关键基础设施人机物协同的安全对抗模型

朱培栋[1], 康文杰[2], 刘 亮[1], 张 瑞[1], 荀 鹏[3]

[1]长沙学院 电子信息与电气工程学院 长沙 中国 410022
[2]湖南警察学院 信息技术系 长沙 中国 410138
[3]国防科技大学 计算机学院 长沙 中国 410073

摘要 智能电网、大型工业系统等国家关键基础设施是大规模人机物融合网络, 安全威胁来自信息域、物理域和社会域, 系统安全性的实现需要多域综合的智能对抗。本文针对攻击者侦察、入侵和破坏等环节的高级威胁, 研究人机物融合的面向可观性与可控性的智能对抗模型和新机制。首先基于关键基础设施对物理子系统的感知与控制功能来刻画人机物融合模型, 然后描述宏观的复杂网络模型和微观的实体关系模型; 刻画的安全威胁模型描述了攻击者如何利用人机物的关联特性和脆弱性, 来实施跨域渗透攻击与多域协同攻击。为应对攻击者侦察阶段的多域目标探测和跨域渗透, 引入人机物多域协同的移动目标防御, 提出多层协同的动态反渗透博弈对抗模型; 为应对多域协同入侵行为的隐蔽性, 设计了人机物多通路完整性监测框架, 引入人机物多域关联的异常检测方法;为应对多域协同攻击, 设计信息物理联动的安全机制和信息安全感知的控制算法, 通过对操作员行为不确定性的调控和认知非理性的修正实现人机物多域联动的安全增强。我们提出完整的以人为中心的安全对抗模型, 通过引入社会域以人为中心的移动目标防御, 减少人作为攻击入口的风险; 通过引入以人为中心的网络行为监测, 实现环路观人的监测结构; 通过消减人在认知过程中感知、注意、记忆、学习、决策等环节的脆弱性, 来增强人在环路的安全控制能力。本文提出的模型将有助于丰富网络空间安全基础模型和发展人机物融合的计算范式, 增强关键基础设施安全性。

关键词 关键基础设施; 人机物融合; 移动目标防御; 入侵检测; 安全博弈
中图法分类号 TP309.1 DOI 号 10.19363/J.cnki.cn10-1380/tn.2025.07.09

Human-Cyber-Physical Collaborative Countermeasure Models for Critical Infrastructure Security

ZHU Peidong[1], KANG Wenjie[2], LIU Liang[1], ZHANG Rui[1], XUN Peng[3]

[1] School of Electronic Information and Electrical Engineering, Changsha University, Changsha 410022, China
[2] Information technology Department, Hunan Police Academy, Changsha 410138, China
[3]College of Computer, National University of Defense Technology, Changsha 410073, China

Abstract Critical infrastructures such as smart grids and large industrial systems are large-scale human-cyber-physical networks, with security threats coming from the cyber domain, physical domain, and social domain. The implementation of system safety requires comprehensive intelligent countermeasures from multiple domains. This paper focuses on advanced threats throughout attacker reconnaissance, invasion, and destruction, and studies intelligent human-cyber-physical countermeasure models and new mechanisms for observability and controllability. We first characterize the human-cyber-physical model based on the monitoring and control functions of critical infrastructure on physical subsystems, and then model it from the perspectives of macro complex networks and micro entity relationships; the proposed security threat model describes how attackers can fully utilize the correlation and vulnerability of human, cyber and physical elements to implement cross-domain infiltration attacks and multi-domain collaborative attack. To cope with the multi-domain target exploration and cross-domain penetration, a multi-domain collaborative mobile target defense model is introduced, and a multi-layer collaborative dynamic game model against penetration is proposed; to address the stealth of multi-domain collaborative intrusion behavior, a human-cyber-physical multi-pathway framework for integrity monitoring is designed, and an anomaly detection method based on human-cyber-physical multi-domain correlation is introduced; to deal with multi-domain collaborative attacks, security mechanisms for cyber-physical collaboration and security-aware control algorithms are to be designed, and by regulating the uncertainty of operator behavior and correcting cognitive irrationality, the security enhancement of human-cyber-physical coordination is achieved. The proposed human-centered security countermeasure model is relatively complete, which reduces the risk of humans as attack entry points by introducing human-centered mobile target defense in the social domain; by introducing a human-centered net-

通讯作者: 康文杰, 博士, 讲师, Email: kangwenjie@nudt.edu.cn。
本课题得到国家自然科学基金(No. 61572514), 湖南省自然科学基金(No. 2023JJ30085, No. 2023JJ40272), 湖南省教育厅项目(No. 22A0599, No. 20A511, No. 22B0938)资助。
收稿日期: 2023-11-03; 修改日期: 2024-01-21; 定稿日期: 2025-06-23

work behavior monitoring approach, a monitoring structure for observing people on the loop is achieved; by reducing the fragility of perception, attention, memory, learning, decision-making in the cognitive process, we can enhance our safety control ability in the loop. These models will help strengthen the fundamental framework of cybersecurity and develop novel paradigms for human-cyber-physical computation, and boost the security of critical infrastructure.

Key words critical infrastructure; human-cyber-physical convergence; mobile target defense; intrusion detection; security game

1 引言

交通系统、电力网络、现代工厂以及重要军事设施等国家关键基础设施(Critical Infrastructure, CI), 随着信息化、网络化和智能化进程的推进, 通过透彻的感知和广泛的互联实现了物理域和信息域的融合, 通过与人类社会的高度交互实现了社会域的渗透, 这些设施正在逐渐形成人机物多域融合的复杂网络。例如, 工业互联网实现人(生产操作员等)、信息系统、物理系统(生产设备等)的互联[1]; 智能电网实现物理系统的信息化, 人作为生产的操作者、产品的消费者已嵌入到电力大闭环系统运行, 形成能源领域的人机物融合系统(Human-Cyber-Physical System, HCPS)[2]。

人机物融合系统可以看作是由人、机、物作为子系统融合而成的一类新的系统。社会子系统、物理子系统基于以计算机技术为核心的信息子系统进行连接与耦合, 通过社会空间、信息空间和物理空间的渗透与关联, 以及社会过程、信息过程和物理过程的交互与融合, 形成一个实现特定目标的完整系统[3]。

近年来关键基础设施不断成为网络空间安全攻击的新目标, 不但扰乱基础设施的运营, 影响向社会提供关键服务, 甚至会造成重大的人员和财产损失[4]。对基础设施的攻击大都充分利用了社会、信息、物理多域的脆弱性, 攻击路径的构建利用了多域的关联关系[5]。例如, 如图 1 所示, 在侦察阶段, 社会域利用基于木马邮件的鱼叉攻击选择基础设施的操作管理人员作为首要目标; 在渗透/入侵阶段, 通过员工及其携带存储载体的移动, 由外网进入内网; 在破坏阶段, 阻断人与物理系统的接口, 例如 StuxNet 在大屏幕回放系统状态, BlackEnergy 对客服中心 DoS 攻击, Industroyer 禁用监控设备; 在恢复阶段, 抑制信息系统的恢复能力和利用操作员的手动恢复缺陷, 例如 Industroyer 预期操作员会在没有继电保护的情况下手工恢复从而产生更大物理破坏。另外, 直接面向基础设施系统操作员认知能力的一类信息拒绝服务(IDoS)攻击[6], 通过产生大量虚假报警来消耗操作员的精力, 使其无暇识别其中隐藏的真实攻击。

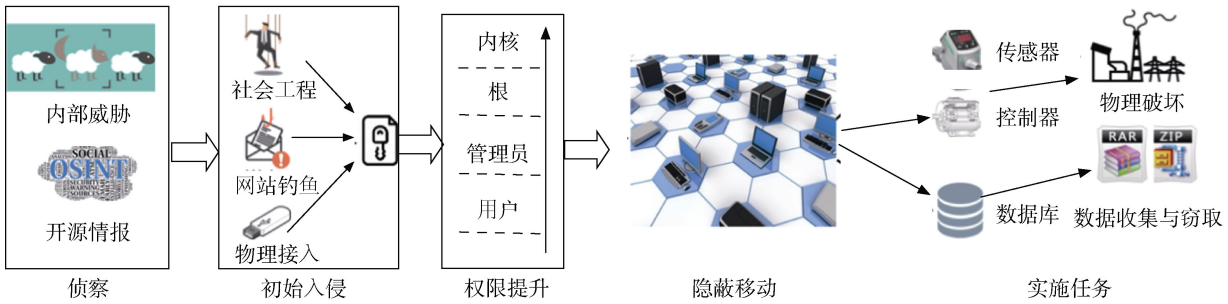


图 1 震网蠕虫等多域渗透攻击过程[5]

Figure 1 Cross-domain penetrating of StuxNet[5]

现有防御技术缺乏人机物充分融合的智能对抗。攻击识别和检测由早期单独信息域的入侵检测, 逐步向基于物理的异常检测和信息物理融合的检测发展, 但是很少考虑社会域及其与信息域、物理域的融合关系, 信息物理融合的检测主要在简单的设备和简化的网络中有所实验, 复杂的基于物理知识和工业设施机理的安全检测和高级攻防对抗还刚起步[7]。

为了有效应对攻击者对人机物融合特点的充分利用, 迫切需要构建相应的高智能的体系化对抗。本文研究的目标就是针对攻击链的侦察、入侵、破坏等环节设计多域协同的对抗模型与机制, 从而通过新型主动防御体系的构建力求在对抗中取得优势。

2 关键基础设施人机物融合模型和人机物多域攻击

为了设计有效的防御和对抗模型, 首先要理解关键基础设施的人机物融合的结构与功能模型, 认清来自人机物多域的安全威胁和攻击行为。

2.1 关键基础设施的网络模型

关键基础设施的人机物融合模型可基于多尺度多维度刻画, 例如, 网络流模型刻画网络信息、物质和能量的流动, 系统动力学模型对系统运行过程进行建模, 实体模型则便于对具体节点进行微观考察。下面首先基于关键基础设施对物理子系统的感知与控制功能来刻画人机物融合模型, 然后从宏观的复杂网络模型和微观的实体关系模型的角度进行描述。

2.1.1 基于感知和控制功能的人机物融合模型

智能电网、工业互联网等关键基础设施中, 系统操作员基于信息子系统对物理设施进行感知和控制, 人机物三者的关系可以表示为“主体(人, *H*)←→信息子系统(机, *C*)←→客体(物, *P*)”。

从哲学上看, 事物演化的相互作用过程一般具有物质形态(*M*)和信息形态(*I*)的双重建构的意义, 信息是物质存在方式和状态的自身显示[8]。“人”包括社会中物质的人(*H*1)及其蕴含的信息(*H*2), *H*2 又包括人作为主体对世界的认识(*H*21)和作为客体被反映的状态(*H*22); “机”是人造信息技术系统(*C*1)及其蕴含的信息(*C*2), *C*2 又包括计算机系统承载的数字化信息(*C*21)和其作为客体被反映的状态(*C*22); “物”是自然界的对象(*P*1)及其蕴含的信息(*P*2), 为人、计算机系统之外的物质所蕴含的信息。这种人机物三元系统的二分解构[3]如图 2 所示。

	物质	信息	
人	H1	H21	H22
机	C1	C21	C22
物	P1		P2
		主体信息	客体信息

图 2 人机物三元系统的二分解构[3]

Figure 2 Binary deconstruction of HCPS[3]

文献[9]从认识论的角度定义三种信息, 其中自在的信息是客观间接存在的物质世界的信息过程, 自为的信息是人在精神层面的主观间接存在, 再生信息是概像信息和符号信息等信息的主体创造。类似地, 在万物互联时代, 我们将表达事物本源运动状态和变化方式的本体信息使用 $I1$ 表示, 计算机存储和处理的再生信息使用 $I2$ 表示, 通过主体感知系统形成的认识论信息表示为 $I3$, 那么将人机物融合系统中人通过信息系统感知物理设施的过程表达为 $I1 \rightarrow I2 \rightarrow I3$; 人的意志 $I3$ 通过信息子系统影响物理子系统的状态的过程表示为 $I3 \rightarrow I2 \rightarrow I1$。自在信息 $I1 = P2 + C22 + H22$, 再生信息 $I2 = C21$, 自为信息 $I3 = H21$。人机物融合系统中信息的三个类型的组成如图 3[3]所示, 其中物质 $M=H1+C1+P1$。

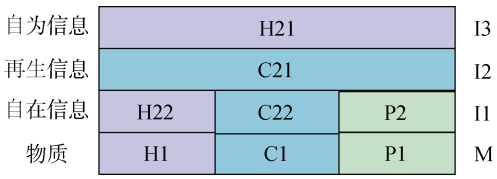


图 3 人机物融合系统中信息的三个类型

Figure 3 Three types of information in HCPS

图 4[3]刻画了人通过信息对物的感知与控制路径, 两条路径都包括人机物三个域的域内节点与跨域节点, 其中 *HH*、*CC*、*PP* 分别表示关键基础设施人、机、物三个域的域内节点, *CH*/*HC*、*PC*/*CP* 表示跨域节点(例如接收终端/控制器、传感器/执行器), 当然也可以有人物直接接口节点 *HP*/*PH*。

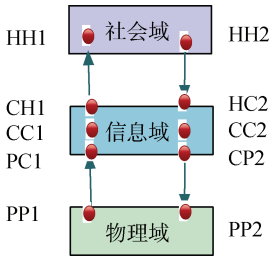


图 4 感知与控制路径

Figure 4 Sensing and control paths

2.1.2 基于复杂网络的宏观模型

宏观网络模型将关键基础设施看作人、机、物多种网络组成的新型的人机物融合网络。如图 5 所示, 物理域的网络是网络化的物理设施, 例如单纯的电力网络; 信息域的网络包括外部的互联网、内部的企业信息网络或工业控制网络; 社会域的网络包括网络化的企业组织结构、有组织的攻防双方力量等人组成的网络。社会域的人通过信息网络或在现场直接感知和操作物理域的设施。

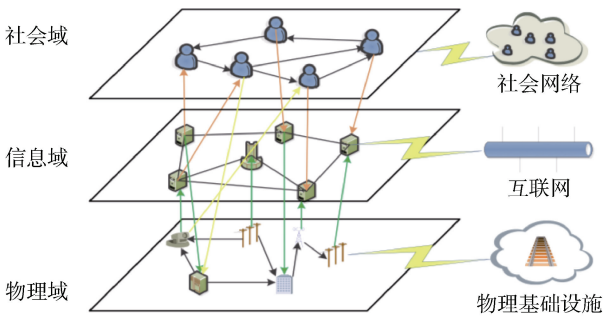


图 5 关键基础设施宏观结构模型

Figure 5 Macro structural model of CI

不少学者运用复杂网络和网络科学的方法, 研究融合后网络脆弱性的变化[10]和网络健壮性的增强方法[11]。多层网络的信息、物质或能量的传递, 包括各个网络内部的扩散和跨域的渗透; 跨域渗透通过域间接口实现。跨越不同域时, 人的意识可以转化为信息网络的数据, 进一步可能促进能源的流动或物质的物理化学变化。

图 5 所示的复杂网络中, 包含社会(人)、信息(机)、物理(物)网络三种异质网络, 分别表示为 $G_h=(V_h,E_h)$, $G_c=(V_c,E_c)$, $G_p=(G_p,G_p)$, 这里 V_h、V_c 和 V_p 分别表示人、机、物网络中的节点集, E_h、E_c 和 E_p 分别表示人、机、物网络的边集。不同类型网络的节点存在相连边, 用以刻画节点之间的映射或关联关系, 例如 R(V_h, V_c)表示人与信息网络节点的关系, 包括人向信息域的映射关系、人对信息节点操作和基于信息节点进行感知等。跨域节点形成的边, 包括 E_{hc} = R(V_h , V_c)={($V_h \to V_c$), ($V_c \to V_h$)}, E_{cp} = R(V_c , V_p)={($V_c \to V_p$), ($V_p \to V_c$)}, E_{hp} = R(V_h , V_p)={($V_h \to V_p$), ($V_p \to V_h$)}。因此人机物融合网络可表示为 G_{hcp}=(V_h, V_c, V_p, E_h, E_c, E_p, E_{hc}, E_{cp}, E_{hp})。

2.1.3 基于多重边实体关系的微观模型

微观模型将人机物融合网络看作多个人机物融合节点组成的多重边网络, 节点之间存在多类连接, 以揭示节点级的多域交互与融合关系, 着重节点实体的具体实现。例如, 智能电网中的智能电表, 与电网一次系统(基本电力网络)有接口, 感知电流、电量; 与电网二次系统(监控网络)有接口, 将感知的电力信息传递给数据集中器, 是信息网络的一部分; 同时有电力公司的员工操控和管理电表, 或者有攻击者破坏电表的工作, 其又与社会域的人密切关联。如图 6 所示, 每个节点在物理域、信息域、社会域都有逻辑存在, 分属于社会网络、信息网络、物理网络三层网络。

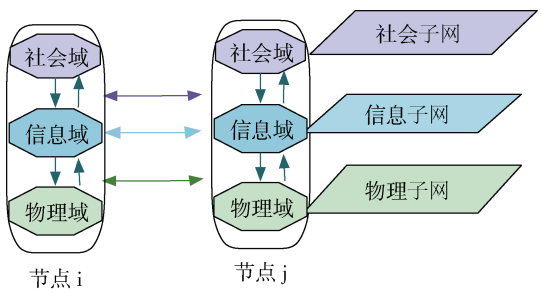


图 6 基于多重网络的微观模型

Figure 6 Micro model based on multiplex network

图 6 中, 每个节点都包含人机物三个域, 人机物融合网络可表示为 $G_{hcp}=\left(V_{hcp},E_{hcp}\right)$, 节点之间同时存在人机物三种类型的关系, 即节点之间的连接是多重连接, 从而形成信息、物理、社会三个逻辑子网。

2.2 跨域渗透攻击与多域协同攻击模型

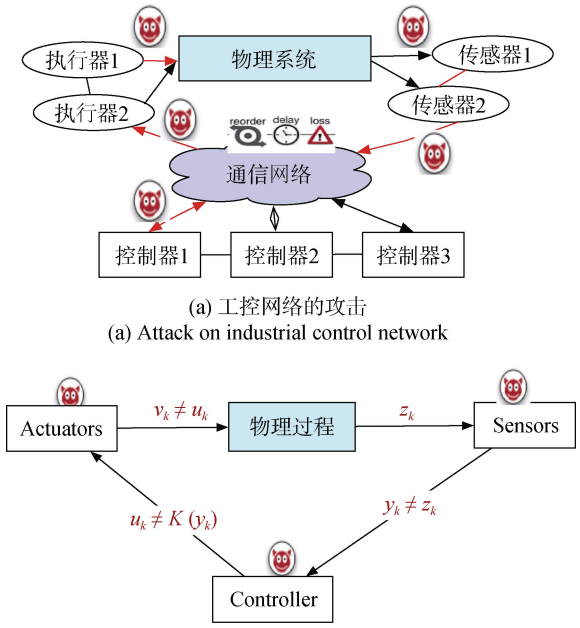
攻击者可以基于上述的多种网络模型, 利用基础设施不同子系统之间的相互关联和相互依赖关系, 包括人机物多域结构相关性和行为关联性[12], 来实施渗透攻击和协同攻击。我们分析了信息物理网络不同耦合模式对系统健壮性的影响[13], 在文献[14]揭示了对控制命令直接修改的威胁模式, 在文献[15]刻画了社会域多人协同进行端系统负载修改从而影响电力网络稳定性的威胁模型。

2.2.1 基于可观性和可控性的安全威胁

可观性和可控性[16], 在控制科学领域有严格的定义, 可观性指的是仅使用来自输出的信息(即传感数据)可以估计当前状态(描述系统时域行为的变量), 采用本文 2.1.1 节的模型则可以理解为基于 $I2$ 可以估计 $I1$; 可控性指的是能够通过控制输入使系统从任意初始状态达到确定的状态[17]。从网络空间安全科学的角度考察, 可控性是指“信息和信息系统时刻处于合法所有者或使用者的有效掌握与控制之下”, 而我们从攻防对抗的角度将可控性理解为“自己可以有效控制, 对手不能及时、准确、完整控制”, 可观性认为是“自己可以充分看到, 而对手看不到、看不准或看不及时”。

图 7(a)为工业控制信息物理系统的入侵模型, 攻击者可能对控制器、传感器、执行器及其通信网络进行攻击。图 7(b)为攻击引起的状态和测量变化, 其中控制器发出的命令为 u_k, x_k 表征物理系统内部的状态, y_k 为系统传感器的输出, z_k 为系统的实际输出, 执行器的命令到达物理设备为 v_k。若控制器被攻击, 发布的命令不能基于反馈实施准确控制, 则对线性控制系统 $u_k \neq K\left(y_k\right)$; 若执行器被攻击, 则不能有效对物理对象和物理过程施加控制, 即 $v_k \neq u_k$; 传感器被攻击, 则无法度量系统状态, 即 $y_k \neq z_k$。

图 7 主要考察了信息域的安全威胁, 另外还有对社会域、物理域的干扰和破坏。基于关键基础设施感知路径和控制路径的完整的威胁结构如图 8 所示, 其中红色×标记的为可能的入侵和破坏点。传感器(PC1)和执行器(PC2)作为跨越信息域与物理域的设备, 控制器(HC2)作为跨越信息域与社会域的设备。



(a) 工控网络的攻击
(a) Attack on industrial control network

(b) 工控网络攻击引起的状态变化
(b) State change caused by industrial control network attack

图 7 工控网络安全威胁模型
Figure 7 Threat model of industrial control network

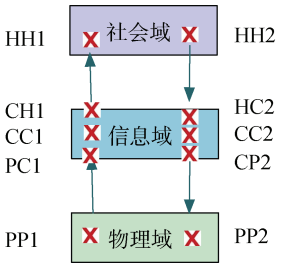


图 8 关键基础设施的入侵破坏点
Figure 8 Intrusion and destruction points of CI

破坏人机物融合系统的感知过程, 即对传感信息的虚假数据注入(False Data Injection-Sensing, 记为 FDI-S)的基本方式是, 将正常的 $I1 \rightarrow I2 \rightarrow I3$ 改为 $I1 \rightarrow I2' \rightarrow I3$, 这是图 8 中在信息域进行的篡改, 篡改了 $I2$ 对 $I1$ 真实的映射关系; $I2$ 的篡改可以发生在传感器、传输网络、接收终端, 即 $I1 \rightarrow I21' \rightarrow I22' \rightarrow I23' \rightarrow I3$; 更进一步, 考虑社会域和物理域的攻击, 则可以直接干扰人的认知状态 $I3$ 和物理系统的内部状态 $I1$。因此, 从完整的人机物融合系统考察对感知路径的侵害, 存在 5 种可能, 即 $I1' \rightarrow I21' \rightarrow I22' \rightarrow I23' \rightarrow I3'$。在这条路径上除了对人机物节点的污染、篡改或干扰, 还包括对节点之间链路的破坏。

类似地, 破坏人机物融合系统的控制过程, 即对控制信息的虚假数据注入(False Data Injection-Control, 记为 **FDI-C**), 也存在 5 种可能, 即 $I3' \rightarrow I21' \rightarrow I22' \rightarrow I23' \rightarrow I1'$, 包括对控制主体(人和人工智能的决策)、对作为通道的信息系统(在控制器、传递路径、执行器)、对物理系统内部状态的破坏。从完整的人机物融合系统考察对控制路径的侵害, 就需要研究 $I3'$ 和 $I1'$。

2.2.2 跨域渗透及其网络扩散

跨域渗透攻击是利用跨域节点的脆弱性实施的攻击, 基于多域依赖关系和安全属性关联, 将脆弱节点基于时空组合形成入侵路径。入侵路径的形成可以基于图 5 宏观结构模型中的边界跨域节点, 也可以通过图 6 的各个多域融合节点。

在人机物融合网络中攻击者对远程物理设备进行操纵或破坏, 从社会域出发, 可以直接通过社会域人际网络到远程节点, 通过信息层网络到远程节点, 也可以通过物理域的网络效应扩散到远程节点, 存在如图 9[3]所示①②③三种不同的客观实在的交互路径, 分别表示为 HHCP、HCCP、HCPP; 也包括 HCCPP 等在多个同类域节点之间扩散的情况。攻击者对关键基础设施攻击的最终目标是对物理子系统功能和性能的破坏, 那么由社会域发起的入侵路径可一般化表示为 $H\{H,C,P\}^nP$, 其中 $n \geq 0$。

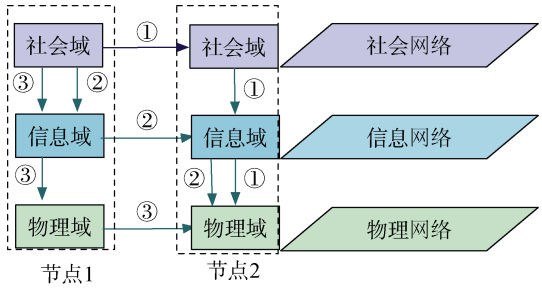


图 9 人和物三种远程交互路径
Figure 9 Remote interaction paths between human and physical objects

在图 10 中, 从人员 h1 经过信息网络对物理节点 p2 的一条攻击路径, 可以表示为 h1 → c11 → c21 → p2。

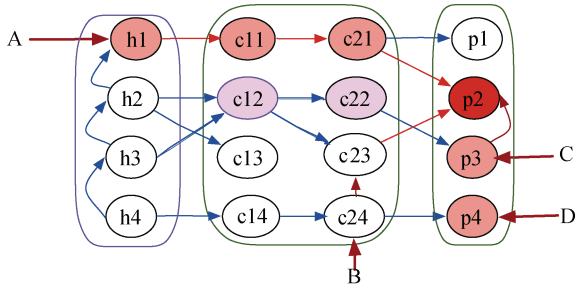


图 10 基础设施攻击模式示意图
Figure 10 Attack mode example for CI

2.2.3 多域协同攻击模型

对基础设施的协同攻击有多种类型, 可以同时对多个或多类节点进行打击, 或者同时采用不同路径对同一目标节点进行打击。例如在图 10 中, 对物理节点 p2 进行攻击可以从 A、B、C 三个入口点同时进行, 攻击点 C 通过物理域内的连锁失效(Cascading failure), 强化了对 p2 的打击效果; p1、p2 可能都由 c21 的失效引发, 存在共因失效(Common cause failure)。文献[18]提出针对智能电网的分布式攻击策略, 攻击者控制多个断路器, 使用多个开关使发电机瞬态失稳, 例如在图 10 中对 p2、p3 和 p4 的同时打击。

智能的协同攻击会在同时或者精确规划的时间间隔发生, 以下是典型的协同攻击行为:

(1) <$X1||X2$>表示同时对 $X1$ 和 $X2$ 进行攻击。

(2) <$X1$-$X2$>表示先对 $X1$ 攻击然后对 $X2$ 攻击。

(3) <$X1\rightarrow X2$>通过 $X1$ 攻击或影响 $X2$, 例如干扰控制信息网络的 QoS 从而造成控制命令的丢失或延迟等。

(4)<$Y|X$>: X 配合 Y 进行协同攻击, 例如<$P_y|C_x$>表示信息物理协同攻击, 对传感信息进行篡改来掩盖对物理域的攻击是其中一种典型形式。

2.3 人机物协同的安全对抗模型

2.3.1 基于可观性和可控性的安全对抗

攻击者对关键基础设施的攻击包括侦察、入侵、破坏三个环节, 前面二个环节是双方可观性的对抗, 第三个环节是可控性的对抗。在侦察阶段, 攻击者对目标系统进行多域探测和渗透(实施 2.2.2 节的威胁模型), 防御者通过多域协同的移动目标防御进行隐蔽、欺骗和阻碍, 双方争夺的是目标系统的可观性; 在入侵阶段, 攻击者使用协同攻击, 通过信息域的篡改掩盖物理域的攻击(实施 2.2.3 节的威胁模型), 防御者则要识别攻击者的欺骗和发现攻击行为, 双方争夺的是对攻击行为的可观性, 对抗的重点是入侵者对目标系统多域特性的综合利用和防御者对入侵行为多域特征的关联。

可控性对抗, 是攻防双方对物理系统控制权的争夺, 攻击者通过信息系统影响控制命令的传递, 通过对社会域操作员人因脆弱性的利用来破坏物理系统的安全; 防御者在信息系统设计时要避免攻击者滥用信息系统的运行机理机制来实施对物理控制的影响, 物理系统运行机制与控制算法的设计要可消解信息安全带来的负面影响, 要能检测社会域的异常并基于信息物理系统弥补或纠正人因脆弱性的影响。

图 11 刻画了上述基于可观性与可控性的人机物对抗模型。攻击者充分利用人机物的关联特性和脆弱性, 运用人机物多域协同的攻击模式; 防御者需要设计人机物多域协同的移动目标防御(Deception, 欺骗)、多域融合的入侵检测(Detection, 检测)、多域联动的安全增强(Deter, 遏制)模型与机制, 形成人机物融合的智能对抗体系。

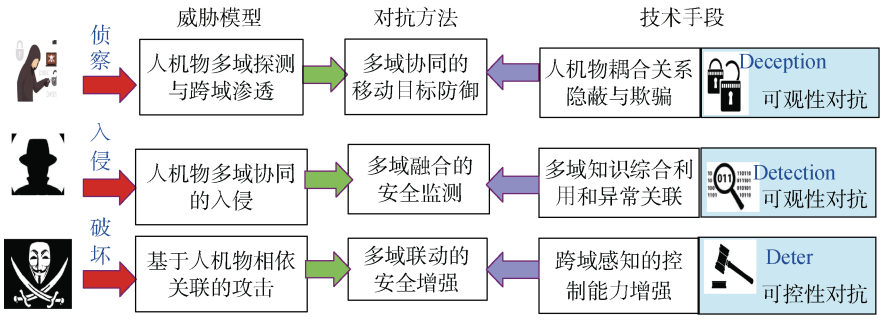


图 11 关键基础设施人机物可观性与可控性对抗模型

Figure 11 Observability and controllability attack-defense model for HCPS

2.3.2 人机物多域的对抗

本文模型的研究, 重点体现在社会域对抗模型的引入, 强化信息物理协同模型, 从而形成人机物融合模型。

如图 12 所示, 主要研究内容和思路如下: (1)为了应对攻击者的侦察和渗透, 在社会域引入以人为中心的移动目标防御, 防止攻击者基于人的脆弱性进入系统(to the loop), 并强化信息物理关联和人机物映射关系的保护, 设计多层协同的动态反渗透博弈对抗方法, 从而实现人机物多域协同的移动目标防御; (2)为了有效识别对基础设施的多种形式的协同攻击, 在社会域引入人在环上(on the loop)的网络空间状态监测, 并突出基于物理机理和工业知识来识别物理系统和物理过程的异常, 实现社会域与信息物理域的异常关联; (3)为了抵抗对物理系统的破坏, 引入社会域操作员行为不确定性的调控和认知非理性的修正, 从而增强人在环内(in the loop)的安全控制能力, 通过物理安全(safety)感知的信息防御

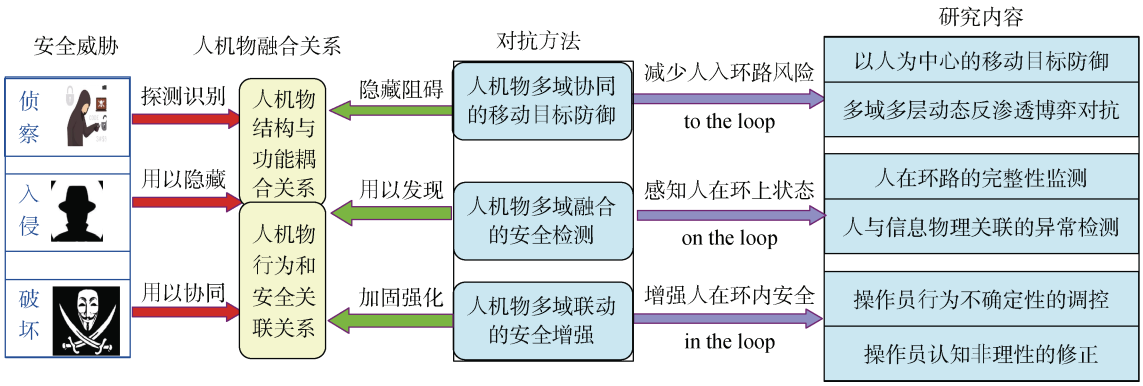


图 12　关键基础设施人机物多域的安全对抗模型

Figure 12　Multiple-domain attack-defense model of HCPS

和信息安全(security)感知的物理系统鲁棒性控制, 实现人机物多域联动的安全增强。

从图 12 可以看到三部分研究内容的关系: (1) 从攻击者来看, 侦察、入侵、破坏是对关键基础设施攻击的三个连续的阶段, 从防御者来看则是针锋相对的三个环节和防御的三个方面, 从而形成完整的攻防对抗体系; (2) 三个环节对抗的核心是“人机物融合关系”, 攻击者探测识别和恶意利用人机物结构功能耦合关系、行为和安全关联, 而防御者要阻碍攻击者探测识别融合关系、发现基于融合关系隐蔽的攻击行为、通过增强人机物融合强度来遏制协同攻击, 从而形成基于人机物融合关系的对抗模型; (3) 三部分内容从社会域及其与信息物理域关联的角度看, 形成了以人为中心的完整的关键基础设施网络空间安全模型, 基于人机物融合关系(of the loop)实现入环、环上、环内三个环节的安全防御机制。

3　人机物多域协同的移动目标防御

移动目标防御(Moving Target Defense, MTD)利用随机切换结构或主动、动态地改变系统参数, 在保持系统完整性和运行正确性的同时, 增加攻击者的知识不确定性, 减少攻击机会窗口。为了应对攻击者对关键基础设施的侦察和渗透[19], 需要在人机物多域采用移动目标防御, 除了传统的信息域 MTD, 还包括: 1) 物理域 MTD; 2) 信息物理协同的 MTD, 通过信息域的变换阻碍攻击者对物理特征的探测与获取; 3) 以人为中心的 MTD, 重点保护人与信息物理系统的耦合关系; 4) 人机物多域多层协同的动态 MTD 博弈对抗, 以有效阻碍跨域渗透攻击和多域协同攻击入侵路径的构建, 从而形成人机物多域协同的移动目标防御体系。不但减少基础设施对敌手的可观性, 并在一定程度上具备对敌手行为的可控性。

基于 2.1.1 节对人机物三元系统的二分解构, 物理域 MTD 影响的是敌手对系统物理状态 $P2$ 的直接感知, 信息物理协同的 MTD 是通过对信息系统作为客体被反映状态 $C22$ 的主动变化(包括通过 $C1$ 结构或参数调整而产生变化), 影响敌手对物理系统状态的间接感知; 以人为中心的 MTD 则通过主动改变人员作为客体被反映状态 $H22$(重点是人员与信息物理系统的耦合关系), 减少人作为攻击入口的可能; 人机物多域多层协同的动态 MTD, 则是综合运用 $H22$、$C22$ 和 $P2$ 的主动变化来影响敌手对 HCPS 系统的认知。

3.1　物理域和信息物理协同的移动目标防御

3.1.1　物理域的移动目标防御

有三种典型的方法来实现物理系统可观映像的多样化、随机化、动态性。一是主动扰乱, 例如, 电力系统采用配电网柔性交流输电系统(D-FACTS), 主动扰动电网的输电线路电抗, 避免攻击者准确探测; 二是采用冗余的或多品牌的传感器、执行器, 在不同厂家的可编程控制器(PLC)、远程终端单元(RTU)、智能电子设备(IED)等之间进行动态切换, 并确保相同的输入和输出; 三是感知到入侵行为后对物理系统参数进行重新配置。工业系统物理域移动目标防御相对于信息域的特殊性, 在于保证功能正确性、可使用性和性能实时性的困难, 以及系统切换的高代价。

3.1.2　信息物理协同的移动目标防御

信息物理协同的移动目标防御, 通过信息域实现对物理域的防护, 不但能够抵抗信息物理协同的目标探测, 而且可以比单纯的物理域移动目标防御具有更小的实现代价。主要采用以下技术途径:

(1) 基于信息域阻碍物理域特征的获取。我们在文献[20]提出自适应转发路径迁移机制, 随机化状态估计所采用的测量集, 使攻击者发动恶意数据注入攻击所需的知识失效, 从而暴露物理攻击的影响

以中断信息物理协同攻击; 我们在文献[21]提出通过多轮随机化攻击者的数据获取过程, 将攻击空间扩展到多个维度, 使攻击者无法通过观测历史测量数据推导出状态估计使用的测量矩阵。

(2) 信息物理接口设备和耦合实体的保护。例如智能电表固件的多样性, 设备运行环境采用指令集随机化和地址空间布局随机化。

(3) 信息物理耦合关系的保护。为了避免被攻击者探知后基于信息物理节点的结构和功能耦合关系发起具有高效连锁效应的攻击, 具体可以使用或拓展信息网络技术的拓扑隐藏和欺骗方法。

(4) 物理系统的数字域存在和信息化映射关系的保护。例如, 在工业设施中对制造执行系统层(MES)蕴含的大量以数字化状态存在的物理系统知识进行防护, 包括对微服务、数据文件等采用随机化、模糊化等 MTD 技术, 或者在存放的系统结构文件中伪造物理系统的拓扑关系等。

3.2 社会域与信息物理域协同的移动目标防御

3.2.1 关键基础设施的人机物多域层次网络结构

人机物融合网络的社会网络、信息网络和物理网络都可以细分为多个层次, 如图 13 所示。

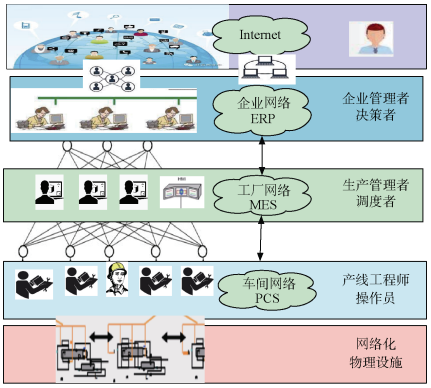


图 13 工业基础设施中的网络结构

Figure 13 Network structure in industrial infrastructure

社会域与信息域、物理域的关联, 主要体现在不同岗位的企业员工对信息资源的拥有和物理资源的掌控关系, 例如, 车间内的操作员与过程控制系统的业务账号、控制网络层 ID / 工位等信息资源的关联, 与控制器、执行器、传感器等设备以及物理系统等的关联。

3.2.2 以人为中心的移动目标防御模型与方法

攻击者可以对基础设施网络中人的关系、人与信息域及物理域的关联关系进行探测, 并基于此形成攻击路径。

基础设施以人为中心的移动目标防御, 即“社会域与信息物理域耦合关系的移动目标防御”, 首先基于在线社交网络等对员工身份进行保护, 例如, 把系统操作员混杂在各类员工中, 在社交网络构建虚假化身; 然后通过模糊化或混淆人与信息物理系统的关联和映射关系进行移动目标防御, 例如, 员工与计算机节点地址的对应关系, 工厂生产网络中生产管理者、生产调度者负责管理的员工集合、负责的信息子网和生产资源等人机物关联关系, 车间控制网络中产线操作员、产线工程师与负责管理的信息网络、操纵的物理系统等对应关系。

通过以上过程, 对社会域及其信息域、物理域的耦合关系进行了模糊化和隐藏, 攻击者的侦察能力和基础设施多域依存关系的利用能力会受到较大程度限制。

3.3 关键基础设施的人机物多域多层协同的移动目标防御

人机物多域协同的移动目标防御, 包括基础设施社会域、信息域、物理域各域多层和多域协同的移动目标防御, 基本结构如图 14 所示。

首先对各域的各层进行基本的移动目标防御配置, 例如, 社会域各层使用“社会域与信息物理域耦合关系的移动目标防御方法”实现以人为中心的移动目标防御; 信息域实现应用层、传输层、网络层、链路层等各层移动目标防御(图 14 中 SDN 为软件定义网络); 物理域实施多层控制结构移动目标防御。

然后, 为了遏制攻击者跨域跨层渗透攻击, 采用多层协同的动态 MTD 博弈对抗, 追求多层防御代价最小化和攻击平面最小化。参照文献[22]等对博弈原理的基本描述, 面向人机物层次化系统动态 MTD 进行拓展。以下方法使用图 4 和图 14 的人机物层次网络模型, 其中相邻层可能属于不同域, 例如通过信息网络最下层可以进入到下面物理系统的最上层。

人机物融合系统多层动态 MTD 博弈方法, 包括基本的动态防御、动态防御中的概率博弈和多层协同的 MTD 三部分。

(1) 基本的动态移动目标防御

在攻击者的渗透入侵过程中, 防御者感知到攻击者利用第 l 层的脆弱性集合, 则动态改变该层的配置; 若新的配置具有的脆弱性不同, 则攻击者无法成功渗透。例如, 如果攻击者针对配置 $c(l,i)$ 对应脆弱性集合 $v(l,i)$ 采取动作 $a(l,i)$, 那么防御者将配置动态调整为 $c(l,j)$。若 $c(l,j)$ 对应脆弱性集合 $v(l,j)$ 与攻击者利用的 $v(l,i)$ 不相同或无交集, 则攻击路径无法继续拓展。

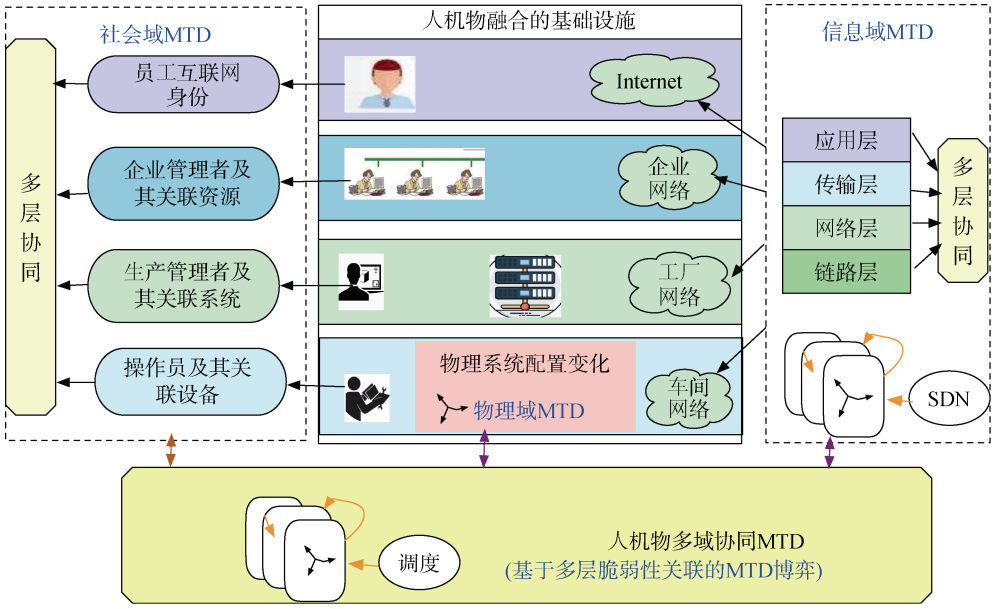


图 14　人机物多域协同的移动目标防御

Figure 14　Human-cyber-physical collaborative MTD

(2) 动态防御中的概率博弈

在实际对抗过程中, 防御者可能不了解攻击者针对的是哪种配置, 攻击者也不知道防御者可能动态调整为哪种配置, 这是一种基于概率的博弈过程。例如, 攻击者从动作集合 A_l 选择 $a(l,k)$ 的概率为 $f(l,k)$, 可能造成的代价或损坏为 $d(l,k)$; 防御者调整配置为 $c(l,j)$ 的概率为 $g(l,k)$, 防御者博弈的目标是最小实现代价, 以及最小化攻击平面, 即 $v(l,k)$ 和 $v(l,j)$ 的交集最小。

(3) 多层协同的 MTD

攻击者构建攻击路径时可能从层次结构的上层或下层到达本层; 防御者在改变当前层配置的时候要考虑前导层的影响, 又要考虑对后续层的影响, 以通过恰当的配置阻止对后续层脆弱性的利用。例如, 对入侵路径 $path =< p_1, p_2, ..., p_k, ... >$, 攻击者可能采取的动作与当前层 p_k 的脆弱性集合、前面 p_{k-1} 和后面 p_{k+1} 的脆弱性集合都有关联, 层 p_k 进行 MTD 要考虑 p_{k-1} 的当前配置与自己的接口及经由该接口带来的脆弱性大小, 考虑 p_{k+1} 哪种配置脆弱性最大从而减少攻击者通往那种配置的可能。

4　人机物多域融合的安全检测

人机物多域融合的安全检测, 主要应对入侵行为在信息域的隐蔽性而在人机物多个域对其进行发现与识别。与传统单纯信息域的安全检测不同, 一是重视物理域状态的感知, 以及信息域和物理域的关联, 二是将人的多模态输入作为安全检测的数据来源并与信息物理系统异常进行关联。

将人的多模态输入用于安全检测, 对关键基础设施操作员的行动、表情、姿态等进行社会域的心理分析与行为辨识, 主要用于二个方面的网络空间异常监测: (1) 基于多维度信息判断基础设施监控关系的完整性, 从而尽早发现操作员对物理系统的监控被攻击者劫持的情况。(2) 内部威胁者进行非法操作时可能表现出紧张、有意遮掩等情态, 基于操作员的心理和行为异常状态与信息物理系统内部异常状态的关联, 通过人机物多域融合实现更准确更完整的安全检测。

4.1　信息物理融合的安全检测

为了识别信息物理协同的隐蔽的入侵和破坏, 检测引擎需要利用更多状态和知识, 包括传感器反映出来的系统内部状态和控制器反映的物理系统模型。信息物理融合的入侵检测, 其目的就是在传感信息(图 3 中的 *I*2)被掩盖时, 利用更多物理域知识来揭示物理域的状态(图 3 中的 *I*1), 并基于不同 *I*1 之间、不同 *I*2 之间、*I*2 与 *I*1 之间的一致性来识别系统异常。

(1) 对不同 *I*1 综合分析: 例如考察能耗、噪声指纹等物理特征的完整性。

(2) 对不同 *I*2 综合分析: 例如基于传感数据之间的相关性[23]或传感网络的熵值是否异动来发现异常变化; 通过对多传感器、多参数、多时间点进行数据空间关系、系统逻辑关系、时序演化模型分析, 判断其蕴含的系统模型与真实模型的一致性或是否偏

离物理系统的常规模型。

(3) 分析 *I*2 和 *I*1 的一致性: 为了获取足够的物理系统状态, 电力系统等配备精细反映物理过程连续状态的电力相位测量单元(PMU)等设备, 从而可以和数据采集与监视控制系统(SCADA)等信息域监控网络的数据进行多模综合分析与关联检测, 判断信息域传递或显示的数据与物理系统的真实状态是否一致。

4.2 “人在环路”完整性监测

关键基础设施运行管理采用的是监控模式(Supervisory control paradigm), 作为监控者的系统操作员在信息物理系统出现异常情况下, 可能终止并重新规划或接管信息物理系统的运行, 因此人对物理系统状态的及时、真实、全面的感知非常重要。

但是, 黑客入侵时为了避免被监控人员发现, 往往会破坏反馈回路, 这方面的例子在本文“引言”部分进行了介绍。如果攻击者欺骗或掩盖了人对物理系统状态的了解和异常的感知, 就会造成错误的控制或不能及时采取有效的控制。

因此, 保证“人在环路”监测结构的完整性非常重要。我们提出“环路观人”(human on the loop)的监测结构和异常检测方法, 实现的是以人为中心的网络行为监测, 持续感知和推断操作员的认知状态及其在执行任务过程中的演变, 以实现人、机对物理系统在认知域的一致性。就核电站等关键设施监控室的安全管理来看, 对操作员的监测包括四个基本的层级: (A1)基于工作现场摄像头判断人是否在工作现场和监控终端前; (A2)基于眼动跟踪, 判断在监控终端前的操作员是否在看物理系统的状态; (A3)基于电脑摄像分析和眼动跟踪数据, 判断操作员是否在认真负责任地感知物理系统的状态, 例如人是否很疲倦、注意力是否集中, 是否受到攻击者胁迫而表现得紧张和有压力, 是否在实施对系统的违规操作或恶意的破坏; (A4)判断看到的是否真实的物理系统状态。

为了防止整个监测状态被攻击者彻底劫持, 可以考虑多个监测通路的设计, 设置物理系统状态的若干直接观测点, 而不只是基于常规监测通路。为了实现 A4, 即判断传递到 *I*3 的 *I*2 是否真实反映 *I*1, 可以采用以下方法: (1) 对不同信息通道获知的 *I*2 进行比对, 判断是否一致; (2) 留有操作员直接观察 *I*1 的通道, 即建立人和物的直接感知与控制接口, 例如伊朗核电站安全事件中就是管理员听到离心机的异常才发现入侵, 也可以说是对物理系统的状态采用多模态的感知途径; (3) 判断 *I*2 是否真实反映 *I*1, 例如在物理系统中添加真实噪声或时间戳等进行识别, 但这样可能造成信息物理系统自身运行约束的破坏。从人机物融合网络系统的角度, 通过操作员的“环路观人”可构建更为完整的入侵监测环路。

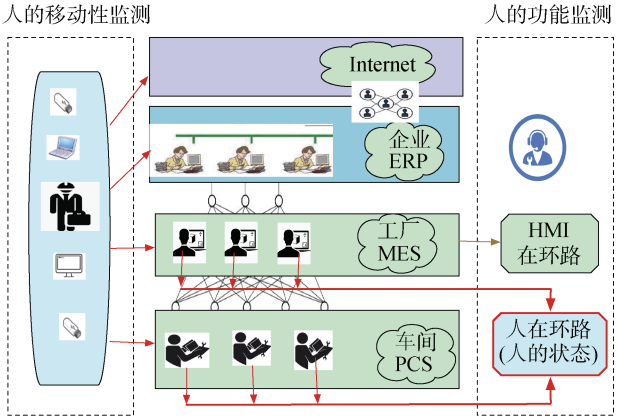


图 15 人在环路的完整性监测框架

Figure 15 Human-on-the-loop integrity monitoring framework

在图 15 中, 我们进一步拓展了对人的监测。左侧是对人作为安全入口的监测, 防止特定工作场所的 U 盘使用、人员移动不符合安全分区要求等。左侧确保人与信息物理系统“不该有的接口没有”, 右侧要保证人与信息物理系统“应该有的感知通路要完整和真实”。这需要形式化定义人与物理系统的感知接口和通路, 并进行完备的可达性测试。

4.3 社会域与信息物理域关联的异常检测

实现人机物融合的异常检测有两种基本模式: 一是对多域原始数据融合, 综合运用多域的关联规则进行异常检测; 二是对各域分别进行异常检测, 然后再对各域的异常进行关联, 如图 16 所示。

人与信息物理域的异常关联能够提高基础设施攻击行为检测的完整性和准确性: 社会域的异常可以触发信息物理空间的安全检测, 从而更加及时发现信息物理域的入侵或破坏; 社会域的异常与信息物理域的异常信号叠加, 可以作为信息物理域异常判断的置信因子。与直接对多域原始数据融合分析的方法相比, 基于异常关联的方法具有较小的复杂性。

社会域与信息物理域异常关联进行安全检测的基本依据是: 人类的思维活动、心理状态与身体语言密切关联, 例如根据读者阅读 WEB 页面的表情, 可以判断是否垃圾网页; 同样地, 存有主观故意的内部攻击者, 在对信息系统进行非法操作或基于信息系统对物理设施进行破坏时, 往往表现出紧张

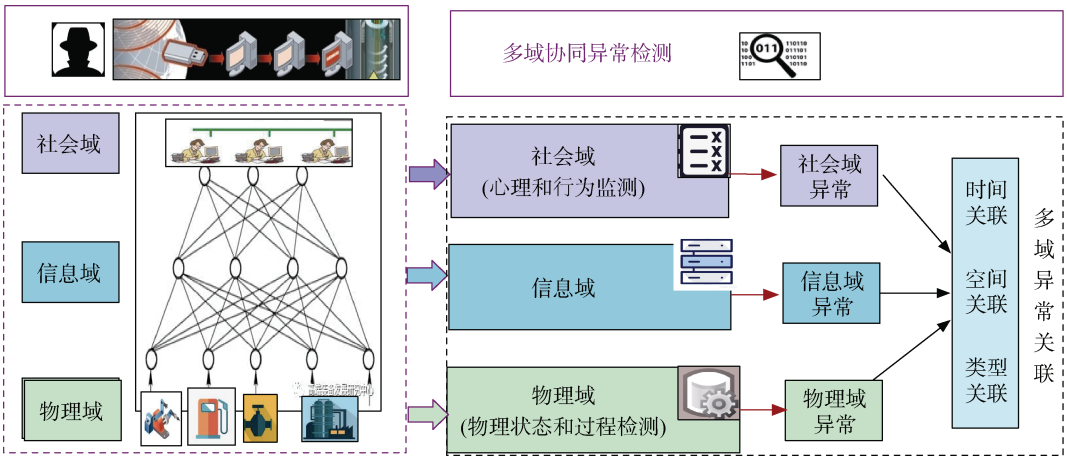


图 16　人机物多域异常关联

Figure 16　Multi-domain anomaly correlation detection

或故意遮掩的情态或预备接受惩罚、不自信等心理表现[24]。因此, 可以基于操作员的表情、眼神和姿态等社会活动空间的状态监测识别人的异常行为, 并与纯粹信息域用户行为的监测进行关联。操作员的非法操作会引起主机数据或物理系统数据的异常, 部分会反映在网络数据中, 因此可以通过信息物理系统内部的数据来共同判断非法操作行为的可疑性; 反过来, 信息域物理域的异常可以通过社会域的数据进一步核实是否攻击行为。

图 16 中人机物多域异常关联方法实现时, 首先各个域基于社会规则、信息约束、物理原理分别检测异常, 然后利用关键基础设施不同子系统之间的耦合和依赖关系, 包括人机物多域结构相关性、功能和行为关联性, 进行异常关联。具体实现框架如下:

(1) 基于视频图像等社会域的数据来判断操作员的行为是否异常;

(2) 信息物理系统分别进行异常检测获取异常数据;

(3) 时间关联: 对相同时间点或时间段的社会域异常与信息物理系统异常进行关联;

(4) 结构关联: 社会域异常与信息物理系统异常在空间上进行关联, 例如不同操作员负责不同信息物理子系统, 同一操作员在不同时间对不同信息物理子系统进行操作, 基于这种空间映射进行异常关联;

(5) 类型关联: 操作员在进行不同类型的非法操作时表现出不同的情态特征, 例如在较长时间拷贝数据时与短暂发出非法控制命令时有不同, 进行细微的非法信息窥探时与实施严重的系统破坏时有明显差异;

(6) 功能和行为关联: 基于人机物多域的功能和行为关联性进行异常关联;

(7) 基于上述多维关联形成综合的安全态势。

5　人机物多域联动的安全增强

基础设施安全的最终目标是整个人机物融合系统的系统安全(System safety)。防御者需要根据攻击者对控制信号的破坏情况并结合传感数据的非法注入情况, 调整系统的控制动作, 从而确保系统的可控性和稳定性[17]。一是强化信息域和物理域的联动, 设计物理安全感知的信息防御方法和信息安全感知的物理控制机制; 二是增强人在环内的安全控制能力。

5.1　信息物理联动的安全能力设计

信息系统是操作员控制物理系统的主要通道, 通过干扰信息系统会破坏操作员对物理系统的有效控制, 例如干扰控制信息网络 QoS 会造成控制命令的丢失或延迟等, 这是信息物理协同的攻击, 因而需要采用信息物理联动的安全设计来应对。如图 17 所示, 信息物理联动的安全能力包括信息安全感知(security-aware)的控制功能增强和物理安全感知(safety-aware)的信息域防御方法。d 为安全防御的机制, u 为安全控制的操作, a 为信息攻击, w 为物理扰

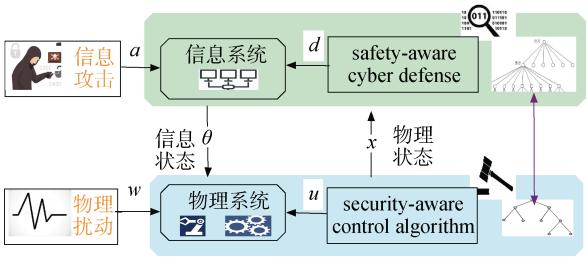


图 17　信息物理联动的安全增强

Figure 17　Cyber-physical coordinated security enhancement

动。信息安全方法的设计要考虑来自物理设施的安全威胁, 物理系统控制算法的设计要考虑来自信息域的安全威胁。

信息物理联动的安全增强, 包括基于通信网络传输能力的控制系统安全增强, 物理系统的控制器的鲁棒性增强和控制系统的弹性增强。

1. 基于通信网络传输能力的控制系统安全增强

(1) 估计通信网络状态和 QoS, 拟定传感器和控制数据的发送策略;

(2) 基于信息物理系统的安全关联关系, 探究基于信息网络的物理破坏机理, 设计信息安全感知的控制算法, 例如, 图 17 右侧所示物理域的博弈树和信息域的博弈树互相耦合。

2. 控制器鲁棒性增强

(1) 部分传感器被恶意数据注入后, 系统仍然能够正常进行状态估计;

(2) 控制器通过产生修正信号消除恶意命令注入的影响;

(3) 物理系统在通信网络被打击时的开环工作能力, 例如 SCADA 被破坏时电网基于物理系统自身的动力学机制运行。

3. 控制系统的弹性增强

(1) 信息网络的应急通信和快速恢复;

(2) 控制器检测到攻击后重新配置, 切换到新的脆弱性小的配置; 若有数字孪生, 则利用其产生的数据支持快速恢复。

5.2 人机物联动的安全控制能力增强

系统操作员作为人, 其行为具有不确定性和不可预测性, 是控制环路中最弱的一环。人的控制可能产生不安全行为, 例如, 在工程领域, 管理者违章指挥、监管人员未核对操作项目、操作员随意解除闭锁装置; 心存不满的内部员工可能恶意操纵大量设施; 由信息域引入的许多安全漏洞大都是由于人的认知错误、有限的推理能力和错误的风险感知造成的。除了规范管理和强化培训来增强员工安全意识和提高安全能力, 也可以设计好的机器学习算法和基于物联网大数据更好地支持人进行决策。而本文从员工现场行为和认知角度, 考虑人与信息物理系统联动的安全能力增强; 人和信息物理系统安全交互的现有研究, 主要涉及机器人[25]、无人机[26]、无人车、医疗器械等单个信息物理融合设备的操纵, 而本文主要面向工业系统等大规模基础设施。情感分析、行为建模、智能态势检测、知识生成等通过计算技术变得更加容易, 使人的深入综合分析变得可行[24]。

5.2.1 操作员行为不确定性的调控

人在工作现场的行为存在很大的不确定性。将人纳入控制环路与环路中的信息物理控制部件不同的是: 人不一定能在相同的刺激下重复或再现相同的精确动作, 这是由于受攻击者干扰、操纵或身心疲劳、有压力等原因造成的。

在控制系统中针对人的精神和心理状态建模, 可分为三个类型: (1) 通用的人的模型, 例如人的控制要比信息物理系统的自动调控有更大时延, 视觉反应延迟至少 40ms(因为人眼的识别连贯图像的速度是 24 帧/秒); (2) 特定操作员的模型, 例如挖掘和学习不同操作员的人格特点和操作习惯, 使用系统辨识技术获得与每个被评估操作员对应的控制器模型; (3) 实时的操作员心理状态感知, 例如, 通过表情分析和眼动跟踪等识别注意力是否集中、存在压力、疲劳或者无聊, 如果超过一定阈值就进行告警并由更高阶的技术系统调控或接管。

图 18 给出可采用的三层控制结构, 第一层是自动化系统(信息系统)对物理系统的运行的自动调控; 第二层是人对自动化系统(信息系统)进行监控, 例如发生严重故障时进行系统重新配置或与自动化系统进行人机协同控制, 是通过人对信息系统能力的增强; 第三层针对操作员受到的攻击进行社会域的调控, 对人的不确定行为进行补偿。

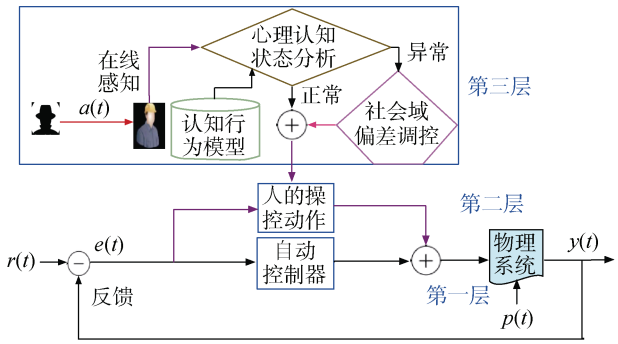


图 18 以人为中心的安全控制能力增强

Figure 18 Human-centric security control enhancement

图 18 采用控制科学的结构来描述, 图中 $e(t)$是系统参考信号 $r(t)$和系统输出 $y(t)$之间的偏差, $p(t)$是物理系统扰动或噪声, 第二层人类操作员的主要工作是减少系统偏差。但是人类操作员自身存在不确定因素, 尤其是受到攻击者操纵或干扰 $a(t)$的影响时, 因此对人自身的偏差也需要进行调控补偿。

我们将操作员的认知行为模型融入人机物融合系统的安全控制。图 18 中的“认知行为模型”包括基于认知科学和经验测量获得的人的通用认知行为

模型, 线下线上学习获知的特定操作员的认知行为模型, 以及实时线上感知的心理认知状态。

对社会域的偏差进行调控和安全增强, 首先要识别操作员异常表现的类型, 分析行为不稳定的心理根源; 然后区分攻击者是对人的哪种认知能力进行攻击, 基于认知过程的感知、注意、记忆、思维处理等环节进行调控或补偿。具体思路如下:

(1) 感知/注意/记忆力调控: 如果信息物理系统运行的偏差在于操作员注意力不集中、过于疲劳等, 这时候提醒操作员为主要任务进行更多的精力分配, 压缩低优先级的任务, 或者帮助提升操作员精神状态, 这里人的精力的调节体现了**人机物融合系统中社会域资源的再分配**。

(2) 思维处理能力调控: 如果操作员具有主观破坏工业设施稳定运行的可能, 而且无法通过物理域和信息域的安全手段进行调控, 这时候可向车间或企业的管理员进行告警; 或者运用控制理论将操作员建模为伺服机构, 基于操作员心理行为特征、控制过程、物理机器的融合进行一体化调控。

5.2.2 操作员认知非理性的修正

关键基础设施的攻防双方运用博弈论的理论和方法可以更加有效地进行人机物安全对抗。经典的博弈论假设参与人是理性的并且追求自身利益的最大化, 并不符合操作员决策的实际情况。操作员由于记忆力、注意力和推理能力的有限性会造成错误, 决策本身可能存在有界理性(Bounded rationality)的倾向; 同时在攻击者的利益引诱和压力施加情况下, 操作员的认知会包含更多非理性因素和更加偏离理性的边界。

行为博弈论(Behavioral game theory)[27]通过心理学成果来修正描述实际行为失真的标准博弈论原则, 介于超理性均衡分析和低理性适应性分析之间。这就给我们基于操作员的认知特征改造传统安全博弈算法和进行新型算法的设计, 即认知模型融入信息物理对抗博弈算法, 提供了理论依据和方法指引。

(1) 博弈过程中的非理性

行为博弈论[27]取消了对抗过程中对参与者的完全理性假设, 通过增加情感干扰、犯错误的可能性、有限的洞察力等因素扩展了分析博弈的内容, 从而可以更加准确理解与预测人的决策行为。

与 5.2.1 节面向行为不确定的操作员建模类似, 对操作员的认知非理性也可在三个层次上进行建模: 通用的人的模型, 特定操作员的模型, 以及实时的操作员认知状态模型。基于行为经济学和行为博弈论研究的成果, 计算通用的人的非理性因子 r^g; 对不同的操作员基于统计学的方法等调查计算各自的非理性因子 r^p, 该因子具有较长的稳定性; 如果受到攻击者心理的操纵, 实时非理性的倾向 r^t 会一时增强。这三类模型分别度量操作员的非理性程度, 综合加权得到操作员认知的非理性因子 r。具体到感知、记忆、处理等认知过程的环节又分别有 r_p、r_m 和 r_d 等分因子。认知处理环节的非理性因子 r_d 主要是决策非理性因子。

(2) 对信息防御和物理安全控制算法的修正

针对不同认知活动中人的特点来研究关键基础安全对抗中的非理性博弈或者有界理性博弈算法, 需要使用非理性因子来修正信息域和物理域的相关博弈算法, 通过设计 Human-Aware(对人进行感知)的信息安全算法和物理安全控制方法, 从而实现操作员认知模型在信息域、物理域安全算法中的直接融入。

基于非理性因子的安全能力增强模型如图 19 所示, 社会域处于中间, 对上面的信息域和下面的物理域都有影响, 即认知的非理性影响信息域、物理域博弈算法的设计。图 19 中 a 表示对信息系统的攻击, m 表示对操作员的操纵, w 表示对物理系统的干扰; d 表示采用新的信息安全防御机制来保护信息系统, u 表示采用新的控制机制来保护物理系统。

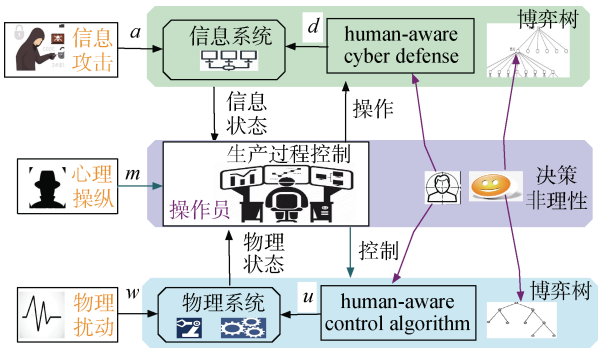


图 19 基于非理性因子的安全能力增强

Figure 19 Security capability enhancement in view of irrational factors

基于决策、记忆与学习、感知等多个认知能力和认知环节的非理性对原有博弈算法修正的具体思路:

1) 考虑认知处理能力非理性的信息物理系统博弈方法

在对抗博弈过程的模拟和对抗方法的设计过程中, 需要对攻防双方的认知处理模型进行假设, 这时候除了采用传统的保守的代价最小化模型、随机化模型等, 还要考虑人类决策固有的非理性经济学

模型, 例如基于双曲线贴现(Hyperbolic discounting)假设, 这是基础设施操作员受到诱惑时很可能采取的决策策略。对现有博弈算法最简单的修正, 就是其决策和行动概率再乘以(1-决策非理性因子), 这样后续对抗过程中会基于前序决策或动作的不成功而有更进一步的补救或强化措施。

2) 考虑记忆与学习能力有限性的信息物理系统博弈方法

基于多次博弈或高级持续威胁(APT)持续博弈算法的研究往往对参与者做了完全理性的假设, 例如了解对方的动作历史并能估计对方在当前回合中最可能的动作, 从而相应地选择最佳反应。本文要考虑参与者记忆能力的有限性和学习能力的现实性, 即既不能完全记住以往的对抗历史也不一定能够有效地从经验教训中学习和提高。这就为新型算法的设计提供了动因和指明了路径, 例如在多轮博弈的学习算法方面可能采用需要掌握的信息量比较少的基于实例的重复社会互动学习认知模型, 或者参考非理性因子 r 按一定概率采取最佳响应策略。这方面的研究值得业界同行进一步深化。

3) 考虑感知能力易变性的非理性博弈方法

经典博弈论中要求参与者对相关变量的赋值具有客观性, 即对环境和攻防双方的感知具有理性。而现实对抗环境中基础设施操作员在受到胁迫或诱惑时其情感状态的非理性, 会带来感知的易变性和不准确性, 因此在相关算法中对环境和对方的感知能力需要给予更多不确定性的假设, 具体如何设置还需要业界同行进一步研究。

6 相关工作

首先介绍基本的人机物融合模型的研究, 重点阐述新型人机物融合计算模型的研究进展; 然后介绍人机物融合系统的安全模型与技术的研究, 考察现有的人机物融合的安全威胁方法和防御机制; 最后结合我们的阶段性成果, 介绍人机物融合安全在复杂网络和网络科学方面的研究进展。

6.1 人机物融合模型的研究

人机物融合系统(HCPS)是基于信息技术的系统工程的最新进展和新工业革命的使能技术[28], 人机物融合计算模型的研究成为新一代软件系统和计算系统研制的新的基本要求, 其核心是在信息领域主动强化物的感知和面向人的需求。李国杰院士提出计算过程从信息空间拓展到包含人机物三元世界, 体现万物互联时代物端计算、信任互联网等新型计算形态[28]。计算机科学需要研究 HCPS 的计算模型、交互机理、集成的元语义模型、规约与验证、程序设计模型和描述语言等[29]; 操作系统的设计要考虑人的感知、认知、决策和行动[30]。智能制造、智能交通、智慧医疗等系统中, 海量多模态的人的数据或知识、信息系统产生的大数据、工程领域和物理世界的数据或知识, 需要彼此融合或协同[31-33]。但是对人机物融合的安全模型缺乏系统深入的研究。

6.2 人机物融合系统的安全模型与技术

人机物融合带来社会域、信息域和物理域的新的安全威胁, 大量的人机物融合系统成为新的攻击目标; 另一方面, 人机物多域协同也是实现网络空间安全的新路径。

美国在智能电网等关键基础设施的安全方面, 最有代表性的研究机构是CIRI(关键基础设施韧性研究院), 现在已开始从人机物多个维度和多学科角度应对关键基础设施的安全挑战。美国国家科学基金会 2020 年 CPS 项目指南[34]开始重视 CPS 的社会技术属性和向社会宏观规模的拓展。

6.2.1 人因安全技术的研究

在信息安全中人的因素(Human factor)方面, 已有研究[35-36]主要考虑如何针对人的粗心、缺乏安全知识, 而采用自动化安全管理工具或最佳安全实践来增强安全性。对单位员工的人格、精神状态、人际关系及其上网行为的分析, 可以帮助识别内部威胁[37]。社会工程学方法利用人的好奇、轻信、贪婪等心理脆弱性来操纵信任关系, 以突破物理隔离和渗透信息系统[38]。

在物理安全中人员风险因素的量化建模方面, 现有研究基于人因可靠性理论计算人为失误概率值, 例如利用上岗时间、技能水平、责任心、历史失误率等基础数据, 以及本次操作的任务强度、时间段、心理情绪、连续工作时长等实时数据, 来计算电网调度过程中的人因风险[39]。随着机器视觉技术的进步, 文献[40]对施工机械安全智能监控, 提出多模态融合的基于场景理解的施工现场工人不安全行为识别方法。

6.2.2 人机物融合安全技术的研究

典型的关键基础设施攻击事件表明, 攻击者在侦察、入侵和破坏三个环节, 充分利用了社会、信息、物理多域的脆弱性, 攻击路径的构建利用了多域的关联耦合关系, 但是防御技术的研究尚未做好充分的应对。

(1) 攻击者对基础设施人机物多域进行全面侦察和利用, 移动目标防御虽然在信息域已有较多研究, 但是在物理域的研究刚刚起步, 社会域和人机

物多域融合的移动目标防御的研究存在空白, 难以应对攻击者对基础设施人机物多域的侦察。

(2) 攻击者对基础设施的跨域渗透入侵和多域协同破坏具有高度隐蔽性, 防御者可基于态势察觉、态势理解和态势投射等构建工业控制系统安全态势感知系统[41]; 但是信息物理融合的检测在异构数据的同构化、多源数据的关联分析、关联建模等方面还存在很多问题[7]。

(3) 攻击者利用基础设施人机物的安全关联和依存关系进行破坏, 物理域传统的容错设计相对成熟, 信息物理协同的安全控制有所进展, 例如设计部分数据丢失和存在延迟情况下保证临近区域间负荷频率的控制算法[42], 文献[43]提出系统在传感器发生故障或通信中断时具有开环工作能力, 但是对存在信息域攻击和敌手存在条件下的对抗性安全控制算法的研究还只在非常简单的场景有所体现; 人在环路的安全控制增强尤其是人的脆弱性如何避免被攻击者利用方面的研究尚未展开。

本文提出的人机物协同的对抗模型, 是在传统信息网络对抗模型的基础上, 引入社会域对抗并强化信息物理协同, 其核心是攻防双方充分利用人机物多域之间的耦合与关联关系。为了有效利用人机物的各类实体和系统运行的众多状态, 我们基于三个环节的安全对抗子模型, 设计新型安全机制, 选择使用不同类别实体的不同类型状态, 通过对状态数据的感知、分析和利用来实现安全防御、检测与增强方法。

在系统对抗中, 可采用以下途径来减少对系统状态数量的要求, 从而避免系统状态的爆炸性增长: (1) 设计高效的安全机制重点选用必要的状态而忽略其他状态; (2) 通过对状态聚类和分类, 减少对细微状态信息的要求; (3) 受防御者资源和实现代价的约束系统真正可能的状态有限; (4) 将基于模型和知识的方法与基于大数据分析的方法结合, 通过模型的抽象能力来屏蔽内部的细节与状态; (5) 通过系统对抗过程中实时的探测与感知, 减少假设状态的不确定性和减少系统可能的状态数量。

6.3 基于复杂网络理论和方法的关键基础设施安全研究

从复杂网络角度对关键基础设施安全的考察, 主要包括人机物关联的健壮性模型与安全威胁评估, 以及新型跨域渗透威胁模型与方法。

我们分析了信息物理网络不同耦合模式对系统健壮性的影响[13], 利用了信息层的威胁传播机制和物理层的失效扩散机制来实时地生成多条可达的跨层攻击路径[44], 对共时攻击、序列攻击与组合攻击等进行安全评估[45], 并设计基于依赖关系分析的人机物安全状态定量评估方法[46]。

面向特定领域的跨越渗透攻击具体方法的研究目前还比较少。我们在文献[15]中研究基于社交网络虚假消息操纵大量电力用户开关电器, 从而影响电力网络稳定性的威胁模型, 实现了社会域发起的物理域攻击, 并且将 AMITT(虚假信息影响技战术)与 ICS ATT&CK(工控系统对抗技战术和常识)融合起来, 构建和刻画完整的 SCAC(Social Collective Attack on CPS) 攻击链。我们在文献[47]中提出基于群体智能的关键基础设施网络自愈机制, 以恒温电器的智能调节来达到电网安全运行的目的; 在文献[48]中研究如何增加相连边和相依边, 以最少的资源实现网络健壮性的增强。

7 结语

本文基于大规模基础设施人机物融合的特点, 针对典型高级威胁安全事件中攻击者在侦察、入侵、破坏多个环节对人机物脆弱性的充分利用和人机物多域协同的攻击模式, 引入社会域安全对抗, 强化和丰富了现有信息物理协同的安全对抗, 逐步构建起人机物多域一体化防御体系。

我们提出比较完整的以人为中心的安全对抗模型: 通过引入社会域以人为中心的移动目标防御, 减少人作为攻击入口(to the loop)的风险; 通过引入以人为中心的网络行为监测和社会域与信息物理域异常关联的检测方法, 实现环路观人(on the loop)的监测结构; 面向认知过程的感知、注意、记忆、学习、决策等环节引发的行为不确定性和认知的非理性, 通过社会域偏差的调控和行为博弈论的方法, 来增强人在环路(in the loop)的安全控制能力。

本文尝试构建了人机物融合安全模型和方法研究的基本框架, 将有助于丰富网络空间安全基础模型和发展人机物融合的计算范式。

参考文献

[1] Zhou J, Zhou Y H, Wang B C, et al. Human–Cyber–Physical Systems (HCPSS) in the Context of New-Generation Intelligent Manufacturing[J]. *Engineering*, 2019, 5(4): 624-636.

[2] Xue Y S, Yu X H. Beyond Smart Grid—Cyber–Physical–Social System in Energy Future[J]. *Proceedings of the IEEE*, 2017, 105(12): 2290-2292.

[3] Zhu Peidong, Dong Wei. Modes and Mechanisms of Human-Cyber-Physical Convergence[J]. *Communications of the CCF*, 2022, 18(10): 59-67.

(朱培栋, 董威.人机物融合的模式与机理初探[J]*中国计算机学会通讯*, 2022, 18(10): 59-67.)

[4] Iaiani M, Tugnoli A, Bonvicini S, et al. Analysis of Cybersecurity-Related Incidents in the Process Industry[J]. *Reliability Engineering & System Safety*, 2021, 209: 107485.

[5] Huang L and Zhu Q. A Dynamic Games Approach to Proactive Defense Strategies against Advanced Persistent Threats in Cyber-Physical Systems[EB/OL]. 2019: ArXiv Preprint ArXiv:1906.09687v2.

[6] Huang L, Zhu Q. RADAMS: Resilient and adaptive alert and attention management strategy against Informational Denial-of- Service (IDoS) attacks[EB/OL]. 2021: ArXiv Preprint ArXiv:2111.03463.

[7] Liu T, Tian J, Wang J Z, et al. Integrated Security Threats and Defense of Cyber-Physical Systems[J]. *Acta Automatica Sinica*, 2019, 45(1): 5-24.
(刘烃, 田决, 王稼舟, 等. 信息物理融合系统综合安全威胁与防御研究[J]. *自动化学报*, 2019, 45(1): 5-24.)

[8] Wang Z S. Information Ontology and the Construction of Information Science Paradigm[J]. *Studies in Dialectics of Nature*, 2018, 34(7): 108-114.
(王振嵩. 信息本体论与信息科学范式的建构[J]. *自然辩证法研究*, 2018, 34(7): 108-114.)

[9] Wu K. Information philosophy: Theory, system and method[M]. Beijing: The Commercial Press, 2005.
(邬焜. 信息哲学: 理论、体系、方法[M]. 北京: 商务印书馆, 2005.)

[10] Milanović J V, Zhu W T. Modeling of Interconnected Critical Infrastructure Systems Using Complex Network Theory[J]. *IEEE Transactions on Smart Grid*, 2018, 9(5): 4637-4648.

[11] Cui P S, Zhu P D, Wang K, et al. Enhancing Robustness of Interdependent Network by Adding Connectivity and Dependence Links[J]. *Physica A*: *Statistical Mechanics and Its Applications*, 2018, 497: 185-197.

[12] Lu Q C, Xu P C, Zhao X M, et al. Measuring Network Interdependency between Dependent Networks: A Supply-Demand-Based Approach[J]. *Reliability Engineering & System Safety*, 2022, 225: 108611.

[13] Kang W J, Hu G, Zhu P D, et al. Influence of Different Coupling Modes on the Robustness of Smart Grid under Targeted Attack[J]. *Sensors*, 2018, 18(6): 1699.

[14] Xun P, Zhu P D, Maharjan S, et al. Successive Direct Load Altering Attack in Smart Grid[J]. *Computers & Security*, 2018, 77: 79-93.

[15] Zhu P D, Xun P, Hu Y F, et al. Social Collective Attack Model and Procedures for Large-Scale Cyber-Physical Systems[J]. *Sensors*, 2021, 21(3): 991.

[16] Jiang G P, Wang X W, Wu X. Survey on Controllability and Observability in Complex Dynamical Networks[J]. *Journal of Nanjing University of Posts and Telecommunications* (*Natural Science Edition*), 2020, 40(5): 178-185.
(蒋国平, 王欣伟, 吴旭. 复杂动态网络可控性及可观性研究综述[J]. *南京邮电大学学报(自然科学版)*, 2020, 40(5): 178-185.)

[17] Barreto C, Cárdenas A A, Quijano N. Controllability of Dynamical Systems: Threat Models and Reactive Security[M]. Decision and Game Theory for Security. Cham: Springer International Publishing, 2013: 45-64.

[18] Liu S, Chen B, Zourntos T, et al. A Coordinated Multi-Switch Attack for Cascading Failures in Smart Grid[J]. *IEEE Transactions on Smart Grid*, 2014, 5(3): 1183-1195.

[19] Yuan Y, Tang X C, Zhou W, et al. Data Driven Discovery of Cyber Physical Systems[J]. *Nature Communications*, 2019, 10: 4894.

[20] Hu Y F, Xun P, Zhu P D, et al. Moving Target Defense Based on Adaptive Forwarding Path Migration for Securing the SCADA Network[J]. *Security and Communication Networks*, 2021, 2021: 1704125.

[21] Hu Y F, Xun P, Zhu P D, et al. Network-Based Multidimensional Moving Target Defense Against False Data Injection Attack in Power System[J]. *Computers & Security*, 2021, 107: 102283.

[22] Zhang T, Zhu Q. A Game-Theoretic Foundation of Deception: Knowledge Acquisition and Fundamental Limits[EB/OL]. 2018: ArXiv Preprint ArXiv:1810.00752.

[23] Kailkhura B, Han Y S, Brahma S, et al. Distributed Bayesian Detection in the Presence of Byzantine Data[J]. *IEEE Transactions on Signal Processing*, 2015, 63(19): 5250-5263.

[24] Makkar A, Park J H. SecureCPS: Cognitive Inspired Framework for Detection of Cyber Attacks in Cyber–Physical Systems[J]. *Information Processing & Management*, 2022, 59(3): 102914.

[25] Portugal D, Pereira S, Couceiro M S. The Role of Security in Human-Robot Shared Environments: A Case Study in ROS-Based Surveillance Robots[C]. *2017 26th IEEE International Symposium on Robot and Human Interactive Communication*, 2017: 981-986.

[26] Zhu H B, Elfar M, Pajic M, et al. Human Augmentation of UAV Cyber-Attack Detection[M]. Augmented Cognition: Users and Contexts. Cham: Springer International Publishing, 2018: 154-167.

[27] 桑吉特·达米(著), 汪丁丁等(译).行为博弈理论[M]. 上海: 格致出版社, 2022.

[28] Li G J. The Long-Term Development Trend of Information Science and Technology and China's Strategic Orientation[J]. *Scientia Sinica Informations*, 2010, 40(1): 128-138.
(李国杰. 信息科学技术的长期发展趋势和我国的战略取向[J]. *中国科学: 信息科学*, 2010, 40(1): 128-138.)

[29] Liu Z M, Wang J. Human-Cyber-Physical Systems: Concepts, Challenges, and Research Opportunities[J]. *Frontiers of Information Technology & Electronic Engineering*, 2020, 21(11): 1535-1553.

[30] Mei H, Cao D G, Xie T. Ubiquitous Operating System: Toward the Blue Ocean of Human-Cyber-Physical Ternary Ubiquitous Computing[J]. *Bulletin of Chinese Academy of Sciences*, 2022, 37(1): 30-37.
(梅宏, 曹东刚, 谢涛. 泛在操作系统: 面向人机物融合泛在计算的新蓝海[J]. *中国科学院院刊*, 2022, 37(1): 30-37.)

[31] Gui W H, Zeng H Z, Chen X F, et al. Knowledge-Driven Process Industry Smart Manufacturing[J]. *Scientia Sinica* (*Informationis*), 2020, 50(9): 1345-1360.
(桂卫华, 曾朝晖, 陈晓方, 等. 知识驱动的流程工业智能制造

[J]. *中国科学: 信息科学*, 2020, 50(9): 1345-1360.)

[32] Dong J, Wang J, Wang Z P. Automatic Ontology Construction for Human-Cyber-Physical Data Fusion in Manufacturing Domain[J]. *Control and Decision*, 2022, 37(5): 1251-1257.
(董津, 王坚, 王兆平. 面向制造领域人机物三元数据融合的本体自动化构建方法[J]. *控制与决策*, 2022, 37(5): 1251-1257.)

[33] Zhai S Y, Guo B, Li R, et al. Cyber-Physical-Social System: A Data-Centric Framework[J]. *Big Data Research*, 2017, 3(6): 85-92.
(翟书颖, 郭斌, 李茹, 等. 信息物理社会融合系统: 一种以数据为中心的框架[J]. *大数据*, 2017, 3(6): 85-92.)

[34] National Science Foundation. Cyber-Physical Systems (CPS) Program Solicitation[EB/OL]. https://www.nsf.gov/pubs/2020/nsf20563/nsf20563.htm.

[35] Oltramari A, Henshel D, Cains M, et al. Towards a Human Factors Ontology for Cyber Security[C]. *10th Conference on Semantic Technology, Defense and Security, Virginia*, 2015.

[36] 16th Workshop on Security and Human Behavior (SHB)[EB/OL]. https://www.heinz.cmu.edu/~acquisti/SHB2023/index.htm, June 2023.

[37] Mohammed M A, Kadhem S M, Maisa'a Abid Ali K. A Novel Approach for Detection Insider Attacker Using Body Language[J]. *Journal of Physics*: *Conference Series*, 2021, 1804(1): 012129.

[38] Wang Z G, Zhu H S, Sun L M. The Concept Evolution Analysis of Social Engineering[J]. *Journal of Cyber Security*, 2021, 6(2): 12-29.
(王作广, 朱红松, 孙利民. 社工概念演化分析[J]. *信息安全学报*, 2021, 6(2): 12-29.)

[39] Lu E, Hu S Z, Zhan C L, et al. Modelling of Human Risk Factors during Power Grid Dispatching Operation[J]. *Automation of Electric Power Systems*, 2016, 40(17): 163-168, 216.
(卢恩, 呼士召, 占才亮, 等. 电网调度操作过程中的人员风险因素建模[J]. *电力系统自动化*, 2016, 40(17): 163-168, 216.)

[40] 侯景严. 基于机器视觉的高危企业生产过程智能监控[D].长安大学,2021.

[41] Zhou M, Lv S C, You J Z, et al. A Comprehensive Survey of Security Situational Awareness on Industrial Control Systems[J]. *Journal of Cyber Security*, 2022, 7(2): 101-119.
(周明, 吕世超, 游建舟, 等. 工业控制系统安全态势感知技术研究[J]. *信息安全学报*, 2022, 7(2): 101-119.)

[42] Peng C, Li J C, Fei M R. Resilient Event-Triggering H∞ Load Frequency Control for Multi-Area Power Systems with Energy-Limited DoS Attacks[J]. *IEEE Transactions on Power Systems*, 2017, 32(5): 4110-4118.

[43] Cardenas A A, Amin S, Sastry S. Secure Control: Towards Survivable Cyber-Physical Systems[C]. *2008 The 28th International Conference on Distributed Computing Systems Workshops*, 2008: 495-500.

[44] Kang W J, Zhu P D, Hu G, et al. Cross-Layer Attack Path Exploration for Smart Grid Based on Knowledge of Target Network[C]. *Knowledge Science, Engineering and Management*, 2018: 433-441.

[45] Cui P S, Zhu P D, Xun P, et al. Robustness of Cyber-Physical Systems Against Simultaneous, Sequential and Composite Attack[J]. *Electronics*, 2018, 7(9): 196.

[46] Liu X X, Zhang J X, Zhu P D, et al. Quantitative Cyber-Physical Security Analysis Methodology for Industrial Control Systems Based on Incomplete Information Bayesian Game[J]. *Computers & Security*, 2021, 102: 102138.

[47] Cui P S, Feng L, Xun P, et al. Load Scheduling of Thermostatical House-Hold Appliances Against Abrupt Changes in Smart Grid[C]. *2017 10th International Symposium on Computational Intelligence and Design*, 2017: 470-475.

[48] Cui P S, Zhu P D, Shao C C, et al. Cascading Failures in Interdependent Networks Due to Insufficient Received Support Capability[J]. *Physica A*: *Statistical Mechanics and Its Applications*, 2017, 469: 777-788.

朱培栋 于 1999 年在国防科技大学计算机科学与技术专业获得博士学位。曾任国防科技大学教授, 现任长沙学院电子信息与电气工程学院院长。CCF 物联网专委会执委。研究领域为网络空间安全、新一代互联网。研究兴趣包括: 物联网安全, 人机物融合网络。Email: pdzhu@nudt.edu.cn

康文杰 于 2018 年在国防科技大学计算机科学与技术专业获得博士学位。CCF 高级会员, 湖南高等院校青年骨干教师, 国防科技大学系统工程学院博士后。现为湖南警察学院信息技术系讲师。研究领域为网络安全、警务大数据。研究兴趣包括: 物联网安全、复杂网络。Email: kangwenjie@nudt.edu.cn

刘亮 于 2004 年在湖南大学控制理论与控制工程专业获得硕士学位。现为长沙学院电子信息与电气工程学院副教授。CCF 专业会员。研究领域为智能控制理论与应用。研究兴趣包括: 工业设施安全控制。Email: liuliang@ccsu.edu.cn

张瑞 于 2021 年在中南大学控制科学与工程专业获得博士学位。现任长沙学院电子信息与电气工程学院讲师。CCF 专业会员。研究领域为智能控制理论与应用。研究兴趣包括: 智能车辆安全控制。Email: ruizhang@ccsu.edu.cn

荀鹏 于2018年在国防科技大学计算机科学与技术专业获得工学博士学位。现在国防科技大学计算机学院网络空间安全系任助理研究员。研究领域为网络体系结构、信息物理系统安全。Email: xunpeng12@nudt.edu.cn