一种面向业务操作防护的工控系统可信行为决策模型

汪京培，程鹏，XXX，孙优贤
（浙江大学，浙江 杭州 310063）

摘要：工业控制系统的逻辑缺陷的认知局限性和安全漏洞利用的便捷性使得基于被动检测的安全防护效果不佳，基于可信计算的主动防御技术为解决工控系统安全防护难题提供了可行思路，然而如何根据操作行为可信度对恶意行为进行有效屏蔽且不影响控制网络的连通性是该技术需要完善的难题。提出一种面向业务操作防护的工控系统可信行为决策模型，分析评估工控系统典型攻击行为，解析工控系统平台、边界、网络上的操作与交互行为，面向主动免疫运行目标提取信任信息并对行为信任建模，建立工控系统网络连通性约束下的信任管理机制，保障工控系统操作可用、可信。以 DCS 组态软件关键控制数据的可信防护为例，对所提的方法进行了理论分析与实验验证，证明了所提方法的有效性和优越性。

关键词：工业控制系统，信息安全，可信，主动免疫，信任管理

A Trustworthy Behavior Decision Model for Industrial Control System for Business Operation Protection

Jingpei Wang, Peng Cheng, XXX, Youxian Sun
（Zhejiang University, Hangzhou 310063, China）

Abstract: The cognitive limitation of logical defects in industrial control systems and the convenience of exploiting security vulnerabilities make the security protection based on passive detection ineffective. Active defense technology based on Trusted Computing provides a feasible way to solve the problem of industrial control system security protection. However, how to effectively shield malicious behavior according to the credibility of operational behavior and does not affect the connectivity of the control network is a difficult problem that the technology needs to be perfected. This paper proposes a credible behavior decision-making model for industrial control system for business operation protection. It analyzes and evaluates the typical attack behavior of industrial control system, and analyzes the operation and interaction behavior of industrial control system platform, boundary and network, and then extracts trust information and models behavior trust oriented to active immune operation target, and finally establishes a trust management mechanism under the

constraints of network connectivity of industrial control systems, ensuring that the operation of industrial control systems is available and credible. Taking the trusted protection of key control data of DCS configuration software as an example, the proposed method is analyzed theoretically and verified experimentally, which proves the effectiveness and superiority of the proposed method.

Keywords: Industrial control system, information security, trustworthiness, active immunity, trust management

1 引言

工业控制系统（简称“工控系统”），广泛应用于核设施、石油石化、水处理、天然气、先进制造等关键基础设施，扮演中枢神经的作用。由于工控系统本身固化大量信息安全漏洞、信息物理系统开放互联以及通用智能化构件广泛应用的趋势，黑客攻击、病毒、木马等威胁正在向工业控制系统扩散，相继发生了“震网”、“毒区”、“火焰”、“蜻蜓组织”、乌克兰停电事故等一系列震惊全球的工业安全事件。与此同时，针对工控系统的攻击方法不断进步，大量可用的渗透工具使得工具门限降低。工控系统攻击严重威胁着关键基础设施的正常运转，对工业控制系统的安全防护已经迫在眉睫[1]。

但是由于工业控制系统具有如下特点：1）工控系统可用性要求高，可用性>完整性>保密性，普通 IT 系统正好相反；2）业务连续性与可靠性要求高，实时通信且不允许重启；3）以业务操作防护为主，最关注业务系统可靠、操作合规、行为可信；4）攻击跨越信息物理空间，与物理与环境互动；5）设备种类繁多，且多基于嵌入式系统（如 VxWorks、WinCE 等）开发和运行、多采用专用通信协议（如 OPC、Modbus、DNP3 等），传统的防火墙、IDS、IPS 存在协议无法识别或干扰可用性的可能。使得普通信息系统中的信息安全理论和方法无法直接应用到工控系统中[2]。

现有关于工控系统的防护主要采用纵深防御的方法，然而基于防火墙、入侵检测、病毒查杀等方法存在如下问题：1）封堵查杀难以应对利用逻辑缺陷的攻击，人们对工控系统的逻辑缺陷的认知局限性，不能穷尽所有组合，只能局限于完成计算任务去设计 IT 系统，必定存在逻辑不全的缺陷；2）被动防御防不胜防，超级权限用户违背安全原则；3）现有防御方法可以被攻击者控制，成为网络攻击的平台，“棱镜门”即是利用防火墙收集情报，病毒库篡改后可以导致

系统瘫痪。因此只有重建主动免疫的可信体系才能有效抵御攻击。沈院士认为：“主动免疫防御即是确保为完成计算任务的逻辑组合不被篡改和破坏，实现正确计算。”基于此，提出基于可信计算的网络安全主动免疫防御方案[3]，为工控系统的内生安全防护提供典型的研究方向。

基于可信计算方案的核心是可信性评估。针对业务操作防护任务，然而如何根据操作行为可信度对恶意行为进行有效屏蔽且不影响控制网络的连通性是该方案需要完善的难题。目前针对工控系统可信防御的研究工作极少，现有工作主要是信任网络架构研究[4-5]，或是普通信息系统可信机制的简单应用[6-7]，或是针对节点信任有限维度的建模[8-10]，尚没有考虑工控系统跨域攻击与可用性优先特征下的确保业务连续性的操作与交互行为信任建模的主动免疫机制。本文提出一种面向业务操作防护的工控系统可信行为决策模型，解析工控系统平台、边界、网络上的操作与交互行为，面向主动免疫运行目标提取信任信息并对行为信任建模，建立工控系统网络连通性约束下的信任管理机制，保障工控系统业务操作可用、可信。

2 研究现状

为了改善工控系统被动防护的防不胜防问题，工控系统主动防御得到广泛关注，比较典型的是可信防御方案[3]，基于可信计算 3.0 的系统安全运行理论，将可信计算平台、可信虚拟动态链、可信免疫架构部署在工控系统计算环境、边界、网络通信、系统管理中，使得漏洞利用无效，为工控系统可信防护提供可行途径，然而可信策略以及可信度量仍需进一步研究。

可信理论被认为是解决复杂异构工业网络安全问题的重要方法，相关研究工作已有开展，Okhravi[4]基于可信网络概念提出了适合于过程控制网络的安全架构，探讨了部件、协议、操作系统以及确保可用性的架构需求分析。Pinto[5]引入 ARM TrustZone 技术作为工业物联网的设备安全防护参考方法，并增强可信执行环境来满足实时性需求。Harshe[6]提出一种可信自主接口保护体系结构(TAIGA)，在可编程逻辑中实现可信组件，与不可信组件隔离并监视物理过程，以检测工控系统的网络和软件重构攻击。Göttert[7]提出一种基于可信硬件支撑可证明安全的分布式可信邻居发现协议（TND），用于监视、检测和定位工业控制系统中的软件配置和控制序列改变的攻击。目前可信技术研究仍停留在普通信息系统领域或工业系统的普通通信网络之中，比较典型的是可信计算平台的应用，尚没有适配的工控系统防护策略的相关报道。

在可信网络度量方面，网络信任评估是可信决策的理论基础，工控系统信任评估的研究处于起步阶段。Fadul[8]将可信理论应用到智能电网的 SCADA 系统安全防护中，利用信任理论过滤恶意节点，在此基础上设计了一个健壮且可配置的信任管理工具箱[9]，基于信任和网络流算法，识别和缓解智能电网故障。Zeng 和 Chow[10]提出基于信誉的内置安全分布式控制方法，来检测和识别分布式网络控制系统中的内部节点不端行为。然而上述算法只是针对智能电网终端设备或机器人控制对象的识别与防护，缺少针对业务操作的控制流和数据流的分析以及异构条件下的信任建模机制。为了处理信任动态评估中相互关联的复杂演算逻辑，需要构建具有普适性的信任评估推理模型。一些基于可信属性的信任评估模型被提出，代表性的包括：基于概率论的信任模型[11]、基于模糊理论的信任模型[12]、基于博弈论的信任模型[13]、基于云理论的信任模型[14]等。现有的信任评估模型在评估的适用性方面存在不足，特别是针对工业网络的结构异构和可用性受限特征，在跨越信息物理域的信任评估，节点和进程的信息流、操作流与交互行为信任建模上，目前缺乏可信评估方法。

3 工控系统业务操作主动免疫机制

工控系统操作流程如图 1 所示。防护的目标主要是保持以控制器、执行器、被控过程、传感器组成的业务操作回路能连续工作，不被非授权的延迟、干扰、修改，保证通过人机界面或远程维护工具的输入不会引起控制器的错误执行或反馈。达到该目标需要保障业务操作的工控系统平台、边界、网络上的操作与交互行为（包括内部误操作）对象和过程的可信性。

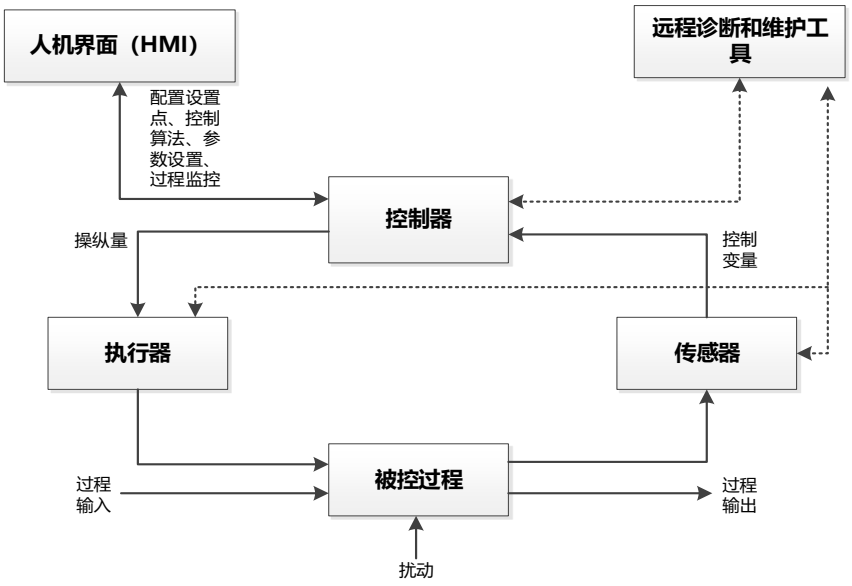

图 1 工控系统操作流程

工控系统的网络架构与威胁如图 2 所示。攻击者可通过外网渗透，依次突破企业防火墙、业务工作站、工程师监控站，通过破坏控制器-执行器-现场设备环路实现信息物理攻击，如“震网”攻击，也可通过维护接口直接从控制系统网络内部发动攻击，或同时对物理侧和信息侧进行攻击和屏蔽，如“乌克兰停电”攻击。还有一种攻击方式是现场工程师有意或无意的误操作。以“乌克兰停电事件”为例，对于跨越信息物理空间的攻击，工控系统固有的安全措施很难应付，SIS 系统与校验机制难以应付合法操作员的不合法或有意的误操作，隔离机制无法防御通过现场维护接口的接入攻击，防火墙检测能力有限，很容易被绕过，且无法过滤未知协议或未知恶意代码。

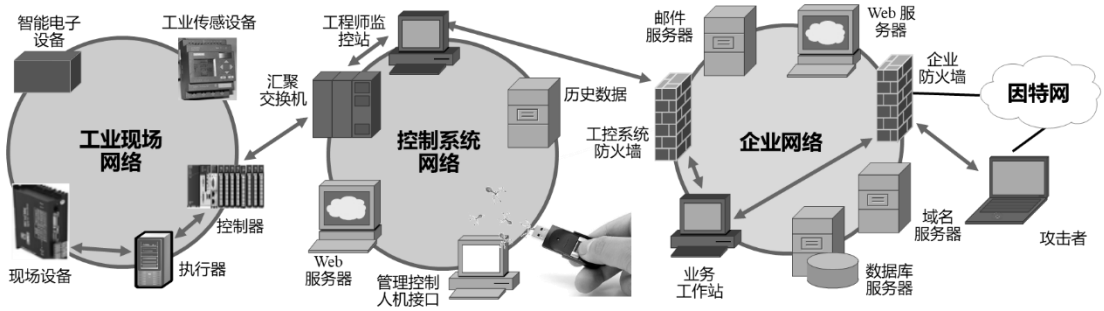

图 2 工控系统网络架构与威胁

工控系统上述脆弱性很容易引发攻击，攻击的目的主要有 4 种：

1）控制指令未授权访问：最典型的攻击目的，利用后门或恶意代码，获取业务操作的控制权限，篡改控制器的指令，执行错误或危险控制。

2）控制操作不可用：以 DDos 攻击为代表，通过对工控系统现场控制器以及其上的工作站的攻击，扰乱正常业务操作，使之无法可靠的提供服务。

3）信息泄露：收集工控系统关键信息、截获关键数据，可在不对工控系统运行造成影响的情况下进行。关键数据的修改以及二次利用会导致控制过程对数据的读写错误。

4）误操作。基于身份认证的操作模式容易引发合法操作员的非法操作，如给现场控制器下达错误指令、直接操作现场设备等行为。

从上述攻击目的看出，工控系统核心数据非法访问与监测、控制软件和指令的修改、内部的恶意行为与误操作等行为不可信是攻击分析、监测、防护的核心。确保在攻击情况下，工控系统按预期运行，使得内外部的恶意访问、修改、误操作等行为自动失效，是主动免疫防护的核心。

因此，定义主动免疫防护如下：

主动免疫防护：确保工控系统正常运行的逻辑组合不被篡改和破坏，系统的运行符合预期。

主动免疫的技术要点是在攻击的场景下，识别核心数据非法访问与监测、控制软件和指令非授权修改、内部恶意行为与误操作等不可信行为，进行屏蔽、阻断、告警，同时确保授权用户或正常操作输出符合预期。梳理业务操作的网络架构、信息流、用户行为集合，解析工控系统平台、边界、网络上的操作与交互行为，面向主动免疫运行目标提取信任信息并对行为信任建模，建立工控系统网络连通性约束下的可信行为决策模型是主动免疫的实现关键。

4 面向业务操作防护的可信行为模型

面向业务操作防护主要保障与业务操作回路相关的数据、指令的访问、通信过程、处理过程等行为的可信，操作回路不被非授权的延迟、干扰、修改，保证通过人机界面或远程维护工具的输入不会引起控制器的错误执行或反馈。

4.1 工控系统业务操作信任建模

工控系统业务操作过程涉及数据流和控制流 2 类主要的信息流，数据流主要包括业务数据、企业数据、管理配置信息，流转路径包括企业管理层的 ERP、数据库、HMI/嵌入式系统/维护设备/终端设备中的存储器，操作动作包括数据的读取、写入、通信、处理等。从一个时间窗口 T 观察数据流转过程，从节点和操作动作 2 个维度建立行为信任。控制流主要包括控制指令、操作信息、控制配置信息，流转路径包括企业管理层的 ERP、工作站/移动工程师站、HMI 组态主机、控制器，操作动作包括控制信息的读取、写入/执行、通信、处理等。也可从点和操作动作 2 个维度建立行为信任。信任信息可在各种嵌入式系统和专用通信协议中的无缝传递。

首先定义一组信任概念。

信任：一定时间窗口内节点对数据流和控制流操作动作的可信性。

信任模型：节点可信性定义、计算、判定、更新方法。

信任管理：基于信任的数据流和控制流访问控制和业务安全控制策略，确保系统主动免疫功能。

信任信息的表示一般采用协议字段方式，如表 1 所示。

表 1 信任基本信息

节点身份	时间	初始信任值	操作动作序列	操作访问阈值	服务能力
Identity	T	T_0	$R_1, R_2,\ldots,R_n$	$T_1, T_2,\ldots,T_n$	*SC*

节点身份：本文设置节点身份为控制流、信息流操作者（包括发起者、传递者、接收者）2 种身份。

时间：考察的控制流、信息流的时间窗口，以此为时间段计算信任信息。

初始信任值：节点加入控制网络时分配的信任值，一般不超过最高信任值的一半，信任值过高容易导致恶意节点攻击，过低会使得有些节点没有表现机会。

操作动作序列：单个节点的操作动作序列，从操作动作细粒度维度考察信任值，用 $R_1, R_2, \cdots, R_n$ 表示。

操作访问阈值：与操作动作对应的访问控制阈值，用 $T_1, T_2, \ldots, T_n$ 表示，通过这些阈值设置访问权限，如某 T_h =1，信任值在 0-1 之间的节点请求访问是不被响应的。

服务能力 *SC*：标注每个节点的服务能力，如带宽、速度限额等，用于应急响应。

信任的建模主要涉及节点操作行为的主观判断及其量化认定，即从观察者角度看，目标对象的操作行为的可信判断依据，以及信任如何量化。节点执行操作会有一些参考属性，从工控系统业务操作防护角度看，可用性>完整性>保密性是整个安全方案防护目标，针对单个节点的信任考察属性，如数据的通信交互，响应时间、通信带宽、交互数量（节点欢迎度）、完整性校验成功率等是影响节点行为是否可信的考察属性。

节点对属性的判断具有一定的模糊性，为便于推理和计算，信任需要量化成数值形式。具体量化值可以根据节点在交易后对属性评分确定，不失一般性，评分的值域范围从 0 到 1，分别表示严重不信任、不信任、最低信任、一般信任、比较信任、完全信任。可信度的判断由量化的属性值的融合得到，最简单的方式是加权平均，属性权重由服务请求者根据个人要求来确定，不关心的属性权重可以设置为 0。节点可信行为会使节点获得更多的交易次数，可信度会良性循环，不可信节点行为会降低信任值，减少交易次数，避免恶意攻击的扩散。

4.2 工控系统业务操作信息收集与异常识别

信任信息收集业务操作过程的数据流和控制流的行为，收集的信息包括历史信息（历史行为序列记录）、当前记录信息。历史信息用于计算操作行为节点当前的信任值，用于决策；当前记录信息用于计算信任的更新。正常的操作按时间窗依次记录。关键在于异常操作的记录

从控制流操作来看，根据上述工控系统上述脆弱性以及攻击方式，梳理面向业务操作防护需求的不可信行为的信任信息采集与建模方法。

1）控制指令未授权访问。控制指令流经网络、节点、进程的信息流、操作流、交互行为等都会对控制指令产生潜在的威胁以及由威胁导致的未授权的获取与修改。

定义数据流的时间窗口 T, 信息采集主要采集时间窗口的节点交互行为流的原始记录。时间分片数量与样本量成正比，时间分片越多，样本越多，建模越精确，但计算量越大。

正常与异常判定：控制指令未授权访问，包括控制指令在未授权的节点或进程中流转；控制指令在非授权用户或进程（读取、写入、通信、处理等）的操作；控制指令在授权用户的不合规交互、修改、截留、欺骗、丢弃等行为，均为异常。其他情况为正常。经过经验或工具无法判断异常的进程将提交入侵检测系统，同时记录为不确定。

2）控制操作不可用：以 DDos 攻击为代表。操作流流经的网络、节点、进程都会对资源敏感的节点发送数据包。控制操作不可用的信任建模方法与控制指令的信任建模类似。

控制操作不可用的正常与异常判定：在控制流层面，发送命令的数据包目的地址持续指向某一个或某几个时间或资源敏感节点；在操作流层面，持续大量的读操作、要求响应操作；交互行为方面，存在选择性转发的行为。以上均为异常。宜调整信任值，根据信任度对攻击节点或进程实施隔离，禁止其参与交互，从而避免系统受到入侵。

3）误操作。基于身份认证的操作模式容易引发合法操作员的非法操作，如给现场控制器下达错误指令、直接操作现场设备等行为。误操作主要考察操作员的非授权或不合理行为。针对操作员节点的交互行为进行信任信息的采集与建模。

收集操作员的交互行为信息，信任建模方法与控制指令的信任建模类似，根据信任度对操作员的非授权或不合理行为及时响应和隔离，禁止其参与交互，从而避免系统受到入侵。

从信息流操作来看，信息非授权篡改和泄露是主要威胁。信息流包括关键操作辅助信息、关键数据信息、企业生产信息等。

1）信息非授权篡改。关键操作辅助信息与关键数据信息流经网络、节点、进程的信息流的操作行为等都会对关键信息产生潜在的威胁以及由威胁导致的未授权的获取与修改。信任建模方法与控制指令的信任建模类似。

正常与异常判定：关键信息未授权访问，包括关键信息在未授权的节点或进程中流转；关键信息在非授权进程的读取、写入、存储等操作；关键信息在授权用户的不合规交互等行为，均为异常。宜根据信任度对信息篡改节点或进程实施隔离，禁止其参与交互，从而避免系统受到入侵。

2）信息泄露：收集工控系统关键信息、截获关键数据、窃取数据库信息和企业管理信息均为异常。对工控系统业务操作来说，信息泄露不会产生直接威胁，但是攻击者可利用窃取的信息，作为攻击的辅助知识。侵害工控企业的知识产权和商业秘密。宜根据信任度对信息泄露节点或进程实施隔离，禁止其参与交互，从而避免系统受到入侵。

所有的异常信息报送审计模块，标记异常节点、异常所属时间段、异常类型、异常内容。

4.3 工控系统业务操作节点信任计算

在工控系统业务操作网中部署信任管理服务器，在每个节点上部署信任计算模块，用于计算和存储信任值，并与信任管理服务器交换信息。利用审计管理模块收集一段时间 T 内的网络节点的交互序列，不失一般性，节点 i 对节点 j 的动态可信值的计算模型如式(1)所示。

$$T_{ij}=\begin{cases}\dfrac{\alpha(t)}{\alpha(t)+\beta(t)}\cdot DT_{ij}+\dfrac{\beta(t)}{\alpha(t)+\beta(t)}\cdot RT_{ij}+RW(t), & N_d(t)<Th_n\\ DT_{ij}+RW(t), & N_d(t)\geq Th_n\end{cases} \tag{1}$$

其中，$\alpha(t)$ 和 $\beta(t)$ 分别用来调整直接信任值和推荐信任值的权重，$RW(t)$ 是奖励信任值。直接信任值和推荐信任值的权重反映了当前节点在时间 T 内对直接信任和推荐信任采信的比重。在考察的时间内当直接交互 $N_d(t)$ 达到一定的次数(门限值)Th_n 时，可以认为交互双方完全建立了信任关系，这时只考虑直接信任，这与人类的心理认知习惯是相符的。RW 奖励信任值由节点持续好评度来确定。一定观察时间 T 内，奖励值为 $RW=0.2\cdot N_{suc}/N$，N_{suc} 为持续成功交互次数，N 为交互总数，即奖励值控制在 0.2 以下，满足加入奖励之后总的 T_{ij} 小于信任评分步长 0.2。

直接信任值 DT_{ij} 的计算：$DT_{ij}=f(R_{ij}, U_{ij}, S_{ij})$，$R_{ij}$、$U_{ij}$ 和 S_{ij} 分别表示节点角色（控制流或信息流），节点交互评价值和节点状态值，节点交互评价值由时间 T 内的成功交互次数与总交互次数的比值，节点状态值根据节点实时状态值得到，由于部分状态值不可量化，采用模糊推理。设信息集合 S ={节点 CPU 利用

率、带宽、能量、传输能力、通信效率}作为论域，设置目标域 V={0-最低信任、0.2-很低的信任、0.5-中等信任、0.7-比较信任、0.9-最高信任}，使用模糊理论推测信任值，隶属度函数可以定义为一个分段函数：

$$\mu_T(x)=\begin{cases}\left(\dfrac{x-t}{1-t}\right)^2, & 0\le x\le\dfrac{a+1}{2}\\ 1-\left(\dfrac{x-t}{1-t}\right)^2, & \dfrac{a+1}{2}\le x\le 1\end{cases} \tag{3}$$

其中 a 是一个定义的门限，给定一个对于 U 的权重向量 W，根据模糊集理论，可以得到最终的信任向量，$V_T=\{v_1, v_2, \ldots, v_m\}$。定义信任 $S_{ij}=\max(V_T)$，其中 $\max(V_T)$表示 T 相对于 V 的最大的隶属度。$f()$表示融合函数，信任值可以直接是三元组形式，也可以采用 U_{ij} 和 S_{ij} 加权和计算得出。

间接信任值 RT_{ij} 由相同的时间段 T 内与节点 j 同处于一个域的节点集合对节点 j 的历史信誉序列的平均值得到。推荐值本质上是第三方节点与节点 j 的直接信任值，计算方法等同直接信任值的计算。

企业管理网络和部分总线网络中的所有节点都可计算相应的信任值，根据可信状态执行访问控制，如大于可信阈值(如 0.5)的节点才能参与服务，从而隔离恶意节点。遍历从服务请求者到服务提供者之间的路径，经过可信节点的路径为有效路径。

在节点交互后，需要对信任进行更新，同时提取审计模块的状态，若没有存在异常报警，则按照固定周期 T 例行检测更新可信值。

若存在异常时，提取标记信息，判断异常节点位置、异常类型（数据流、控制流）、异常内容（读取、写入、通信、处理等），确定异常级别。控制流异常级别确定规则如表 2 所示，信息流的规则类似。由数字 1-3 表示，数字越大，威胁程度越严重。威胁主要根据可用性>完整性>保密性的策略考察 2 个因素：节点位置的重要性或离核心控制器的距离、控制流/信息流操作影响优先级。

表 2 控制流异常级别规则

异常	非授权写入	非授权处理	非授权篡改	非授权通信	非授权读取
1 级节点/1 跳距离	3	3	3	2	2
2 跳距离	3	2	2	2	1
3 跳距离	2	2	2	1	1
4 跳距离	2	2	1	1	1

根据审计模块提供的节点或进程信息，修改其当前的信任值 TV_i，定义 ΔT_i 为

TV_i 和 TV_{i-1} 的差值，如果 ΔT_i <0，则修改信任值 $TV_{i+1}=TV_i+\eta\cdot\Delta T_i$，否则 $TV_{i+1}=TV_i+1/\eta\cdot\Delta T_i$，其中 η 是惩罚因子且 η >1，参考表 2 异常级别值确定。该设置符合人类的交往习惯，即信任的增加是缓慢的，下降是迅速的，对于恶意节点，需要更多次诚实的交易才能恢复到原信任水平。

如果节点存在高威胁异常级别（如节点被攻陷、控制器指令被篡改、控制器无法保障实时性等业务操作无法进行的情况），信任管理服务器直接将该节点的信任值设置到不信任临界点(全网节点的最小的资源访问阈值)以下，并通报全网，同时启动系统修复措施。信任值经过奖惩评估和信任值调整之后，按照 DHT 方式将评价节点信任值存储到对象标识符对应的若干节点上，用于下一轮服务请求。

4.4 工控系统业务操作的可信行为决策方法

部署信任管理服务器，对业务操作的信息流和控制流的节点和行为进行信任标记，以及正常和异常行为下的信任值调整。基于信任值开展决策。决策方式根据异常类型确定，主要考察以下 3 个方面：

1）基于信任值的分级访问控制

针对访问控制型的操作，如外部节点对控制指令未授权访问、DDOS 操作、信息非授权获取等。根据信任值执行隔离操作或访问控制操作，典型访问控制操作如下：

根据节点信任值，对不同级别进程执行不同访问控制策略。定义一组服务决策函数：设总体信任度有 P 个等级划分 $t_1,t_2,...,t_p$，且满足 $t_1<t_2<...<t_p$，$t_i\bigcap t_j=\varnothing(i\neq j)$，对应有 P+1 个级别的服务 $S=\{s_0,s_1,...,s_p\}$，则 T 和 S 之间的服务决策函数可以定义为：

$$S(T)=\begin{cases}s_p, & T>t_p\\ s_{p-1}, & t_{p-1}\le T<t_p\\ \vdots & \vdots\\ s_1, & t_1<T\le t_2\\ s_0, & T\le t_1\end{cases} \tag{4}$$

当节点请求服务（读取、写入、通信、处理等）时，首先根据节点信任度决定它能得到何种质量的服务。例如可定义服务 S={拒绝服务，部分服务，正常服务}，对应的 P={0.2,0.5}，如果用户的信任 T=0.8，是可以得到所需服务的。系统可以根据服务的重要性灵活设置访问信任阈值，来避免系统的访问风险。

在现场网络和监控网络可能没有足够的冗余节点，基于访问控制的信任管

理应保证控制网络关键业务操作不被拒绝。

根据控制网络的信任值和阈值对接入网络的节点或进程、来自上层网络的控制节点、来自下层网络的感知节点进行访问控制，不可信节点拒绝访问或受限访问。采用标记的方法识别关键业务操作，以及对应的最小权限，对关键业务操作指令流经网络、节点进行遍历，若发现存在节点信任值在最小权限访问阈值以下，则对关键操作节点调整信任值，$T_j = T_g \bigcup Th_{R_t}$，$T_g$ 表示当前信任值，T_j 表示调整后信任值，使其满足正常传递的基本信任值，然后对异常节点做如下处理：（1）若存在冗余节点，对比冗余节点信任值（初始值一般在阈值之上），则切换到冗余节点；（2）将关键指令以外的数据加壳或转换成不可操作行为，阻止附带攻击信息的传播；（3）启动预警机制，将异常行为传递到入侵检测模块作进一步检测。

2）基于行为隔离的信任管理

针对内部行为操作，如控制指令未授权操作和虚假通信、信息非授权获取、误操作等。在企业管理网络和冗余带宽的总线网络上部署信任管理机制，根据信任度对恶意行为予以拒绝，屏蔽恶意节点，从而避免系统受到入侵。

基于隔离的信任管理应不影响控制网络的连通性，根据控制网络的信任值和阈值确定可信节点或进程参与服务。根据网络拓扑，对符合需求的节点集合 P 进行遍历，若包含工控系统的关键路径节点，则选择最短路径节点集合 $P_i \subset P$ 作为部署的位置。若不包含工控系统中的关键路径节点，则调整可信阈值 Th_{R_t}，扩大选择范围，选择最短路径节点集合 $P_i \subset P$ 作为部署的位置。若 $R_T \subset R$ 中没有符合防护要求部署能力的要求节点，则保证关键节点覆盖前提下上移防护手段，一般而言，工控现场设备层、监控层、企业管理层的能量能力依次增加，上层可部署更复杂的防护策略和设备。

3）可用性约束优化的信任管理方法

信任管理应保障业务的可用性不受影响或在可控范围，一个关键因素是信任计算与决策不应显著增加延迟而影响系统响应时间。

信任计算与决策应该是轻量级的，对系统的延迟在可控范围。一种可行方式是确保节点部署可信机制的能量需求不超过节点的潜在能力。对于工控系统中节点列表，$R=\{R_1, R_2, \ldots, R_n\}$，$n$ 是节点数，每个节点都具有一个可用资源参数集合，表示为 $R_p=\{p_1, p_2, \ldots, p_m\}$，典型的 3 种属性是计算能力($p_c$)、存储能力($p_s$)、传输能力($p_b$)。计算能力定义为 CPU 剩余计算空间占全部计算空间的比值，存储

能力定义为剩余存储空间占总内存的比值，传输能力为剩余带宽占总带宽的比值。

计算信任管理机制的能量需求。如在部署信任计算模块时，能量需求序列为 $E(R_p) = \{E(p_1), E(p_2),\ldots, E(p_m)\}$。满足如下式的节点才能部署：

$$\{R_i \mid \bigcup_{i=1}^{n}(E(p_1) < \lambda_1 p_1) \bigcap (E(p_2) < \lambda_2 p_2) \bigcap \cdots \bigcap (E(p_m) < \lambda_m p_m)\} \tag{5}$$

其中 $\lambda_1, \lambda_2, \ldots, \lambda_m$ 表示调整系数，如需求能量不超过剩余能量的 80%，则调整系数为 0.8。根据上述要求确定可部署防护策略的部分节点集合 $R_T \subset R$。

若没有满足的节点或部署的信任机制超过节点能力威胁可用性时，则需要调整信任管理方案，一方面优化信任管理方案，降低能量消耗强度。如简化可信评估方法设置（如缩小推荐信任计算范围或简化直接信任值的计算）等；另一方面，保证关键节点覆盖前提下上移防护手段，工控系统现场设备层、监控层、企业管理层的能量能力依次增加，上层可部署更复杂的防护策略和设备。

4.5 优化的信任管理方法

综合上述分析，提出本文的信任管理方案，如图 3 所示。

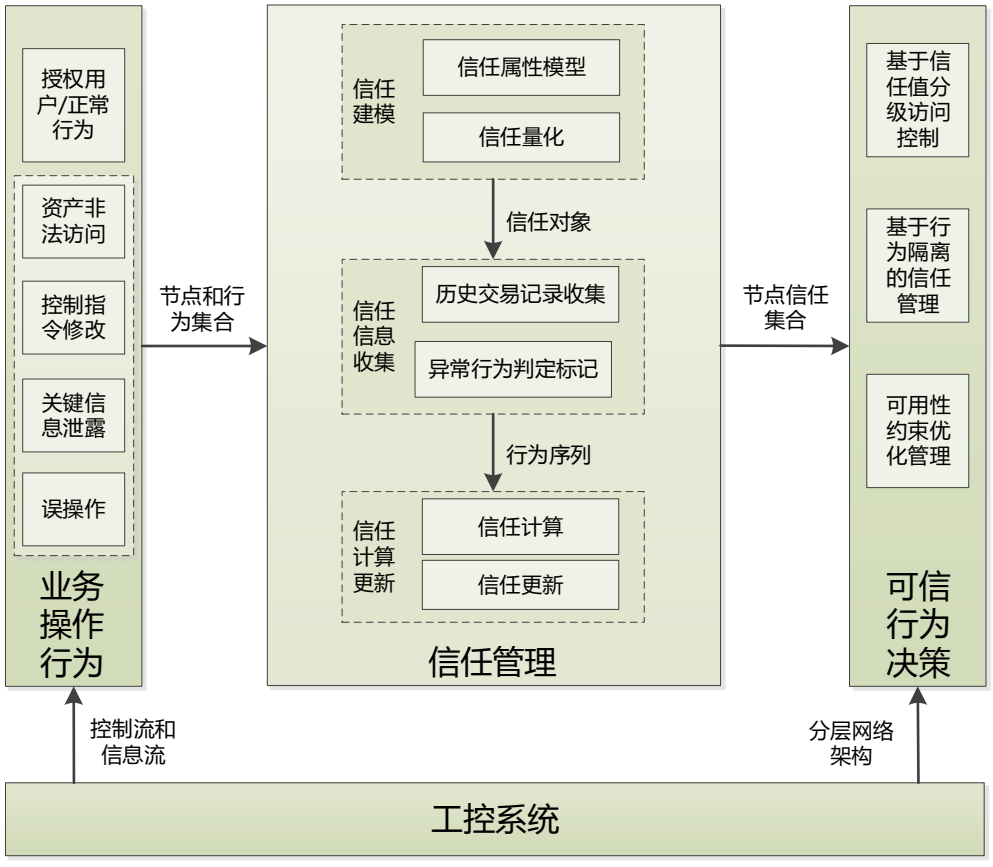

图 3 面向工控业务操作的信任管理方案

所提的信任管理的步骤如下：

1）针对给定的工控系统及其业务流程，识别控制流和信息流以及路径节点，标记关键业务操作以及其路径节点；

2）定义业务操作行为的正常行为和异常行为；

3）基于行为序列集合，对信任进行建模与量化；

4）收集信任信息，包括历史交易记录收集，用于历史信任值计算；异常行为收集、判定、标记，用于分类更新当前信任值；

5）信任计算与更新，得到节点信任集合；

6）基于信任值的决策。针对给定的工控系统分层的网络架构，在现场层和监控层部署基于信任值分级访问控制机制，在部分监控层和企业网络层部署基于行为隔离的信任管理。同时计算工控系统节点能力集合，基于可用性约束，确定 2 种信任决策机制部署的分界线，优化部署信任管理机制。

5 理论分析与仿真实验

5.1 方案功能分析

所提的信任管理方案对典型攻击应对如下：

1）未授权访问：最典型的攻击目的，利用后门或恶意代码，获取控制权限，篡改控制数据，执行错误或危险控制。信任管理机制定义了关键指令和数据的访问控制规则，不可信节点不可访问关键数据，至少是受限访问；不可信用户不可参与传递关键指令，因此，未授权的访问被屏蔽。

2）系统不可用：以 DDos 攻击为代表，通过对工控系统的攻击，使之无法可靠的提供服务。通过基于可信度量的访问控制机制隔离攻击节点。对不符合预先定义的控制流、信息流、行为不予响应，并通过审计管理机制预警阻断，可最大程度缓解攻击。

3）信息泄露：收集工控系统关键信息、截获关键数据，可在不对工控系统运行造成影响的情况下进行。基于访问控制的信任管理机制确保符合预期的可信用户的正常使用数据，不符合预期的访问予以阻断，避免了敏感数据泄露。

4）误操作：可信节点的错误操作。基于行为约束的信任机制能及时发现错误操作行为，通过调整信任值隔离节点或受限访问方式，阻止错误的传播。

所提方案的轻量化主要通过可用性约束函数的能量约束部署来实现,通过能量函数确定节点的剩余能力，与防护策略的预期能力进行匹配。需求的各项能

力（带宽、计算、存储等）均要求低于剩余能力的一部分，且某一方面能力最接近剩余能力，且遍历的路径包括关键路径节点的最短路径的节点集合作为部署的位置。一旦无节点可选，确保关键路径有所防护的基础上，简化信任计算方法。

5.2 方案性能分析

防护方案的时间复杂度：

可信网络计算的复杂度为 $O(n_1^2)$，n_1 为传输节点数。能力匹配算法的计算复杂度为 $O(n_2^2)$，n_2 为拟部署防护方案的节点数目。信任管理的时间复杂度计算：$O(n_1^2) + O(n_2^2)$。监测的规则数目、传递路径节点数量为常数级，信任模块传输节点数、匹配算法计算的节点数目与网络规模相关，因此防护方案时间复杂度为 $2O(n^2) + O(n)$，n 为线性计算复杂度的平均数量。防护方案的空间复杂度：空间复杂度主要是增加了可信管理服务器以及相应的数据存储服务器，内存增加均为线性的，复杂度为 $O(n)$。

5.3 仿真实验

以 DCS 网络为例，设置仿真参数如表 3 所示。N_1，N_2，N_3 表示 DCS 网络中节点数量，N_1 是 DCS 企业层网络大小，N_2 是监控层网络大小，N_3 是现场层网络大小，M_1 表示关键控制指令条数，T_0 是初始信任值，$\alpha_0(t)$和 $\beta_0(t)$是直接信任值和间接信任值的初始权重，Th_n 是直接交互次数门限，P 是访问控制等级数，m 是能力衡量数目，$\lambda_1,\lambda_2,\ldots,\lambda_m$ 采用相同的调整系数 λ。

表 3 仿真参数设置

参数	N_1	N_2	N_3	M_1	T_0	$\alpha_0(t)$	$\beta_0(t)$	Th_n	P	m	λ
值	30	15	10	2	0.5	0.7	0.3	5	3	3	0.8

模拟恶意行为：

1）控制指令未授权访问：模拟从外部 ERP 中某个主机渗透到 DCS 内部的工作站、数据库服务器、传输服务器、PLC 控制主机，试图读取控制权限、窃取核心或敏感数据、篡改关键指令；模拟从内部监控工作站渗透到内网数据库服务器、传输服务器、PLC 控制主机发动攻击，试图获取控制权限、窃取核心或敏感数据、篡改关键指令；模拟误操作，向 PLC 控制主机发送执行命令。路径上部分节点配合恶意节点执行写入、处理、篡改、通信、读取等操作。

2）控制操作不可用：在 ERP 层某几个主机模拟 DDOS 攻击，向监控主机

所在地址发送大量请求消息，监控主机执行读操作使其消耗超过能力阈值。路径上部分节点配合传递消息。

3）信息泄露：在 ERP 层某几个主机模拟信息泄露攻击，非授权处理关键数据、截获数据、虚假反馈、选择性的转发包/丢包等行为。

基于定义的场景和参数，入侵到 DCS 系统 20 次，以读取和破坏控制文件作为攻击成功的标准。仿真结果如图 4 所示。

仿真结果待定

仿真 1：正常操作下信任值分布效果。横坐标为时间（交互次数），纵坐标为节点信任值，对比实验为选取的典型 3 个关键节点的信任值变化曲线。

仿真 2：攻击条件下关键控制指令操作链路上节点信任变化效果。横坐标为时间（交互次数），纵坐标为节点信任值，对比实验为选取的典型 3 个关键节点的信任值变化曲线。

仿真 3：安全性实验/攻击检测和阻断效果。横坐标为攻击次数，纵坐标为免疫成功率，定义为正常执行业务操作次数占总的业务操作次数的比例。针对不同攻击对象（工作站、数据库服务器、传输服务器、PLC 控制主机），面向不同的目的（获取控制权限、窃取数据、篡改关键指令、DDOS、误操作）随机攻击 20 次。对比实验为 2 种攻击（外部渗透、内部渗透）的主动免疫效果。

仿真 4：实时性影响度量/部署前后的延时。横坐标为功能覆盖率，定义为关键节点覆盖比例占总节点的比例，纵坐标为 DCS 系统部署前后关键指令操作与数据读取的延时之和。基于该实验可确定功能覆盖率和延迟的关系。设置监控层最大跳数为 5，对比实验为访问控制防护和隔离防护的效果。

仿真 5：对比实验/免疫效果。横坐标为不同比例恶意节点数目，纵坐标为交互成功率，反应指令和数据无干扰运行的比例。主要针对恶意节点和恶意行为进行屏蔽。对比实验为本文算法、Fadul 方案、Okhravi 方法效果。

仿真 6：对比实验/轻量级效果。横坐标为节点数量，以 10 个节点为初始值，每次增加 5 个，最大节点为 50。纵坐标为防护方案的时间复杂度。对比实验为本文算法、Fadul 方案、Okhravi 方法效果。

6 总结

为解决工控系统的安全防护难题，提出一种面向业务操作防护的基于可信

机制的工业控制系统主动免疫方法。利用工控系统跨域攻击与可用性优先特征下的操作与交互行为信任建模的主动免疫机制，对工控系统控制指令的未授权修改、核心数据非法访问与监测、内部的恶意攻击与误操作等威胁进行有效防护，确保了系统的运行符合预期的主动免疫效果。结合控制网络的连通性需求，根据基于能量的可用性约束函数衡量和调整信任管理方案，解决了工控系统可用性约束问题。分析表明所提方案具有较好的功能、性能以及可控的复杂度，仿真分析表明所提方案的优越性。

主要参考文献

[1] Serpanos D. Secure and Resilient Industrial Control Systems [J]. IEEE Design & Test, 2018, 35(1): 90-94.

[2] Cheminod M., Durante L., Valenzano A. Review of Security Issues in Industrial Networks [J]. IEEE Transactions on Industrial Informatics, 2013, 9(1): 277-293.

[3] Chen D, Li Q. Recent advances on trusted computing in China[J]. Science Bulletin, 2012, 57(35): 4529-4532.

[4] Okhravi H, Nicol D M. Application of trusted network technology to industrial control networks [J]. International Journal of Critical Infrastructure Protection, 2009, 2(3): 84-94.

[5] Pinto S, Gomes T, Pereira J, et al. IIoTEED: An Enhanced, Trusted Execution Environment for Industrial IoT Edge Devices[J]. IEEE Internet Computing, 2017, 21(1):40-47.

[6] Harshe O A, Teja Chiluvuri N, Patterson C D, et al. Design and implementation of a security framework for industrial control systems[C]// International Conference on Industrial Instrumentation and Control. IEEE, 2015:127-132.

[7] Göttert N, Kuntze N, Rudolph C, et al. Trusted neighborhood discovery in critical infrastructures[C]// IEEE International Conference on Smart Grid Communications. IEEE, 2015: 976-981.

[8] Fadul J, Hopkinson K, Sheffield C, et al. Trust Management and Security in the Future Communication-Based "Smart" Electric Power Grid[C]. Proceedings in System Sciences (HICSS), 2011 44th Hawaii International Conference on IEEE, 2011: 1-10.

[9] Fadul J E, Hopkinson K M, Andel T R, et al. A Trust-Management Toolkit for Smart-Grid Protection Systems[J]. IEEE Transactions on Power Delivery, 2014, 29(4):1768-1779.

[10] Zeng W, Chow M Y. A Reputation-Based Secure Distributed Control Methodology in D-NCS[J]. IEEE Transactions on Industrial Electronics, 2014, 61(11): 6294-6303.

[11] Begriche Y., Khatoun R., Khoukhi L., et al. Bayesian-based model for a reputation system in vehicular networks [C]. SSIC 2015, shanghai, China: IEEE, 2015: 1-6.

[12] Jadidoleslamy H., Reza Aref M., Bahramgiri H. A fuzzy fully distributed trust management system in wireless sensor networks [J]. AEU- International Journal of Electronics and Communications, 2016, 70(1): 40-49.

[13] Feng R, Che S, Wang X, et al. An incentive mechanism based on game theory for trust management[J]. Security & Communication Networks, 2015, 7(12):2318-2325.

[14] Jiang J, Han G, Shu L, et al. A Trust Model based on Cloud Theory in Underwater Acoustic Sensor Networks[J]. IEEE Transactions on Industrial Informatics, 2017, 13(1): 342-350.