率函数,观测与状态间的关系通过观测概率体现,即

$$O(s',a,z)=P(z\mid a,s'), \tag{13}$$

表示Agent采取行动a后到达状态s'观测到z的概率,可通过分析系统历史运行数据得到.

6) R为即时收益,$R(s,a)$为状态处于s时,Agent采取行动a后获得的即时收益. 本文的即时收益为防御者采取防御策略后获得的收益值,通过上文中防御收益量化步骤得到.

7) γ为折扣因子,$0<\gamma<1$,用来减少未来收益对总收益的影响,根据实际情况选择即可.

2.2.2 安全策略决策

安全策略决策是在构建好POMDP模型后,利用实时观测到的现场层异常信息对模型进行求解得到最优防御策略的过程.

POMDP模型是马尔可夫决策(Markov decision processes, MDP)模型的扩展. MDP模型是在对系统状态完全已知的情况下建模,其求解过程为状态到动作的映射. 而POMDP模型是在对系统状态不确定的情况下建模,Agent只能从环境中获得观测信息作为状态的参照,所以它必须根据所有的观测和执行动作的历史序列$(a_0,z_1,\dots,a_{t-1},z_t)$来决策下一个动作$a_t$. 随着时间的推移,这种历史序列会变得很长,Åström[22]提出可通过信念分布进行总结,b为一个代表状态上概率分布的向量,如下式所示:

$$b_t(s)=P(s_t=s|z_t,a_{t-1},z_{t-1},\dots,a_0), \tag{14}$$

其中t时刻的信念点b_t可根据贝叶斯规则更新,只涉及上一步的信念状态b_{t-1}、采取的动作a_{t-1}和得到的观测z_t,如下式所示:

$$b_t(s') = \tau(b_{t-1}, a_{t-1}, z_t) = \frac{O(s', a_{t-1}, z_t)\sum_s T(s, a_{t-1}, s')b_{t-1}(s)}{P(z_t|b_{t-1}, a_{t-1})}, \tag{15}$$

$$P(z_t|b_{t-1}, a_{t-1}) = \sum_{s'} O(s', a_{t-1}, z_t) \times \sum_s T(s, a_{t-1}, s')b_{t-1}(s). \tag{16}$$

因此, POMDP 模型的求解过程可看作信念状态到动作的映射: $\pi(b) \to a$, b 为信念分布, a 为策略 π 选择的动作. 最优策略可由贝尔曼方程迭代获得, 即

$$Q_{t+1}(b, a) = \sum_s b(s)R(s, a) + \gamma \sum_z P(z|b, a)V_t^*(\tau(b, a, z)), \tag{17}$$

$$V_{t+1}^*(b) = \max_a Q_{t+1}(b, a), \tag{18}$$

$$\pi_{t+1}^*(b) = \arg\max_a Q_{t+1}(b, a). \tag{19}$$

其中: Q 值函数 $Q_{t+1}(b, a)$ 为 t 步视野内在当前信念点 b 执行 a 的收益值, $V_{t+1}^*(b)$ 为 $Q_{t+1}(b, a)$ 选择动作取得的最大收益值, $\pi_{t+1}^*(b)$ 为最大收益值时的最优策略.

由于信念空间是连续的, POMDP 模型求解不能像 MDP 一样直接迭代解出, 为了求解 POMDP 模型, Smallwood 等[23] 表明任何有限视野 t 上的值函数均可用一组向量表示: $\Gamma_t = (\alpha_0, \alpha_1, \ldots, \alpha_m)$, 每个 α 向量为一个 $|S|$ 维超平面, 并定义信念有界区域上的值函数为

$$V_t(b) = \max_{\alpha \in \Gamma_t} \sum_s \alpha(s)b(s). \tag{20}$$

许多近似 POMDP 解决方案利用特定或少数的信念点更新收益值, 通过增加迭代次数保证算法效果来获得计算优势. 本文采用基于点的值迭代求解算法求解 POMDP 模型, 步骤如下.

step 1: $\alpha_b = \arg\max_a \sum_s R^a(s)b(s),\ b \in B$;

step 2: $\Gamma_0 = \bigcup_{b \in B} \alpha_b$ // 计算初始向量 Γ_0;

step 3: policy $= \Gamma_0^a \to a$ // 向量对应的动作 a;

step 4: $t = 0$;

step 5: while $t <$ itermax;

step 6: $\alpha^{a,z}(s) = \gamma \sum_{s'} T(s, a, s')O(s', a, z)\alpha(s')$, 其中 $\alpha(s') \in \Gamma_{t-1}$;

step 7: $\alpha_b = \arg\max_a \left[\sum_s R(s, a)b(s) + \sum_z \max \times \left[\sum_s \alpha^{a,z}(s)b(s)\right]\right],\ b \in B$;

step 8: $\Gamma_t = \bigcup_{b \in B} \alpha_b$ // 更新向量 Γ_t;

step 9: $t = t + 1$;

step 10: end;

step 11: V max $= \max_{\alpha \in \Gamma_t} \sum_s \alpha(s)b(s)$ // 最大收益;

step 12: policy$_{\text{best}} = V$ max $\to$ policy // 最优策略.

3　实验验证和结果分析

本节以简化的田纳西-伊斯曼 (Tennessee Eastman, TE) 过程控制系统[19] 为对象来阐述所提出方法的执行过程, 并通过与 MDP 模型对比验证所提出方法的有效性. 如图 5 所示, 该系统主要由监控层、通信网络、现场控制系统和边缘节点组成. 其中: 监控层包括工程师站 ES 和人机界面 HMI, 负责监控现场控制系统的生产运行情况和接收上层信息对现场控制系统进行配置; 边缘节点部署安全策略决策模块; 现场控制系统包括控制器 (成分控制器 CC、流量控制器 FC、压力控制器 PC) 和 TE 物理系统, TE 物理系统包括反应釜、传感器 (成分传感器 IS、流量传感器 FS、压力传感器 PS)、执行器 (阀门 V_1、V_2、V_3) 以及管道等. 简化的 TE 过程主要模拟反应釜的化学反应过程: 化学物 A 和 C 在 B 的作用下反应生成产品 D. 该过程的控制目标是产品生产速度维持在 100 kmol / h, 反应釜压力维持在 2 700 kPa, 物料 A 浓度维持在 47 mol %, 共涵盖 3 个控制闭环, 分别控制釜内压强、产物流量和反应物纯度, 其中压强的安全阈值为 3 000 kPa[24].

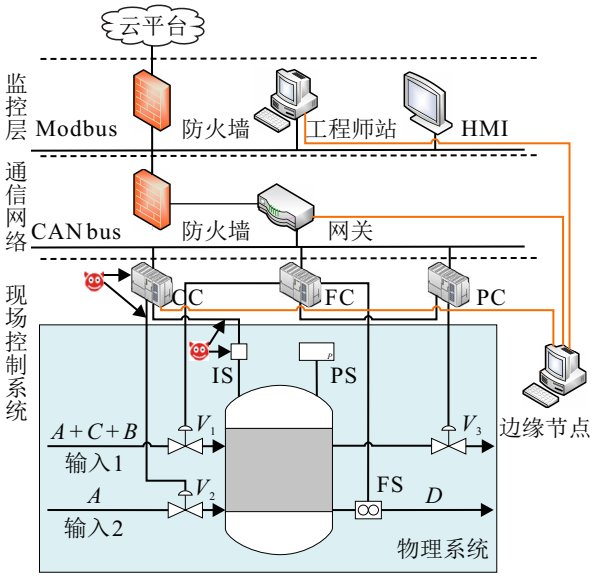

图 5　简化的 TE 过程控制系统

如图 5 所示, 当外部的网络攻击已渗透至 TE 现场控制系统时, 为破坏反应釜运行过程, 攻击者通常会利用漏洞对控制器或传感器进行拒绝服务攻击 (DOS 攻击) 和完整性攻击. DOS 攻击会导致数据传输中断, 这种情况下设备会沿用前一时刻的数据作为

当前数据,该攻击会扰乱反应釜的正常生产但是危害较小. 完整性攻击是对传输的数据进行篡改,为使得伤害最大,攻击者常采用最大最小值攻击,该攻击可能会造成反应釜压力过大发生爆炸,对系统的危害较大. 当现场设备遭受攻击时,由于设备间的耦合性会导致多个运行数据发生异常,很难准确判断被攻击的设备. 但是设备间的互相影响存在一定的滞后性和范围性,本文通过构建POMDP模型利用实时产生的异常数据计算被攻击设备的信念概率,再依据信念概率决策出最优安全策略,具体的安全策略决策过程如后文所述.

3.1 对TE过程控制系统构建攻击防御树

首先对TE过程控制系统构建攻击防御树,根据上文的分析: 目标层为破坏反应釜; 中间层为达到目标层的要素,通过攻击控制回路设备(控制器和传感器)均可造成生产数据异常从而破坏反应釜; 攻击层为攻击事件,对控制回路设备的攻击包括DOS攻击和完整性攻击; 防御层为针对攻击事件可采取的防御策略. 构建的攻击防御树如图6所示,各节点含义如表5所示.

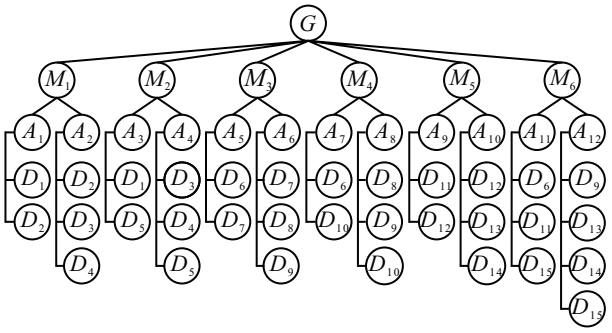

图 6 攻击防御树

表 5 攻击防御树节点含义

标识	含义	标识	含义
G	破坏反应釜	A_{11}	对PS发起DOS攻击
M_1	攻击CC	A_{12}	对PS发起完整性攻击
M_2	攻击IS	D_1	关闭CC和GW的通信
M_3	攻击FC	D_2	CC控制指令校验
M_4	攻击FS	D_3	CC和GW通信加密
M_5	攻击PC	D_4	重启CC
M_6	攻击PS	D_5	启用IS备份传感器
A_1	对CC发起DOS攻击	D_6	关闭FC和GW的通信
A_2	对CC发起完整性攻击	D_7	FC控制指令校验
A_3	对IS发起完整性攻击	D_8	FC和GW通信加密
A_4	对IS发起完整性攻击	D_9	重启FC
A_5	对FC发起DOS攻击	D_{10}	启用FS备用传感器
A_6	对FC发起完整性攻击	D_{11}	关闭PC和GW的通信
A_7	对FS发起DOS攻击	D_{12}	PC控制指令校验
A_8	对FS发起完整性攻击	D_{13}	PC和GW通信加密
A_9	对PC发起DOS攻击	D_{14}	重启PC
A_{10}	对PC发起完整性攻击	D_{15}	启用PS备用传感器

按照第2.1节的步骤计算防御节点收益值. 首先在攻击层利用FAHP计算攻击节点风险值,根据表3攻击属性等级评分标准对攻击节点打分,再通过式(7)计算得到攻击节点风险值,结果如图7所示.

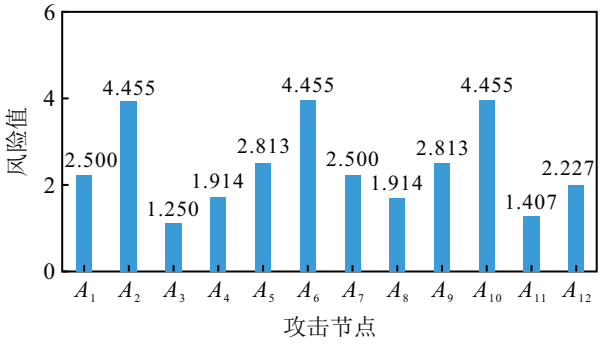

图 7 各攻击节点风险值

然后在防御层利用FAHP计算防御节点收益值,根据表4防御属性等级评分标准对防御节点打分,再通过式(9)计算得到防御节点成本值,最后利用式(8)计算得到防御节点收益,如图8所示.

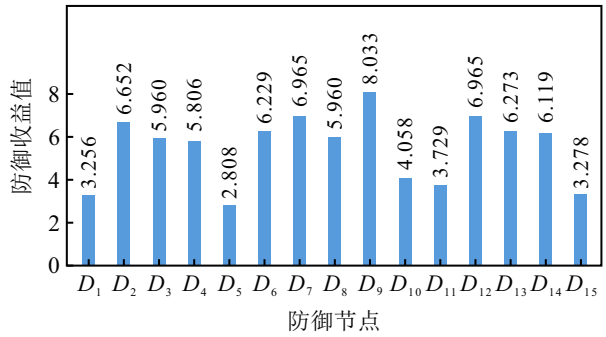

图 8 各防御节点收益值

3.2 基于POMDP的安全策略决策

构建POMDP模型对TE过程现场控制系统进行实时的安全策略决策. 选择控制系统设备被攻击构成状态空间,其中包括正常运行状态,共有7个系统状态. 通常不能直接判断系统的安全状态,可通过观测状态推理,选择告警信息构成观测状态空间,其中包括未观察到异常,共有7个观测. 系统状态及其可观测到的告警信息如表6所示. 行动空间包括

表 6 状态空间及其对应的观测

标度	含义
s_1: CC被攻击	z_1: A料(输入2)异常、z_7: 未观察到异常
s_2: IS被攻击	z_1: A料(输入2)异常、z_2: A料(输入1)异常、z_3: B料异常、o_4: C料异常、o_7: 未观察到异常
s_3: FC被攻击	z_2: A料(输入1)异常、z_3: B料异常、z_4: C料异常、z_7: 未观察到异常
s_4: FS被攻击	z_5: 产品D异常、z_7: 未观察到异常
s_5: PC被攻击	z_6: 压力异常、z_7: 未观察到异常
s_6: PS被攻击	z_2: A料(输入1)异常、z_3: B料异常、z_4: C料异常、z_6: 压力异常、z_7: 未观察到异常
s_7: 正常运行状态	z_7: 未观察到异常

第3.1节中的防御策略($D_1 \sim D_{15}$)和不采取防御策略(D_{16}),共16个行动. 即时收益为第3.1节中的防御收益,其中不采取防御策略的收益值为0,当防御策略不能抵御攻击时,其收益为防御成本的负值.

马尔可夫决策过程(MDP)模型是在系统当前安全状态完全已知的情况下作出决策,因此可用MDP决策的结果验证所提出方法的有效性.

实验设定了多种攻击场景,针对每种攻击场景系统检测到的观测状态,利用POMDP模型作出实时动态安全策略决策,将得到的最优防御策略及其对应的收益值,与安全状态已知情况下MDP决策的结果进行对比,实验结果如下所示.

图9为成分传感器IS受到攻击时,POMDP模型和MDP模型的决策结果,图中展示了迭代200次时,在2种决策模型下16个防御策略的收益值,其中收益值最大的即为最优策略. 由图9可见,2种决策模型的防御策略收益值相近,且最优策略均为防御策略D_4. 图10为不同攻击场景下2种决策模型的决策结果对比. 由图10可见,在部分观测状态下POMDP模型的最优防御策略,与在安全状态完全已知下MDP模型的最优防御策略完全一致,即验证了POMDP模型安全决策的有效性.

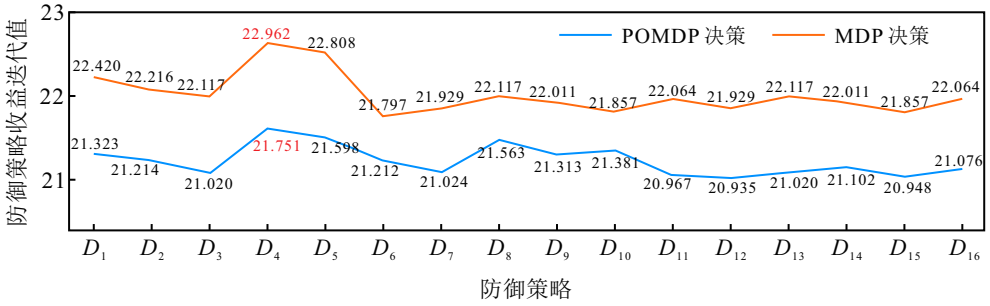

图 9　成分传感器IS受到攻击时2种模型对比

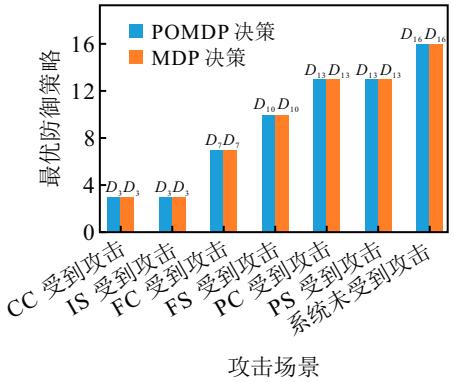

图 10　不同攻击场景下2种模型决策结果对比

图11和图12为不同攻击场景下2种决策模型的防御策略收益值迭代曲线以及决策结果. 由图11和图12可见,POMDP模型决策出的最优策略与MDP模型一致,且POMDP模型迭代的收益值不断向MDP模型趋近. 可得出:在对系统安全状态不完全已知的情况下,POMDP模型利用部分观测状态,以牺牲较少的收益值可决策出最优安全策略.

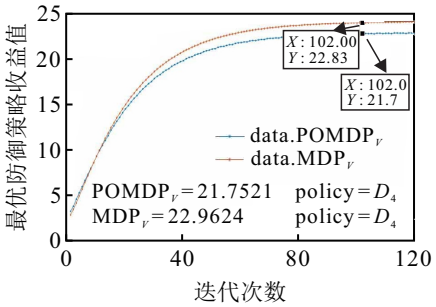

图 11　成分传感器IS受到攻击时2种模型的迭代收益值

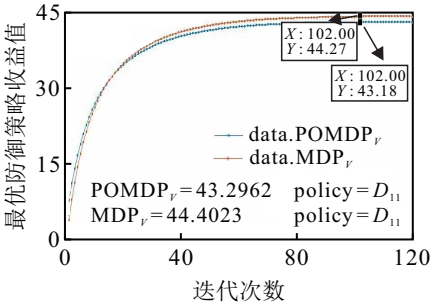

图 12　流量传感器FS受到攻击时2种模型的迭代收益

4　结　论

本文首先对云环境下工业信息物理系统现场层安全防护需求进行分析,针对现场层安全防护需求高但是缺乏适用于新型架构的安全策略决策相关研究的现状,提出了部署于边缘计算节点的现场层区域性安全策略决策框架. 利用边缘节点的就近优势实现现场层的快速安全决策,通过对工业现场划分区域,减小决策规模来减少边缘设备计算消耗;然后,针对现场区域构建攻击防御树,在攻击层和防护层引入模糊层次分析法,综合考虑防御因素和攻击因素,更合理地评估防御策略的防护能力;最后,再对现场区域构建POMDP模型,利用实时的异常数据求解出最优安全策略,实现了系统安全状态不完全已知下的最优

安全策略决策. 本文考虑到边缘节点的资源约束提出在边缘计算节点上实现现场区域性最优安全决策, 但不是全局最优, 后续的研究工作中需要结合云端的资源优势与全局优势, 通过云边结合的形式以实现工业信息物理系统的全局快速安全策略决策.

参考文献(References)

[1] Colombo A W, Karnouskos S, Kaynak O, et al. Industrial cyberphysical systems: A backbone of the fourth industrial revolution[J]. IEEE Industrial Electronics Magazine, 2017, 11(1): 6-16.

[2] 中国信息物理系统发展论坛. 信息物理系统白皮书[EB/OL]. (2017-03-01)[2022-1-20]. http://www.cesi.cn/201703/2251.html.

[3] 工业互联网产业联盟(AII). 工业云安全防护参考方案[EB/OL]. (2017-04-01)[2022-1-20]. http://www.aii-alliance.org/index/c145/n99.html.

[4] Chen J, Su C, Yeh K H, et al. Special issue on advanced persistent threat[J]. Future Generation Computer Systems, 2018, 79: 243-246.

[5] Denning D E. Stuxnet: What has changed?[J]. Future Internet, 2012, 4(3): 672-687.

[6] Baker T, Mackay M, Shaheed A, et al. Security-oriented cloud platform for SOA-based SCADA[C]. The 15th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing. Shenzhen, 2015: 961-970.

[7] Zhou C J, Hu B W, Shi Y, et al. A unified architectural approach for cyberattack-resilient industrial control systems[J]. Proceedings of the IEEE, 2021, 109(4): 517-541.

[8] Xing H, Zhou C J, Ye X H, et al. An edge-cloud synergy integrated security decision-making method for industrial cyber-physical systems[C]. IEEE the 9th Data Driven Control and Learning Systems Conference. Liuzhou, 2020: 989-995.

[9] Fessi B A, Benabdallah S, Boudriga N, et al. A multi-attribute decision model for intrusion response system[J]. Information Sciences, 2014, 270: 237-254.

[10] Li X, Zhou C J, Tian Y C, et al. A dynamic decision-making approach for intrusion response in industrial control systems[J]. IEEE Transactions on Industrial Informatics, 2019, 15(5): 2544-2554.

[11] Hao J Y, Kang E, Sun J, et al. An adaptive Markov strategy for defending smart grid false data injection from malicious attackers[J]. IEEE Transactions on Smart Grid, 2018, 9(4): 2398-2408.

[12] Hu Z S, Zhu M H, Liu P. Adaptive cyber defense against multi-stage attacks using learning-based POMDP[J]. ACM Transactions on Privacy and Security, 2021, 24(1): 1-25.

[13] Miehling E, Rasouli M, Teneketzis D. A POMDP approach to the dynamic defense of large-scale cyber networks[J]. IEEE Transactions on Information Forensics and Security, 2018, 13(10): 2490-2505.

[14] 工业互联网产业联盟. 工业互联网平台白皮书(2019)[EB/OL]. (2019-06-05)[2022-1-20]. http://www.aii-alliance.org/index/c145/n63.html.

[15] Schneier B. Attack trees[J]. Dr. Dobb's Journal, 1999, 24(12): 21-29.

[16] Kordy B, Mauw S, Radomirović S, et al. Foundations of attack-defense trees[C]. International Workshop on Formal Aspects in Security and Trust. Heidelberg, 2010: 80-95.

[17] 张吉军. 模糊层次分析法(FAHP)[J]. 模糊系统与数学, 2000, 14(2): 80-88.
(Zhang J J. Fuzzy analytical hierarchy process[J]. Fuzzy Systems and Mathematics, 2000, 14(2): 80-88.)

[18] Stouffer K, Falco J, Scarfone K. Guide to industrial control systems security[J]. NIST Special Publication, 2011, 800(82): 16-16.

[19] 黄慧萍, 肖世德, 孟祥印. 基于攻击树的工业控制系统信息安全风险评估[J]. 计算机应用研究, 2015, 32(10): 3022-3025.
(Huang H P, Xiao S D, Meng X Y. Attack tree-based method for assessing cyber security risk of industrial control system[J]. Application Research of Computers, 2015, 32(10): 3022-3025.)

[20] Pineau J, Gordon G, Thrun S. Point-based approximations for fast POMDP solving[R]. Montreal: Technical Report, SOCS-TR-2005.4, School of Computer Science, McGill University, 2005: 1-45.

[21] Huang K X, Zhou C J, Qin Y Q, et al. A game-theoretic approach to cross-layer security decision-making in industrial cyber-physical systems[J]. IEEE Transactions on Industrial Electronics, 2020, 67(3): 2371-2379.

[22] Åström K. Optimal control of Markov processes with incomplete state information[J]. Journal of Mathematical Analysis and Applications, 1965, 10: 174-205.

[23] Smallwood R D, Sondik E J. The optimal control of partially observable Markov processes over a finite horizon[J]. Operations Research, 1973, 21(5): 1071-1088.

[24] Ricker N L. Model predictive control of a continuous, nonlinear, two-phase reactor[J]. Journal of Process Control, 1993, 3(2): 109-123.

作者简介

朱美潘(1997−), 女, 工程师, 硕士, 从事工业互联网安全防护技术的研究, E-mail: meipanzhu@hust.edu.cn;

杨健晖(1998−), 男, 硕士生, 从事云计算环境下工控系统安全防护技术的研究, E-mail: yangjianhui@hust.edu.cn;

李欣格(1996−), 女, 博士生, 从事工业控制系统信息安全、人工智能等研究, E-mail: xingeli@hust.edu.cn;

杜鑫(1996−), 男, 博士生, 从事工控系统异常检测技术、安全控制技术、数字孪生技术等研究, E-mail: xdhust@hust.edu.cn;

周纯杰(1965−), 男, 教授, 博士生导师, 从事工业互联网及工业信息物理系统安全等研究, E-mail: cjiezhou@hust.edu.cn.