控制与决策
Control and Decision

网络攻击下多智能体系统一致性安全与隐私保护研究综述

卢剑权, 邢梦平, 张晶

引用本文:
卢剑权, 邢梦平, 张晶. 网络攻击下多智能体系统一致性安全与隐私保护研究综述[J]. 控制与决策, 2025, 40(11): 3201-3219.

在线阅读 View online: https://doi.org/10.13195/j.kzyjc.2025.0291

网络攻击下多智能体系统一致性安全与隐私保护研究综述

卢剑权[1†]，邢梦平[2]，张 晶[3]

(1. 东南大学 数学学院，南京 210096；2. 国防科技大学 电子对抗学院，合肥 230037；
3. 南京师范大学 数学科学学院，南京 210023)

摘 要: 多智能体系统通过智能体间的信息交互与协调实现整体目标, 其一致性研究不仅揭示了自组织与分布式控制的特点, 也在计算机科学和人工智能等领域展现出重要的应用价值. 智能体间的相互作用高度依赖网络通信条件, 而数据传输过程中网络攻击的频繁发生对系统的安全稳定运行提出了严峻挑战. 鉴于此, 总结网络攻击下多智能体系统一致性安全与隐私保护的研究进展. 首先, 简要介绍攻击者对信息传输的保密性、完整性以及可用性产生的影响, 揭示系统在窃听攻击、虚假数据注入攻击和 DoS 攻击这 3 类典型的攻击模式下的脆弱性; 然后, 针对不同网络攻击类型的特点, 从防御者的角度对多智能体一致性研究中的安全控制与隐私保护问题进行讨论, 并进一步梳理应对攻击常用的检测与防御保护机制; 此外, 综合讨论不同攻击在分布式控制相关研究分支的发展现状; 最后针对多智能体系统一致性安全与隐私保护研究中存在的技术难题与挑战进行分析, 并对未来研究方向进行展望.

关键词: 多智能体系统；网络攻击；隐私保护；安全控制；一致性；分布式控制

中图分类号: TP13 **文献标志码:** A

DOI: 10.13195/j.kzyjc.2025.0291

引用格式: 卢剑权, 邢梦平, 张晶. 网络攻击下多智能体系统一致性安全与隐私保护研究综述 [J]. 控制与决策, 2025, 40(11): 3201-3219.

A survey on secure and privacy protection of multi-agent systems consensus under cyber attacks

LU Jian-quan[1†]，*XING Meng-ping*[2]，*ZHANG Jing*[3]

(1. School of Mathematics, Southeast University， Nanjing 210096， China； 2. College of Electronic Warfare，National University of Defense Technology，Hefei 230037，China；3. School of Mathematical Sciences，Nanjing Normal University，Nanjing 210023，China)

Abstract: Multi-agent systems achieve overall objectives through information interaction and coordination among agents. Research on consensus of multi-agent systems not only reveals the characteristics of self-organization and distributed control but also demonstrates significant practical application value in fields such as computer science and artificial intelligence. However, as interactions among agents heavily rely on network communication conditions, the frequent occurrence of cyber attacks during data transmission poses severe challenges to the secure and stable operation of systems. Therefore, this paper aims to summarize the research progress on consensus security and privacy protection of multi-agent systems under cyber attacks. First, the impact of attackers on the confidentiality, integrity, and availability of information transmission is briefly illustrated, highlighting the vulnerabilities of systems under three typical attack patterns: eavesdropping attacks, false data injection attacks, and denial-of-service (DoS) attacks. Subsequently, based on the characteristics of different types of cyber attacks, the paper discusses secure control and privacy protection issues in consensus research from the defender's perspective, while further summarizing commonly used detection and defense mechanisms. Additionally, the study comprehensively reviews the development status of distributed control-related research branches, focusing on secure control strategies under false data injection attacks and DoS attacks, as well as the design of consensus privacy protection algorithms under eavesdropping attacks. Finally, the technical challenges and issues in the research of consensus security and privacy protection for multi-agent systems are analyzed and research directions worth exploring are further outlined.

收稿日期: 2025-03-20；**录用日期:** 2025-06-09.

基金项目: 江苏省自然科学基金攀登项目 (BK20240009)；国家自然科学基金面上项目 (623731050)；江苏省应用数学科学研究中心项目 (BK20233002).

[†]通信作者. E-mail: jqluma@seu.edu.cn.

Keywords: multi-agent systems；cyber attacks；privacy protection；secure control；consensus；distributed control

0 引 言

多智能体系统是人工智能领域的一个重要分支, 它指由多个智能体及相应的组织规则和信息交互协议构成, 能够完成特定任务的一类复杂系统[1-2]. 组成系统的各个智能体通过相互通信、合作、竞争等方式, 完成单个智能体不能完成的、大量而又复杂的工作. 多智能体系统的出现克服了单一智能体在信息处理和执行、传感和通信等方面的局限性, 对解决大型、复杂、分布式及难预测问题意义重大. 多智能体系统的应用范围极为广泛, 涵盖了工业、商业、医疗、交通、军事等多个关键领域[3-5]. 随着人工智能和物联网技术的持续进步, 多智能体系统的应用前景将不断拓展, 其应用范围也将日益扩大.

近年来, 多智能体系统的分布式协同控制问题受到了来自不同领域学科专家的广泛关注, 它利用邻域的局部信息, 为每个智能体设计简单的控制策略, 使网络化系统能够完成共同的任务或实现预定的集体目标[6-7]. 其中, 一致性控制问题作为多智能体之间协同合作的基础, 是传感器融合、智能电网经济调度等应用的有效工具[8-9]. 对多智能体系统的一致性进行研究不仅可以推动相关技术的发展, 而且可以为其他领域提供新的思路和方法.

多智能体系统在完成共同任务或实现集体目标时, 对数据通信高度依赖, 数据的高效可靠传输是保障系统安全稳定运行的关键[10-12]. 然而, 随着现代化通信网络的使用, 开放的数据传输环境使得系统面临严重的网络攻击威胁. 同时系统的分布式传输也使得攻击能通过交互作用相互影响, 导致攻击带来的故障问题进一步扩大[13-15]. 攻击者一旦突破防御措施成功实现攻击目标, 就有可能对系统带来灾难性损害, 进而造成难以估计的社会和经济损失. 近年来, 随着计算机和通信技术的飞速发展, 网络安全事件频发, 带来的危害和影响也日益加剧. 2010 年, 针对工业控制系统的 Stuxnet 蠕虫病毒入侵了伊朗核设施, 造成数百台离心机失效, 该事件标志着国家间网络战争的开始[16]; 2013 年, 黑客通过入侵 Target 与第三方供应商之间的网络窃取了几千万客户的信息数据, 造成大规模数据泄露, 带来了巨额经济损失[17]; 2015 年, 攻击者经过长期潜伏准备, 对乌克兰电网采用了复杂的多阶段攻击, 使得数十万用户长时间失去电力供应[18]; 2023 年 DDoS 攻击导致 OpenAI 的 API 和 ChatGPT 服务在 24 小时内出现周期性中断, 数以万计的用户无法正常访问 ChatGPT 服务. 表 1 列举了近年来世界各地发生的重大网络安全事件. 从以上安全事件可以看出, 网络攻击对社会经济发展具有极大的破坏力, 甚至会对人身安全构成严重威胁. 因此, 在研究多智能体系统的一致性控制问题中, 考虑网络攻击对数据传输的保密性、完整性以及可用性的影响具有重要的现实意义.

表1 近年来世界各地发生的重大网络安全事件

年份	地点	安全事件
2010	伊朗	针对核武器计划的Stuxnet蠕虫病毒攻击
2015	乌克兰	Black Energy软件对电网的攻击
2021	美国	控制水处理设施的计算机被黑客渗透
2022	德国	约两千台涡轮机的控制系统受攻击瘫痪
2023	美国	ChatGPT因遭受DDoS攻击而服务中断
2024	以色列	40家重要组织受到数据擦除器攻击

根据网络攻击的实现方式和攻击目的, 目前针对多智能体系统的典型攻击类型有 3 种: 窃听攻击、欺骗攻击以及 DoS 攻击[19-21], 分别会破坏数据的保密性、完整性以及可用性. 其中, 窃听攻击主要利用网络漏洞窃取多智能体之间交互的数据, 导致敏感信息泄露[22]; 欺骗攻击主要通过发送虚假数据或篡改数据, 影响智能体之间的协作与决策, 通常包括虚假数据注入攻击和重放攻击为特例[23-24]; DoS 攻击通过大量的请求使智能体之间的通信网络超负荷, 或使智能体的计算资源被大量消耗, 使得数据传输的实时性和可用性被严重影响, 进而导致正常的服务请求无法被及时响应[25].

多智能体系统一致性安全问题的研究主要围绕安全控制和隐私保护展开. 在虚假数据注入攻击和 DoS 攻击影响下, 一致性控制方案很容易受此影响而失效, 使得智能体的状态失控. 对此, 需要引入合适的攻击检测和防御保护机制, 使异常数据或伪造数据被有效识别处理. 此外, 需要设计鲁棒控制方案以及容错算法, 提高系统的抗攻击能力, 使得智能体数据传输的完整性和可用性被破坏时, 系统仍然可以保持一定的正常功能. 针对多智能体系统在虚假数据注入攻击下的一致性问题, 攻击者通常致力于设计能最大化一致性误差的攻击模型[26], 而保护者的研究重点通常集中在以下 3 个方面: 设计攻击检测与隔离方案, 使得系统在发生虚假数据注入攻击时, 检测器能准确识别攻击并定位异常链路, 进而将

目标攻击智能体与被攻击链路隔离[27]; 建立防御机制如引入冗余通信链路, 使得数据传输被攻击者破坏时, 额外添加的链路能被激活用于通信, 以提高数据传输的可靠性[28]; 设计鲁棒控制方案, 使得控制器在系统存在虚假数据注入攻击时仍能有效作用于智能体, 使其实现一致性目标[29]. 此外, 相关的安全估计与控制研究在实际情境 (如电力系统) 中的应用问题也得到了深入探索[23,30-33]. 对于多智能体系统在DoS 攻击下的一致性问题, 一方面大量研究致力于建立合适的数学模型刻画 DoS 攻击发生的特点, 如周期/脉冲宽度调制信号、非周期信号、伯努利分布、马尔科夫过程、平均驻留时间概念等[34-40]; 另一方面DoS 攻击下的弹性一致性问题也被众多学者广泛讨论[41-44]. 在隐私保护方面, 研究者们旨在设计有效合理的隐私保护方案, 以确保智能体的敏感信息得到充分地保护, 同时保证系统的一致性目标和收敛精度. 关于多智能体系统的隐私保护, 当前研究主要集中在两个方面: 一方面是基于不同的隐私保护方法实现一致性目标, 包括但不限于同态加密方法[45]、差分隐私方法[46]、状态分解方法[47] 等; 另一方面, 除了从不同的隐私保护方法视角入手, 研究者们还从一致性问题本身入手, 基于不同的一致性目标进行隐私保护一致性的探究, 如有限时间一致性[48-50]、量化一致性[51] 和二部一致性[52-53] 等.

如图 1 所示, 本文以多智能体一致性问题为研究中心, 围绕网络攻击对数据传输以及系统运行的影响展开讨论, 阐述了多智能体系统在网络攻击下一致性安全与隐私保护的研究现状. 首先说明多智能体系统一致性中考虑数据传输安全性问题的重要性; 随后对信息安全研究中常见的攻击类型进行简单介绍, 列出应对攻击常用的检测与防御保护机制; 基于上述讨论, 对虚假数据注入攻击和 DoS 攻击下多智能体系统的一致性安全控制的研究现状进行分析, 并对多智能体系统一致性问题研究中的隐私保护问题的研究进展进行阐述, 表明多智能体系统一致性问题中考虑安全控制与隐私保护的重要意义. 最后, 针对安全控制方案建立与隐私保护算法设计等方面的难题和挑战, 对未来可能发展的方向做出展望.

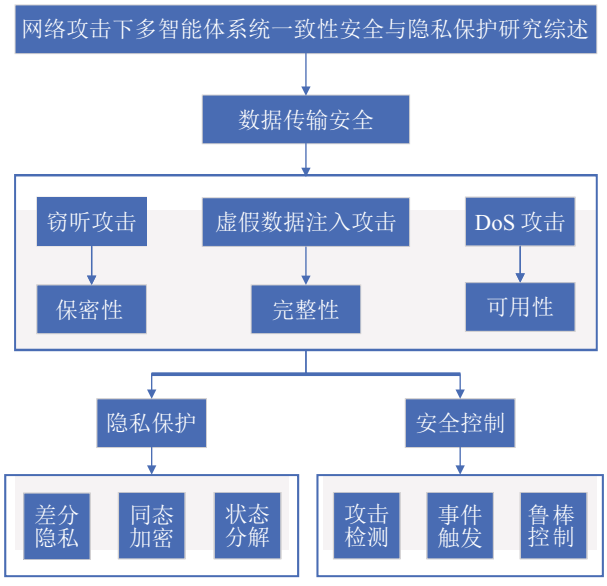

图1　网络攻击下多智能体系统的研究框图

1　多智能体系统一致性问题描述

考虑一个具有N个智能体的多智能体系统, 每个智能体i $(i=1,2,\ldots,N)$在时刻t的状态表示为$x_i(t)$, 其中$x_i(t)$可以是一个标量 (如位置), 也可以是一个向量 (如位置和速度的组合). 在多智能体系统中, 智能体之间的通信关系可以通过图论来描述. 定义一个有向图或无向图$G=(V,E)$. 其中: $V=\{1,2,\ldots,N\}$表示所有智能体的集合, E表示智能体之间的通信链路的集合. 如果智能体i能够接收到智能体j的信息, 则称智能体j是智能体i的邻居, 记作$j\in\mathcal{N}_i$. 矩阵$A=[a_{ij}]\in\mathbb{R}^{N\times N}$称为图的权重邻接矩阵, 元素$a_{ij}$与图中的边相关, 即$(j,i)\Leftrightarrow a_{ij}>0$, 否则$a_{ij}=0$. 拉普拉斯矩阵$L=[l_{ij}]\in\mathbb{R}^{N\times N}$定义为$l_{ii}=\sum_{k=1}^{N}a_{ik}, i=1,2,\ldots,N$, $l_{ij}=-a_{ij}, i\neq j$.

一致性协议描述了智能体之间如何交换信息并更新自己的状态. 文献 [54] 建立了一致性基本框架, 将图论、矩阵论、李雅普诺夫稳定性理论等相关知识引入到多智能体系统一致性的研究中, 为一致性研究打下理论基础. 首先构建每个智能体的连续动态方程为

$$\begin{cases}\dot{x}_i(t)=u_i(t),\\ y_i(t)=C_ix_i(t),\ i=1,2,\ldots,N.\end{cases}\tag{1}$$

其中: $\dot{x}_i(t)$为第i个智能体在时刻t的状态变化率; $u_i(t)$为智能体i在时刻t的控制输入; $y_i(t)$为第i个智能体在时刻t的测量输出, 对应的实际传输信号用$z_i(t)=D_ix_i(t)$表示.

离散时间动态方程为

$$\begin{cases}x_i(t+1)=x_i(t)+\epsilon u_i(t),\\ y_i(t)=C_ix_i(t),\ i=1,2,\ldots,N.\end{cases}\tag{2}$$

其中: $\epsilon\in(0,1/d^*)$为更新步长, $d^*=\max\limits_{i\in V}\{l_{ii}\}$.

控制输入$u_i(t)$通常与智能体i的邻居智能体的状态有关, 具体形式取决于一致性协议的设计. 经典的控制协议[54] 一般设计为

$$u_i(t)=\sum_{j\in\mathcal{N}_i}a_{ij}(x_j(t)-x_i(t)). \tag{3}$$

关于多智能体系统的一致性目标实现, 常用以下两个公式化定义.

定义 1[54] 对于任意智能体i和j, 若满足

$$\lim_{t\to\infty}||x_i(t)-x_j(t)||=0, \tag{4}$$

则称多智能体系统实现一致. 即对于任意两个智能体i和j, 其状态差随着时间的推移趋于 0.

定义 2[54] 对于任意智能体i, 若有

$$\lim_{t\to\infty}x_i(t)=c,\ i=1,2,\ldots,N, \tag{5}$$

则称多智能体系统实现一致, 其中c为任意实数或实向量. 特别地, 若$c=\dfrac{1}{N}\sum\limits_{i=1}^{N}x_i(0)$, 则称系统实现平均一致; 若$c=\max\limits_{i=1,2,\ldots,N}(\min)\{x_i(0)\}$, 则称系统实现了最大 (最小) 一致.

多智能体一致性问题模型包括智能体状态、邻居关系与通信拓扑、一致性协议以及一致性条件等方面, 这些元素共同构成了多智能体一致性问题的基本框架, 为后续研究和应用提供了基础.

2 信息安全三要素

智能体实现一致性的过程中, 与邻居之间的信息传输起关键性的作用, 信息传输中信息安全的三要素是保密性、完整性和可用性.

2.1 保 密 性

保密性是指确保信息只能被授权的人或实体访问和使用, 防止信息泄露给未经授权的用户或实体. 这要求信息在保存、使用和传输过程中不被泄露给非授权方. 保密性主要由加密技术保证, 常用的保密技术包括信息加密、访问控制、身份验证和网络安全措施等[55-56]. 其中, 信息加密是在密钥的控制下, 用加密算法对信息进行加密处理, 即使对手得到了加密后的信息也会因没有密钥而无法读懂有用信息.

随着各种技术的不断进步和应用需求的不断扩大, 多智能体系统中一致性算法也被要求提供更加严格的安全性保障. 经典的一致性算法要求智能体与它们的邻居显式共享状态, 忽略了保密性的要求. 在当今数据大爆炸时代, 保护个体信息的隐私尤其重要[57]. 一致性问题中的保密性主要是指每个智能体的状态最终实现一致, 同时还保证敏感信息的隐私性. 这里的敏感信息可以是交通系统中用户的位置轨迹[58]、社交网络中个体的初始观点[59]、多智能体会合问题中智能体的初始位置[60] 等. 因此, 在多智能体一致性算法中考虑隐私保护具有重要意义.

2.2 完 整 性

多智能体系统信息通信的完整性是指智能体之间或传感器与控制器等组件之间, 传输的数据不会被篡改、伪造或丢失, 即接收方收到的数据与发送方发送的数据完全一致. 一方面, 通过引入校验码、哈希函数以及数字签名等技术, 可以检测并标识数据在传输过程中的更改[61-62]; 另一方面, 基于对数据的概率统计分析, 引入卡方检测或 K-L (Kullback-Leibler) 散度检测[63-64], 可以识别数据分布的异常.

在多智能体系统中, 对数据完整性的破坏可能会使智能体根据错误信息做出误导性决策, 例如在无人机编队中, 某个无人机的位置或运动速度信息被篡改, 可能导致整个编队以错误的运行速度偏离预定航线, 进而增加智能体碰撞以及任务失败的风险. 因此, 在多智能体系统的实际应用中, 如无人机集群、智能交通、智能制造等领域, 需要采取有效的安全措施, 如数据加密、认证、检测等, 确保数据传输过程中的完整性, 减少错误数据对系统稳定运行的影响.

2.3 可 用 性

多智能体系统信息通信的可用性是指接收方 (智能体或组件) 有效访问和使用数据的能力. 相对而言, 完整性主要关注传输数据的准确性和一致性, 确保数据在传输和存储过程中不被篡改, 而可用性则关注数据的可访问性, 确保在需要时数据能被有效使用.

对于多智能体系统, 其个体之间相互协作完成整体任务的关键是各个智能体能够访问和使用必要数据. 不论是邻居智能体传输的数据还是传感器传送到控制器的数据, 都在系统协作与决策过程中发挥重要作用, 数据可用性被破坏会对系统的正常运行和任务执行带来严重影响. 为了应对数据可用性破坏的威胁, 可以采取多种检测和防御机制. 例如: 建立冗余通信或多路径传输方案, 提高通信链路的可靠性[28]; 建立故障检测机制实时监测智能体的通信状态, 一旦检测到数据不可用, 则切换到备用机制 (如利用零阶保持器或在接收端引入新的数据更新迭代规则等)[65-66].

3 常见的网络攻击

3.1 窃听攻击

窃听攻击作为网络安全威胁的一种, 属于网络攻击中的被动攻击. 它是指攻击者在数据传输过程中拦截、截获和窃取通信内容的攻击行为, 其通常很难检测, 因为它不会主动向网络注入攻击数据, 不会

改变网络的流量特征, 也不会触发入侵保护系统的预警. 窃听攻击的危害极大, 它可能导致个人隐私泄露、商业机密被盗取、敏感信息被篡改或破坏等严重后果. 针对多智能体一致性问题中可能存在的窃听攻击, 根据其在网络中的位置分为以下两种[47]:

1) 诚实但好奇的智能体. 位于多智能体网络内部, 正确遵循所有协议步骤, 但试图根据收集的数据获取其他智能体的信息. 诚实但好奇的智能体拥有的信息是自己的内部状态和从邻居处接收到的信息.

2) 外部窃听者. 位于多智能体网络外部, 可以通过入侵通信信道来访问信息. 窃听者可以窃取网络拓扑和网络中所有共享的数据, 但无法获取没有在网络中共享的状态.

3.2　欺骗攻击

欺骗攻击的主要目的是破坏数据的完整性, 攻击者通过篡改或伪造个体或组件之间传输的数据, 让系统在接收到伪造数据后做出错误的判断或决策, 从而影响系统的一致性或决策. 例如在智能交通系统中, 车辆如果接收到来自邻居车辆带有欺骗性的错误位置信息, 则可能导致碰撞事故即交通混乱问题. 通常欺骗攻击模型的设计目标有两个: 对检测机制保持隐形, 以及给系统正常运行带来最大破坏. 根据篡改数据方式的不同, 目前针对多智能体系统的欺骗攻击类型主要有两种: 重放攻击和虚假数据注入攻击[67-68], 对应的数学模型如下:

$$z_i^{\tau}(t)=\sum_{h_i^t=0}^{\tau_i(t)}\alpha_i(h_i^t)z_i(t-h_i^t), \tag{6}$$

$$z_i^{a}(t)=f_i^t(z_i(t))+a_i(t). \tag{7}$$

其中: $z_i(t)$为智能体i在时刻t的传输信号, $z_i^{\tau}(t)$和$z_i^{a}(t)$分别为对应接收端在t时刻获得的重放攻击信号和虚假数据注入攻击信号; $\tau_i(t)\geqslant 0$为攻击者i在时刻t可获得历史待传输数据的窗口长度; $\alpha_i(h_i^t)\in\{0,1\}$, 当$\alpha_i(h_i^t)=0$时, 表示攻击者未利用历史数据$z_i(t-h_i^t)$生成重放攻击信号; $f_i^t(\cdot)$为虚假数据注入攻击的攻击函数, 通常可通过引入攻击比例矩阵$T_i(t)$设计成$T_i(t)z_i(t)$的形式; $a_i(t)$是注入的随机噪声信号.

由式 (6) 可以看出, 重放攻击是攻击者截获合法的通信数据, 然后重新发送截获的历史数据包, 以此扰乱系统的正常行为. 这一过程并没有对原始数据进行修改, 而是通过重复旧数据来实现干扰目的, 因此更容易规避检测器的检测, 但由于利用的是系统的已有数据, 对系统正常运行的破坏性不够理想. 针对这一点, 如式 (7) 所示的虚假数据注入攻击可以直接伪造或篡改通信数据, 其设计上通常具有一定难度, 例如需要获知系统的部分信息, 才能有针对性地设计出对检测器保持隐形且对系统产生最大破坏的攻击形式, 但是其具有更高的灵活性, 可以根据攻击目标, 直接更改传输数据, 导致系统做出错误的决策和控制. 为了提升系统的安全性和稳定性, 需要引入合适的攻击检测与防御机制来识别处理潜在的恶意行为和数据异常. 针对重放攻击目前常用的防御措施有: 在传输数据中加入时间戳或序列号, 或者对传输数据进行加密处理并在接收端进行身份认证. 而对于虚假数据注入攻击, 则可以利用数据完整性校验 (如哈希函数、数字签名)、基于统计的异常检测方法 (如卡方检测、K-L 散度检测) 以及加密认证 (对称加密、公钥加密) 等检测防御措施进行处理.

3.3　DoS 攻击

DoS 攻击通常通过消耗被攻击对象的资源, 让信息传输时发生拥堵而无法正常通信, 由于这种攻击方式的成本低且形式简单易于实现, 在实际情况中是网络安全最主要且最持久的威胁[69]. 不同于欺骗攻击对传输数据的篡改或伪造, DoS 攻击会导致其作用期间数据的直接丢失, 从而破坏数据的可用性, 因此相关研究更集中于探索攻击频率或持续时间对系统稳定性的不利影响. 目前, 对于 DoS 攻击模型的刻画方式主要有脉冲宽度调制、伯努利分布、马尔科夫过程、平均驻留时间等, 其具体形式如下所述:

1) 脉冲宽度调制刻画的 DoS 攻击模型[34]. 假设t时刻智能体j向智能体i的待传输信号为$z_{ij}(t)$, 智能体i实际接收到的来自智能体j的信号为$z_{ij}^a(t)$, 引入一组如下所示的能量受限的干扰信号$\beta_{ij}(t)$来表示 DoS 攻击对数据传输的破坏:

$$\beta_{ij}(t)=\begin{cases}0,\ T_{ij}^{n-1}\leqslant t<T_{ij}^{n-1}+T_{oij}^{n-1};\\ 1,\ T_{ij}^{n-1}+T_{oij}^{n-1}\leqslant t<T_{ij}^{n}.\end{cases} \tag{8}$$

其中: 区间$[T_{ij}^{n-1},T_{ij}^{n-1}+T_{oij}^{n-1})$表示在此时间段内, 智能体$j$到$i$的信道受 DoS 攻击影响, 数据无法正常传输; 区间$[T_{ij}^{n-1}+T_{oij}^{n-1},T_{ij}^{n})$表示攻击停止, 智能体$j$到$i$能正常通信. 因此, 智能体$i$实际接收的来自智能体$j$的信号可进一步表示为

$$z_{ij}^a(t)=\beta_{ij}(t)z_{ij}(t). \tag{9}$$

2) 伯努利分布刻画的 DoS 攻击模型[36]. 将式 (9) 中$\beta_{ij}(t)$重新定义为以下形式, 则其变为伯努利分布刻画的 DoS 攻击模型: $\beta_{ij}(t)\in\{0,1\}$为服从伯努利分布的随机变量, 当其取值为0时表示通信信道中有

攻击, 取值为1时表示没有攻击, 且取值满足条件

$$\Pr\{\beta_{ij}(t)=1\}=\bar{\beta}_{ij},\ \Pr\{\beta_{ij}(t)=0\}=1-\bar{\beta}_{ij}. \tag{10}$$

其中: $\bar{\beta}_{ij}\in(0,1]$, $\Pr\{\cdot\}$为事件发生的概率.

3) 马尔科夫过程刻画的 DoS 攻击模型[37,70]. 在用马尔科夫过程刻画 DoS 攻击发生的特点时, 针对其直接阻断智能体之间通信的特点, 在 DoS 攻击下智能体之间的信息交互可以由时变拓扑来描述. 引入马尔科夫过程$s(t)\in\{1,2,\dots,M\}$描述通信拓扑在 DoS 攻击下的转移概率, 其中$\{1,2,\dots,M\}$对应M个时变拓扑图$\{G_1,G_2,\dots,G_M\}$, 则t时刻对应的通信拓扑为$G_{s(t)}$. 而对于马尔科夫过程$s(t)$, 其跳变可以由转移速率矩阵$\Pi=[\pi_{uv}]_{M\times M}$确定, 具体地, 有

$$\Pr\{s(t+h)=v|s(t)=u\}=\begin{cases}\pi_{uv}h+o(h),\ u\neq v;\\ 1+\pi_{uu}h+o(h),\ u=v.\end{cases} \tag{11}$$

其中: $o(h)$为h的高阶无穷小, 即$\lim\limits_{h\to0}o(h)/h=0$, $h>0$; 若$u\neq v$, 则$\pi_{uv}\geqslant0$为$s(t)$从模态u跳变到v的转移速率, 而$\pi_{uu}=-\sum\limits_{v=1,u\neq v}^{M}\pi_{uv}$.

4) 平均驻留时间刻画的 DoS 攻击模型[38]. 利用 Hespanha 和 Morse 在文献 [71] 中提出的平均驻留时间概念描述的 DoS 攻击模型, 由于其在刻画攻击发生频率和持续时间特征上的优越性, 近年来得到广泛研究. 对于智能体j到i之间的通信链路, 类似上述引入变量$\beta_{ij}(t)\in\{0,1\}$表示 DoS 攻击信号, 该链路的第l个 DoS 攻击间隔表述为

$$H_{ij}^{l}\triangleq\{h_{ij}^{l}\}\bigcup[h_{ij}^{l},h_{ij}^{l}+\xi_{ij}^{l}). \tag{12}$$

其中: $\{h_{ij}^{l},l=0,1,\dots\}$为 DoS 攻击发起时刻序列集合, $\{\hat{h}_{ij}^{l}=h_{ij}^{l}+\xi_{ij}^{l},l=0,1,\dots\}$为对应的攻击暂停时刻序列集合, $\xi_{ij}^{l}\geqslant0$为对应的攻击区间长度. $n_{ij}(t_1,t_2)\triangleq\mathrm{card}\{m|h_{ij}^{m}\in[t_1,t_2)\}$为攻击者针对智能体$j$到$i$的链路, 在时间段$[t_1,t_2)$内发起的总攻击次数. 此外, 在此区间内, 将 DoS 攻击区域的并集表示为$\Xi_{ij}(t_1,t_2)\triangleq[t_1,t_2]\bigcap(\bigcup\limits_{l\in\{1,2,\dots\}}H_{ij}^{l})$, 而非攻击区域的并集为$\bar{\Xi}_{ij}(t_1,t_2)\triangleq[t_1,t_2]-\Xi_{ij}(t_1,t_2)$. 并且将 DoS 攻击的总并集表示为$\mathcal{T}_{ij}=\bigcup\limits_{l\in\{1,2,\dots\}}H_{ij}^{l}$, 而$|\Xi_{ij}(t_1,t_2)|$和$|\bar{\Xi}_{ij}(t_1,t_2)|$分别表示对应的区域长度. 若$t\in\mathcal{T}_{ij}$, 则$\beta_{ij}(t)=0$, 表示$t$时刻智能体$j$到$i$之间的通信被 DoS 攻击破坏, 否则$\beta_{ij}(t)=1$表示无攻击影响. 基于以上符号说明及平均驻留时间的概念, 文献 [38] 给出以下两个假设对攻击发生的频率和持续时间进行限制.

假设 1[38] (DoS 攻击频率) 对于给定的标量$\eta_{ij}\geqslant1,\tau_{ij}^{D}>0$, DoS 攻击在时间段$[t_1,t_2)$内发生的攻击次数$n_{ij}(t_1,t_2)$受限于如下不等式:

$$n_{ij}(t_1,t_2)\leqslant(t_2-t_1)/\tau_{ij}^{D}+\eta_{ij}. \tag{13}$$

假设 2[38] (DoS 攻击持续时间) 对于给定的标量$\mu_{ij}\geqslant0,T_{ij}^{D}>1$, DoS 攻击在时间段$[t_1,t_2)$内持续的时间$|\Xi_{ij}(t_1,t_2)|$受限于如下不等式:

$$|\Xi_{ij}(t_1,t_2)|\leqslant(t_2-t_1)/T_{ij}^{D}+\mu_{ij}. \tag{14}$$

针对不同 DoS 攻击模型下多智能体系统的稳定性分析, 需要深入考虑攻击特征 (如周期性、随机性、能量约束等), 以选取合适的方法进行稳定性判据构建和控制方案设计. 在 PWM 刻画的 DoS 攻击场景下, 引入时间触发控制并构建简洁的 Lyapunov 函数, 即可实现对系统的有效镇定[34]; 对于随机突发的 DoS 攻击, 则需采用随机微分方程以及随机跳变系统理论, 通过概率方法分析系统的均方稳定[36]; 当攻击者能量受限导致频率和持续时间受约束时, 切换系统理论和基于平均驻留时间的多 Lyapunov 函数成为首选工具, 可以有效处理攻击导致通信间歇性失效问题, 并显式地给出攻击参数约束的稳定性条件[38]; 针对局部节点遭受 DoS 攻击导致拓扑结构动态变化的情况, 则可以借助图论中的拉普拉斯矩阵谱分析技术, 量化攻击对系统一致性的影响[72]. 此外, 在多智能体系统一致性判据的构建过程中, 不仅要全面考虑各类攻击模式的特点, 还需深入分析系统本身的动态特性, 如系统的非线性程度、维度、时延等, 进而选取合适的矩阵不等式处理技巧来构建稳定性条件, 这些综合分析方法共同决定了所得稳

表2 不同 DoS 攻击模型下的建模分析方法及对系统性能影响

DoS攻击模型	常用建模方法	稳定性分析工具	系统性能
PWM型攻击[34]	时间触发控制	时变Lyapunov函数	稳定性和保守性受脉冲宽度直接影响
随机性攻击[36-37, 70]	随机微分方程	均方稳定性	稳定性依赖攻击发生的统计特征
能量约束型攻击[38]	切换系统理论	平均驻留时间和多Lyapunov函数	涉及攻击频率及持续时间与控制效果的折中
局部节点攻击[72]	图论拓扑分析	拉普拉斯矩阵特征值分析	稳定性受连通性阈值限制

定性判据的保守性和适用性. 基于以上讨论, 进一步提供了表 2, 以更清楚地展示不同攻击模型下的建模分析方法以及对系统性能的影响.

4　常用的防御机制与攻击检测方法

4.1　常用防御机制

为了应对网络中的窃听攻击, 研究人员提出了许多隐私保护方法, 旨在保护智能体之间交互过程中的敏感信息不被泄露或滥用. 下面简单介绍几种常见的隐私保护方法.

1) 差分隐私方法.

差分隐私作为一种隐私保护技术, 通过在数据集中添加噪声来确保数据查询结果的隐私性. 具体而言, 如果两个数据集仅在一条记录上存在差异, 则对这两个数据集执行相同的查询操作后, 查询结果应当是不可区分的.

定义 3[46]　对于任意给定的$\delta \geqslant 0$, 如果存在某个$k \in \{1, 2, \ldots, N\}$使得

$$|x_i - x_i'| = \begin{cases} \delta, & i = k; \\ 0, & i \neq k. \end{cases}$$

则向量x, x'被称作是δ-邻接的.

定义 4[46]　对于任意给定的一对δ-邻接的初始状态$\theta(0), \theta'(0)$, 如果对观测序列和噪声序列的任意集合O和Ω, 有下式成立:

$$P\{X_{\theta(0)}(\eta) \in O | \eta \in \Omega\} \leqslant e^{\epsilon} P\{X_{\theta'(0)}(\eta') \in O | \eta' \in \Omega\}, \tag{15}$$

则称随机系统是能保持ϵ-差分隐私的.

定义 4 中的参数ϵ被称为隐私损失或隐私预算, 用于控制添加到原始数据集的噪声. ϵ值越小, 隐私保护级别越高, 但数据准确性可能降低; 反之, ϵ值越大, 数据准确性越高, 但隐私保护级别越低.

差分隐私方法提供可量化的隐私保护水平, 通过调整隐私预算ϵ的值可以控制隐私保护的程度. 但由于噪声的引入会影响最终收敛结果的正确性, 未来可以通过研究差分隐私的改进算法, 设计更合理的噪声扰动策略等方式, 提高隐私保护的效果.

2) 同态加密方法.

同态加密是一种密码学技术, 属于安全多方计算范畴, 它允许用户对密文直接执行特定形式的代数运算, 得到的结果以加密形式保存, 将其解密所得到的结果与对未加密数据 (即明文) 执行同样的操作时所产生的输出一样. 在公钥加密算法的基础上, 同态加密算法一般由以下 3 个函数组成:

① 密钥生成函数, 以安全参数λ作为输入, 输出加密公钥和私钥k_p, k_s.

② 加密函数$E(k_p, m)$, 以公钥k_p和明文m作为输入, 输出密文c, 即$c = E(k_p, m)$.

③ 解密函数$D(k_s, c)$, 以私钥k_s和密文c作为输入, 输出明文m, 即$m = D(k_s, c)$.

同态加密是指两个明文m_1、m_2满足

$$D(k_s, E(k_p, m_1) \odot E(k_p, m_2)) = m_1 \oplus m_2,$$

其中$\odot$、$\oplus$分别对应密文和明文域上的运算. 当$\oplus$代表加法时, 称该加密为加同态加密算法; 当$\oplus$代表乘法时, 称该加密为乘同态加密算法. 同时满足加同态和乘同态性质, 并且可以进行任意多次加和乘运算的同态加密算法称为全同态加密算法.

同态加密方法可以提供高维的安全性, 很好地应对内部和外部的窃听攻击. 然而, 由于通信和计算压力过大, 使其无法适用于资源有限或具有快速演化行为的系统[73].

3) 状态分解方法.

文献 [47] 提出的状态分解方法为隐私保护领域提供了另一种创新性解决方案. 与传统加密和加噪机制不同, 该方法创新性地从网络拓扑结构维度切入. 一方面, 通过对节点初始状态分解施加限制, 确保系统最终能收敛至正确的全局一致值; 另一方面, 构建了同时抵御内部好奇节点推理和外部窃听者攻击的双层防护体系. 此外, 该方法操作简单, 计算量小, 可有效应用于资源有限系统, 具体方案介绍如下.

每个节点将其状态x_i分解为两个子状态x_i^{α}和x_i^{β}, 其中$x_i^{\alpha}(0)$和$x_i^{\beta}(0)$可从实数集中任取, 但需要满足如下约束条件:

$$x_i^{\alpha}(0) + x_i^{\beta}(0) = x_i(0). \tag{16}$$

在新的网络拓扑结构下, 子状态x_i^{α}接替了原始状态x_i在节点间交互的角色, 是节点i的邻居唯一能接收到的状态值. 另一个子状态x_i^{β}只与同一节点的子状态x_i^{α}交互, 对其他节点完全不可见. 以图 2 (b) 中的节点 1 为例, x_1^{α}在节点间交互中表现得同x_1一样, 而x_1^{β}对除节点 1 以外的其他节点是不可见的, 尽管它影响了x_1^{α}的演化. 两个子状态x_i^{α}和x_i^{β}之间的耦合权重是对称的, 记作$a_{i,\alpha\beta}(k)$.

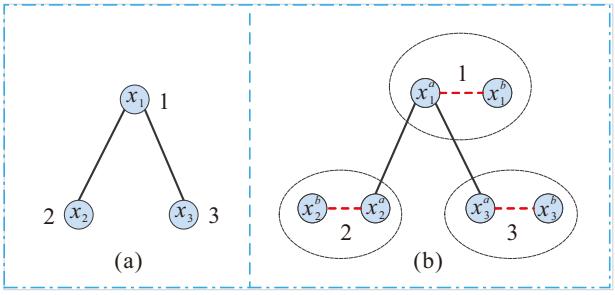

图2　状态分解示例

表3 不同防御机制对比

方法	适用场景	计算复杂度	防御效果
差分隐私方法	适用于数据发布、数据分析等场景, 当需要在保护数据隐私的同时提供一定的数据分析功能时较为适用, 例如政府统计数据发布、企业用户行为分析等	相对较低, 通常只需要在数据处理过程中添加一些噪声来实现隐私保护, 计算开销较小, 能在较短时间内处理大量数据	能有效防止基于数据差异的隐私攻击, 通过限制数据的变化对查询结果的影响, 保护个体数据的隐私性, 但可能会对数据的准确性有一定影响, 导致结果存在一定误差
同态加密方法	适用于云计算、数据外包等场景, 数据所有者希望在加密数据上进行计算, 而无需解密, 例如医疗数据在云端的处理、金融数据的外包计算等	较高, 因为涉及到复杂的加密和解密操作, 以及在加密数据上进行计算, 计算量较大, 处理速度相对较慢	能提供很强的隐私保护, 在密文上进行计算, 攻击者无法获取明文信息, 即使数据在传输或存储过程中被窃取, 也能保证数据的保密性, 但对计算资源要求较高, 可能会影响系统的性能
状态分解方法	适用于一些需要对系统状态进行分析和保护的场景, 如电力系统状态估计、网络系统状态监测等, 通过将系统状态分解为多个子状态, 分别进行处理和保护	取决于具体的系统规模, 一般来说计算复杂度适中. 如果系统规模较小, 则计算复杂度相对较低; 但对于复杂的系统, 计算复杂度可能会增加	能在一定程度上保护系统状态的隐私性, 通过分解状态使得攻击者难以直接获取完整的系统状态信息, 但对于一些针对特定子状态的攻击可能防御能力有限, 需要结合其他安全措施

在状态分解方法下, 节点的更新动态变为

$$\begin{cases} x_i^{\alpha}(k+1)= \\ x_i^{\alpha}(k)+\epsilon\displaystyle\sum_{j\in\mathcal{N}_i}a_{ij}(k)(x_j^{\alpha}(k)- \\ x_i^{\alpha}(k))+\epsilon a_{i,\alpha\beta}(k)(x_i^{\beta}(k)-x_i^{\alpha}(k)), \\ x_i^{\beta}(k+1)=x_i^{\beta}(k)+\epsilon a_{i,\alpha\beta}(k)(x_i^{\alpha}(k)-x_i^{\beta}(k)). \end{cases} \tag{17}$$

基于该结果, 一些扩展性的状态分解方法相应地被提出[74-75]. 此外, 状态分解方法还被广泛地应用于其他场景, 如弹性一致问题[76]、分布式优化问题[77]和智能电网中的经济调度问题[78] 等.

表 3 简单总结了 3 种不同防御机制在适用场景、计算复杂度和防御效果的比较结果.

4.2 攻击检测方法

在多智能体系统中, 攻击检测方案的设计是保障系统安全性和鲁棒性的重要环节, 有效的检测方案能够在攻击发生时快速、准确地识别异常并进行溯源, 为进一步设计合适的防御机制提供基础, 进而有助于减少攻击对系统性能的影响. 常用的攻击检测方案通常包括水印技术、卡方检测和 K-L 散度检测等. 下面对这 3 类典型的主动和被动检测方法进行说明, 揭示攻击检测在系统中的作用原理及重要意义.

1) 水印技术.

水印技术是一种主动防御技术, 通过向系统的控制信号或通信数据中注入特殊的、可验证的标识(即水印), 实现对数据的标记. 随后在接收端, 分析接收到的数据中是否包含水印信息, 根据水印信号是否缺失或篡改来判断传输数据是否受到可用性和完整性攻击. 这种主动检测可以有效应对隐蔽攻击, 且具有较高的鲁棒性, 难以被伪造, 但由于向系统中注入了水印信号, 可能会对系统性能造成一定影响, 需要兼顾水印强度与系统正常运行之间的平衡. 在存在虚假数据注入攻击的情况下, 为了提高卡方检测器对攻击的检测率, 文献 [79] 对传输的新息值进行了水印技术处理. 将待传输的新息值记为$z(t)$, 则类似文献 [79], 经过水印技术处理后的数据为

$$z_w(t)=mz(t)+w(t). \tag{18}$$

其中: m为常数因子, $w(t)$为独立同分布的零均值高斯随机变量. 考虑$z_w(t)$在传输过程中会受到如式 (7) 所示的比例-噪声形式的虚假数据注入攻击影响, 远程端获得的数据变为

$$\bar{z}_w(t)=T(t)z_w(t)+b(t). \tag{19}$$

基于水印信号, 对接收到的数据进行恢复处理, 可以得到

$$\tilde{z}_w(t)=T(t)z(t)+[(T(t)-I)w(t)+b(t)]/m. \tag{20}$$

由式 (18) 可以看出常数因子m和随机变量$w(t)$有效标记了攻击者注入的噪声信号$b(t)$, 通过对m和$w(t)$的合适设计, 可以提高虚假注入攻击被检测出的概率, 但同时在攻击存在的情况下, 水印信号也改变了远程端接收到的数据, 因此可能对系统的性能产生不利影响. 根据以上水印信号的设计原理, 多智能体系统基于水印技术进行欺骗攻击检测的问题被深入研究[43,80-81].

2) 卡方检测技术.

作为被动检测方案的代表, 卡方检测通过收集充足的样本, 基于统计分布设计检测条件, 用于判断所接收数据的分布是否偏离正常分布. 其形式简单易用, 对显著性偏差敏感, 但高度依赖对高斯分布的应用, 不适用于连续分布或复杂数据模式. 当待传输的残差信号$z(t)$是零均值、方差为Σ的独立同分布的高斯随机变量时, 考虑在远程端安装如下基于残差$z(t)$的卡方检测器:

$$\begin{cases} H_0:\ \lambda(t)=\displaystyle\sum_{k=t-h+1}^{t} z^{\mathrm{T}}(k)\Sigma^{-1}z(k)<\gamma, \\ H_1:\ \lambda(t)=\displaystyle\sum_{k=t-h+1}^{t} z^{\mathrm{T}}(k)\Sigma^{-1}z(k)\geqslant\gamma. \end{cases} \tag{21}$$

其中: 原假设H_0表示所接收的数据没有触发警报, 值得信任; H_1表示数据不可信; h为选取数据的窗口大小, γ为检测阈值. 可以看出, $\lambda(t)$是归一化残差平方和, 由于残差$z(t)$是零均值独立同分布的高斯分布, 若其维度为m, 则$\lambda(t)$是自由度为mh的卡方分布. 当网络攻击存在时, $z(t)$经过信道传输到达远程端检测器会变为$z^a(t)$, 若攻击对残差方差的影响比较大, 则远程端的检测信号$\lambda(t)$很容易超过检测阈值γ, 进而触发警报. 在研究网络攻击下多智能体系统安全一致性问题时, 考虑到智能体之间信息交互的复杂性, 基于高斯分布的残差设计的卡方检测方案使用并不广泛[43].

3) K-L 散度检测技术.

K-L 散度检测是一种基于信息论的分布差异测量方法, 用于判断当前数据分布与参考分布之间的偏差, 相比于卡方分布其灵活性更高, 可用于连续分布和复杂场景, 但对概率分布估计的要求较高. K-L 散度的定义如下.

定义 5[82-83]　随机变量Θ的观测分布P_Y与参考分布P_Z之间的 K-L 散度为

$$D_{\mathrm{KL}}(Y||Z)=\int_{\theta\in\Theta}P_Y(\theta)\log\left(\frac{P_Y(\theta)}{P_Z(\theta)}\right)\mathrm{d}\theta. \tag{22}$$

基于以上定义, 文献 [84] 通过引入以下两组误差序列建立基于 K-L 散度的攻击检测方案:

$$u_i=\left\|\sum_{j\in\mathcal{N}_i}a_{ij}(x_j^a-x_i^a+d_{ij})\right\|, \tag{23}$$

$$v_i=\sum_{j\in\mathcal{N}_i}\left\|a_{ij}(x_j^a-x_i^a+d_{ij})\right\|. \tag{24}$$

其中: x_i^a为智能体i受攻击影响的状态信号, d_{ij}为零均值高斯分布噪声. 检测条件设计为

$$\begin{cases} H_0:\dfrac{1}{h}\displaystyle\int_k^{k+h-1}D_{\mathrm{KL}}(u\_\{i\}||v\_\{i\})\mathrm{d}\kappa<\gamma_i, \\ H_1:\dfrac{1}{h}\displaystyle\int_k^{k+h-1}D_{\mathrm{KL}}(u\_\{i\}||v\_\{i\})\mathrm{d}\kappa>\gamma_i. \end{cases} \tag{25}$$

其中: γ_i为检测阈值, h为检测窗口, 原假设H_0表示传输数据没有受攻击影响. 在网络攻击环境下, 多智能体系统的攻击防御与安全控制研究中, 基于 K-L 散度的检测方案设计问题也得到了深入探索[84-85].

在实际研究及应用中, 为了提高系统对攻击的检测精度和防御能力, 通常会综合使用多种检测防御方法, 实现多层次的攻击检测, 提高系统的安全性和鲁棒性.

5　隐私保护下的多智能体系统一致性问题

5.1　不同类型隐私保护方法

隐私保护一致性研究是当前计算机科学、网络安全、数据科学等领域的重要研究方向, 旨在通过设计合理的隐私保护方案, 确保智能体的隐私信息得到充分地保护, 同时保证系统的一致性目标和一致结果的准确性可以实现.

首先介绍两类传统的隐私保护方法. 第 1 类为加密方法[55-56], 利用密码学理论, 将智能体的真实状态进行加密, 从而达到隐私保护的目的. 安全多方计算是在无可信第三方的情况下, 多个参与方协同计算一个约定的函数, 并且保证每一方仅获取自己的计算结果, 无法通过计算过程中的交互数据推测出其他任意一方的输入数据. 因此, 一些传统的安全多方计算方法被使用, 如 Yao 的混淆电路[86]、Shamir 的秘密共享算法[87-88] 等. 但这些方法大多基于完整网络, 智能体需要与网络中的所有其他智能体进行交互. 此外, 支持密文直接计算, 解密结果可以达到与明文操作相同的同态加密方法也被用来实现隐私保护. 文献 [89] 描述了当系统中至少存在一个可信智能体时如何使用同态加密进行隐私保护. 在文献 [45] 中, 同态加密首次以完全分散的方式实现而无需第三方的协助. 文献 [90] 提出了一种基于具有半同态特性的 Paillier 算法来解决二阶多智能体的隐私一致问题. 加密方法带来的隐私保护效果好, 可提供高维的安全性, 有效防御外部和内部窃听者的窃听. 然而, 由于通信和计算压力过大, 基于加密理论的方法无法适用于资源有限或具有快速演化行为的系统[73].

除了加密方法, 第 2 类使用较多的方法为加噪方法, 通过注入精心设计的噪声来掩盖真实状态. 这其中, 差分隐私方法提供了对隐私保护级别的数学严格评估, 在对后处理和侧信息的弹性、对手模型的免疫力和较低的计算成本方面具有明显的优势. 因此, 在文献 [46] 中, 首次将差分隐私机制用于多智能体系统一致问题, 即在每个智能体的状态中加入指数衰减的拉普拉斯噪声. 这之后, 差分隐私方法被广泛应用到各种类型的一致问题中, 如正一致问题[91]、最大一致问题[92]、量化一致问题[93-94]、二部一致问题[95] 和弹性一致问题[96] 等. 但差分隐私机制会影响最终收敛结果的准确性[97], 对于一些对结果精度要求比较高的应用不太适用. 之后, 学者们相继提

出了许多改进方法来克服隐私保护与结果准确性之间的平衡问题. 例如: 文献 [98] 设计了一种新的隐私保护算法, 通过在一致过程添加零均值高斯随机噪声来确保准确的收敛结果; 文献 [99-100] 中, 节点可以在迭代过程中添加任意随机噪声, 但需要保证添加的总噪声在某一步中最终消除.

除了经典的隐私保护方法外, 基于不同的设计思路也涌现出一系列新型方法. 文献 [47] 提出了一种巧妙的点分解方法, 令每个节点将其状态分解为两个子状态来避免隐私泄露. 基于该结果, 一些扩展性的点分解方法也相应地被提出[74-76]. 文献 [101-102] 提出了一种保证初始状态和时不变的策略, 在一致协议更新前, 通过与邻居进行随机值的传输, 构造出一个新的虚假初始状态序列, 只需要保证新的初始状态的和与原始初始状态的和相等即可. 文献 [103-104] 通过构造满足一定条件的函数来隐藏智能体的初始状态, 即智能体之间进行传输的不再是自己真实的状态值, 而是作用于这个函数之后的值. 文献 [105] 提出了一种基于边的扰动信号方法, 通过隐私保护的协调扰动阶段和收敛阶段实现隐私保护的目标. 更多地, 对状态的隐私保护也不再仅仅局限于初始状态, 实时状态保护也得到了关注[106]; 另一方面, 联邦学习技术, 作为“数据可用不可见”的应用新范式, 也被考虑到一致性问题的研究中[107].

当前, 如何克服加密方法所带来的计算和通信压力过大以及差分隐私方法导致的收敛结果不准确问题, 并设计更加全面有效的隐私保护策略, 逐渐成为研究者关注的主要方向之一.

5.2 不同类型多智能体模型

除了从不同的隐私保护方法视角入手, 研究者们还从不同类型的多智能体一致性模型考虑, 进行隐私保护研究.

1) 有限时间隐私保护一致性研究.

相较于渐近一致性, 有限时间一致性具有更强的抗干扰能力和更快的收敛速度[108], 是推动多智能体系统优化与高效运行的重要方向. 考虑有限时间一致下的隐私保护成为研究的热点之一. 基于同态加密机制, 文献 [48] 研究了时变变换网络下的有限时间平均一致性问题, 通过构造合适的输出掩码函数, 综合考虑连续时间多智能体系统的隐私保护和收敛性能. 文献 [49-50] 提出了新的有限时间隐私保护一致控制协议, 使系统在有限时间内达到一致.

2) 量化通信下隐私保护一致性研究.

在数字交互网络中, 由于通信带宽和计算能力有限, 即智能体之间单位时间内传输的数据量是有限的, 智能体交换的信息往往需要在传输前进行量化, 量化一致性问题成为一个值得研究的热门课题. 当前, 同时考虑隐私保护和量化通信的文献较少. 在量化通信环境下, 差分隐私方法被用于保护初始状态[93-94], 但结果表明, 收敛到的一致值是初始状态平均值邻域内的随机变量, 并不能保证精确收敛. 文献 [109] 设计了一种动态量化方案以确保系统在存在对手的情况下可以达成一致, 但其只考虑了外部窃听者, 内部好奇节点会直接获得节点的真实初始状态. 文献 [110] 直接假设初始状态为整数值, 因此最终收敛结果也是量化后的平均值. 此外, 文献 [51] 使用了加密方法解决隐私保护量化一致问题, 但它通常导致更多的计算和通信压力. 基于以上讨论可以发现, 隐私保护和量化通信的直接扩展是不平凡的, 主要涉及的挑战和困难是: 量化和隐私保护机制的同时存在使得最终收敛结果的正确性更难实现; 节点的初始状态可以是任何实数; 需要保证更全面的隐私保护 (内部好奇节点和外部窃听者). 此外, 所设计的机制应该执行简单, 计算和通信成本低.

3) 带有竞争交互的隐私保护二部一致性研究.

无论是在人类社会、自然界、微生物界, 还是工程应用中, 存在许多多智能体系统, 其群体内部既存在合作又存在竞争关系. 合作竞争网络是对传统合作网络的进一步深入, 在个体间合作关系的基础上, 也关注到了个体间竞争因素. 与合作网络中的一致性问题类似, 合作竞争网络中建立了二部一致性问题[111], 即要求所有智能体收敛到大小相同但符号相反的最终状态. 二部一致可以很好地解释许多现象, 例如, 两极分化经常发生在持相反观点的两个联盟团体中[112]. 其他如两党政治体系、双寡头市场、商业卡特尔都可以体现二部一致. 研究二部一致问题的隐私保护同样具有重要意义. 文献 [52, 95] 将差分隐私方法应用到了二部一致性问题中, 以确保最终的二部一致, 同时实现隐私保护. 文献 [113] 将合作者邻居与竞争对手邻居区分开, 通过只与合作者邻居进行随机数的交换, 构造出新的虚假初始状态. 这既保留了符号网络的特性, 也更符合实际场景. 文献 [53] 研究了一类离散时间非线性多智能体系统的二部一致隐私保护问题, 通过 Paillier 加密方案进行加密从而保护传输过程中的数据.

除上述列举的不同的多智能体模型外, 还有许多其他一致性问题受到学者的广泛关注, 如最大一致性[114]、弹性一致性[115]、动态平均一致[116] 等.

6　DoS 攻击下的多智能体系统一致性问题

6.1　事件触发控制方案

对于网络攻击下的多智能体系统, 建立安全一致性方案的另一个关键方向为: 设计鲁棒控制器, 提高系统的抗攻击能力. 此外, 讨论多智能体系统安全一致性时, 通信约束问题不容忽视. 由于通信资源的限制, 智能体之间的数据传输容易发生网络诱导现象, 如时滞、丢包、量化、信道衰落等[117-120], 这些现象的发生可能会导致数据传输不准确或失败, 若不采取合适的应对措施, 可能会使系统的性能变差, 甚至破坏系统的稳定性, 导致系统无法正常运行, 因此建立合理有效的方案来调度智能体之间的信息传输近年来也引起了众多学者的广泛关注. 目前, 很多研究工作都致力于设计有效的数据传输机制, 保证多智能体系统发生 DoS 攻击时, 基于数据传输的控制方案仍能有效作用于系统, 实现一致性控制目标. 从降低数据传输量或频率, 减少不必要数据通信的角度, 周期或非周期采样[121]、轮询协议[122]、Try-once-discard (TOD) 机制[123] 以及事件触发机制[35,124-125] 被相继引入到多智能体系统一致性问题的研究中. 具体地, 文献 [121] 讨论了随机 DoS 攻击下异构线性多智能体系统的非周期采样控制问题; 文献 [122] 利用轮询协议决定每一时刻哪个节点可以传输数据给观测器, 观测器利用获得的数据对原系统的状态信息进行估计, 再利用获得的估计状态构建控制器, 以实现对原系统的分布式准一致性控制; 文献 [123] 利用滑模控制方法分析了多智能体系统在 TOD 传输机制下的一致性问题; 文献 [35,124-125] 详细讨论了 DoS 攻击下多智能体系统的事件触发控制方案设计问题.

在轮询协议或 TOD 传输机制的作用下, 由于每一时刻只有一个节点可以获得数据传输权限, 大大降低了每一时刻的数据传输量, 实现了节约有限通信带宽的目的. 然而, 轮询协议实际上是一种静态数据传输策略, 在各个采样时刻以循环的方式将传输权限平等地分配给各个节点, 虽然原理简单, 但需要在每一个采样时刻都进行数据传输, 没有基于节点数据特征的灵活调度. 相比之下, TOD 机制在每个采样时刻动态地将传输权限分配给与上一次传输具有最大数据差值的节点, 然而这种动态方法的灵活性是对于节点的而非采样时刻, 这可能无法在 DoS 攻击存在时有效地减轻数据传输风险. 可以发现, 只在特定事件发生时才进行数据传输的事件触发机制, 相比于轮询协议或 TOD 机制具有更灵活的动态适应能力, 在增强对 DoS 攻击的抵抗能力方面具有显著优势. 因为事件触发机制是“按需”进行数据传输, 通常利用当前数据以及与上一次更新数据之间的误差建立触发条件, 只有当误差数据足够大时, 触发器才会执行数据传输动作. 考虑一种极端的情况: 若 DoS 攻击的时间区域恰好落在相邻两次事件触发时刻之间, 则 DoS 攻击相当于完全失效, 由此可见事件触发机制能更有效地应对 DoS 攻击存在的情况.

自文献 [126] 设计出基于事件的 PID 控制器, 体现了事件触发控制的优越性后, 事件触发镇定反馈控制问题便得到了控制等领域学者的广泛关注. DoS 攻击下多智能体系统的触发条件通常可以设计为如下形式:

$$t_{k+1}^i = \inf\{t > t_k^i | f(x_i(t), x_j(t), j \in \mathcal{N}_i) > 0\}, \tag{26}$$

其中 $f(x_i(t), x_j(t), j \in \mathcal{N}_i)$[127] 可以设计为

$$f(x_i(t), x_j(t), j \in \mathcal{N}_i) = \|\mathrm{e}^{A(t-t_k^i)} x_i(t_k^i) - x_i(t)\| - \gamma_i.$$

或者设计为[128]

$$f(x_i(t), x_j(t), j \in \mathcal{N}_i) = \mathrm{e}_i^{\mathrm{T}}(t)\Theta e_i(t) - \sigma q_i^{\mathrm{T}}(t)\Theta q_i(t).$$

这里 $e_i(t) = q_i(t_k^i) - q_i(t)$, 且

$$q_i(t) = \sum_{j=1}^{N} a_{ij}(x_j(t) - x_i(t)) + d_i(x_0(t) - x_i(t)).$$

触发条件还可以设计为更复杂的形式[129-130], 值得一提的是, 常规事件触发机制需要触发器对当前状态连续检测, 并持续计算判断触发条件是否满足, 这显然会带来繁重的计算负担. 为了应对此问题, 自触发机制被提出用于避免持续监测状态和触发条件. 从生成传输序列的角度看, 事件触发机制是被动的, 而自触发是主动的, 因为不同于事件触发机制通过实时监测触发条件来决定是否传输数据, 自触发机制根据当前传输时刻收集到的数据, 提前预测计算出下一个传输时刻. 文献 [131] 详细说明了 DoS 攻击下自触发机制的设计原理, 如下所示:

$$t_{k+1}^i = t_k^i + \Gamma(x_i(t_k^i)),\ \Gamma(\cdot): \mathbb{R}^{n_x} \to \mathbb{R}^+. \tag{27}$$

基于对系统模型和 Lyapunov 函数的分析利用, $\Gamma(x_i(t_k^i))$在 DoS 攻击和非攻击区域通常会有不同的设计形式. 利用类似的分析方法, 文献[132-133] 研究了多智能体系统自触发安全一致性问题.

另一方面, 在事件触发机制的设计过程中, 一个关键的问题是如何构造合适的触发律, 保证触发序列不会出现芝诺现象. 芝诺现象是指在有限的时间

内做出无限次数据传输更新. 因此, 对于多智能体系统, 为了避免事件触发机制在 DoS 攻击区域发生芝诺现象, 文献 [125] 在攻击区域引入一组正常数 $\vartheta_i, i \in \{1, 2, \ldots, N\}$, 以保证连续两个触发时刻之间的间隔. 为了进一步拓宽相邻两次触发之间的时间间隔, 文献 [134] 提出了动态事件触发原理, 具体如下:

$$t_{k+1} = \inf\{t > t_k | \eta(t) + \theta(\sigma\alpha(||x(t)||) - \gamma(||e(t^-)||)) \leqslant 0\}. \tag{28}$$

在此研究成果基础上, 文献 [135] 针对 DoS 攻击下的多智能体系统设计了动态事件触发律. 进一步地, 文献 [136-137] 综合讨论了 DoS 攻击下分布式事件触发和自触发机制的设计原理.

6.2 不同多智能体模型下的安全一致性研究

1) DoS 攻击下固定时间安全一致性研究.

在多智能体系统中, 固定时间一致性是一个重要的控制目标, 旨在对于任意系统初值, 所有智能体状态能在有限且可预先设定的时间内达到一致. 当发生 DoS 攻击时, 数据传输会被中断, 进而导致控制方案中的信息交互不连续. 尽管近年来关于多智能体系统固定时间镇定问题的研究取得了很大进展, 但大多数现有结果都是通过分析具有随时间呈单调递减趋势的 Lyapunov 函数建立的. 然而, 当考虑 DoS 攻击导致通信故障时, 构造一个随时间单调递减 (不一定严格单调递减) 的 Lyapunov 函数很难实现, 因此相关研究具有极大的挑战性. 考虑 DoS 攻击会破坏通信拓扑, 文献 [138] 讨论了拓扑连通性保持和被破坏两种情况下, 非线性多智能体系统安全跟踪一致性问题, 设计了分布式固定时间观测器和固定时间控制器. 对于有向图描述的高阶领导-跟随多智能体系统, 文献 [139] 提出了一种检测 DoS 攻击的算法, 并建立了固定时间事件触发分布式观测器来估计领导者的状态, 进而设计了固定时间领导-跟随一致性控制策略. 文献 [140] 通过建立具有低保守性的固定时间稳定条件, 设计了分布式事件触发控制方案, 实现了 DoS 攻击下非线性多智能体系统的固定时间平均一致性.

2) 带有竞争交互的安全一致性研究.

在带有竞争交互的多智能体系统中, 二部一致性是一种特殊的协同控制目标, 要求智能体在两组对立的子集 (即二部图结构) 内分别实现状态一致, 同时保证两组之间状态符号相反 (例如, 一组趋于正值, 另一组趋于负值). 前述讨论已说明二部一致性具有重要的实际应用价值, 因此关于 DoS 攻击下多智能体系统的二部一致性研究近年来也得到广泛关注, 相关研究工作主要致力于对多智能体系统和攻击模型进行建模、处理拓扑图中权重符号相反的问题、设计鲁棒二部一致性控制方案、分析建立稳定性条件、仿真验证理论结果. 据此, 针对 DoS 攻击下的多智能体系统, 文献 [141] 研究了事件触发二部一致性问题, 文献 [142] 讨论了动态事件触发二部一致性问题, 文献 [143] 考虑了事件触发自适应二部一致性问题, 文献 [144] 进一步讨论了数据驱动事件触发二部一致性问题.

7 实际应用

1) 多智能体隐私保护应用案例.

在智能交通领域, 车辆协同驾驶的数据隐私与车主个人信息、人身安全息息相关. 针对这一问题, 文献 [145] 构建了基于差分隐私的内部安全信息粒度模型, 以解决轨迹共享时的数据隐私泄露问题. 文献 [146] 引入联邦学习保护自动驾驶的车辆数据, 通过移动边缘计算服务器共享训练模型参数; 在智能电网领域, 分布式能源交易有利于家庭或企业通过共享用电数据优化调度, 但用电数据的泄露容易导致针对性的攻击和精准诈骗. 因此, 文献 [147] 在同态加密方案中设计了一种掩码方法, 并提出了直流和交流模型的隐私保护状态估计协议. 在医疗协作领域, 医院或研究机构协作训练疾病预测模型, 需对患者医疗记录进行保密. 文献 [148] 设计了一种新的隐私保护马氏距离比较方法来提高医学图像检索的准确性, 然后结合基于马氏距离的模糊 C-均值算法, 以实现对医学图像检索过程中的隐私保护. 在军事领域, 无人机执行协同搜救等共享位置的任务时, 需要避免军事设施等敏感区域被暴露. 为了解决位置隐私问题, 文献 [149] 基于 SM4 的轻量级对称加密算法, 提出了面向无人机网络的安全传输方法.

2) 多智能体安全控制应用案例.

在工业领域, 文献 [150] 指出, 工业控制系统对关键基础设施的稳定运行至关重要, 但其日益增长的复杂性和互联性要求综合的安全与防护措施. 因此, 文中提出了自动标记语言建模与贝叶斯信念网络风险评估方法, 解决了因数据分散导致的安全与防护综合风险评估难题. 在军事领域, 无人集群在对抗环境中的协同围捕时, 需要同时抵御干扰或欺骗攻击, 在执行搜救、军事侦察时, 需要避免碰撞和外部干扰并保持队形覆盖目标区域. 为防御攻击、抵抗干扰, 文献 [151] 研究了抗干扰的无人机路径规划策略, 通过学习最优轨迹来躲避干扰攻击. 在物流仓储

中, 自动导引车在仓库协同搬运货物, 需要协同优化路径并避免死锁. 为了实现这一目标, 文献 [152] 设计了无死锁的仓储系统调度问题, 文献 [153] 利用动态随机网络分析路径网络, 在 A*启发式算法中引入概率时间约束并设计了适合启发式规划算法的多导引车冲突避免策略.

这些实际应用案例表明, 在工业、军事、医疗、交通、电网等场景中, 都离不开多智能体一致性安全与隐私保护的相关研究. 随着人工智能、通信技术的不断发展, 隐私保护机制和安全控制策略将得到进一步的优化和发展.

8　总结与展望

随着信息技术的快速发展和智能化应用的普及, 多智能体系统在协同控制、优化决策等方面展现出了巨大的潜力. 然而, 在实现一致性的过程中, 其面对各种网络攻击的安全性问题日益凸显, 成为制约其进一步应用的关键因素. 本文对网络攻击下的多智能体系统一致性安全与隐私保护的研究进展进行总结, 从窃听攻击、虚假数据注入攻击和 DoS 攻击这 3 类典型攻击模式入手, 从防御者角度给出了常见的隐私保护方法、攻击检测方法以及安全控制方法. 进一步地, 详细介绍了当前基于隐私保护的多智能体系统一致性研究成果和 DoS 攻击下的多智能体系统一致性问题研究成果.

尽管隐私保护一致性研究已取得显著进展, 但仍存在若干关键性技术挑战亟待突破:

1) 跨域数据协作中, 异构隐私策略的规则冲突消解机制尚未建立;

2) 量子计算时代来临对现有加密体系造成的长期安全威胁;

3) 数据效用保持与隐私保护间动态平衡难题;

4) 缺乏系统化的评估框架来量化隐私一致性级别.

针对这些挑战, 未来可以从以下方向开展深入研究:

1) 设计跨隐私计算框架的一致性协议. 构建支持多方计算的隐私协议栈, 重点突破异构算法间的协同问题, 混合部署场景下的动态一致性保障机制以及联邦学习中的跨域策略冲突实时检测技术.

2) 研究抗量子攻击隐私保护方案, 开展抗量子攻击的密码学体系研究.

3) 建立可验证的隐私一致性模型. 开发隐私损失与一致性级别的量化关联模型, 设计面向医疗、金融等高敏感场景的端到端验证协议.

4) 构建多维评估体系. 从基础层、协议层和隐私层等角度出发, 建立性能指标.

针对多智能体系统在网络攻击下的安全控制问题, 当前研究的难点和技术瓶颈主要集中在:

1) 攻击模型的建立、攻击检测以及防御机制的设计;

2) 动态威胁环境下的实时性;

3) 系统规模扩展性.

为了克服上述挑战, 未来的研究重点如下:

1) 精细化攻击建模, 设计高精度、轻量化检测机制, 设计主动弹性防御机制;

2) 通过边缘计算、云计算、机器学习等技术, 设计低时延检测与响应机制, 保证防御策略的动态环境自适应性, 进行实时资源调度;

3) 研究分布式可扩展架构, 通过图采样等技术提升计算与通信效率, 设计跨协议、跨平台的安全控制算法提升异构系统的兼容性.

参考文献 (References)

[1] Sharf M, Zelazo D. Analysis and synthesis of MIMO multi-agent systems using network optimization[J]. IEEE Transactions on Automatic Control, 2019, 64(11): 4512-4524.

[2] Xiao S Y, Dong J X. Distributed adaptive fuzzy fault-tolerant containment control for heterogeneous nonlinear multiagent systems[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2022, 52(2): 954-965.

[3] Karnouskos S, Leitao P, Ribeiro L, et al. Industrial agents as a key enabler for realizing industrial cyber-physical systems: Multiagent systems entering industry 4.0[J]. IEEE Industrial Electronics Magazine, 2020, 14(3): 18-32.

[4] Bodin O N, Ubiennykh A G, Bezborodova O E, et al. Improving the information reliability in medical information system based on multi-agent technology[C]. The 21st International Conference of Young Specialists on Micro/Nanotechnologies and Electron Devices. Chemal, 2020: 427-431.

[5] Haydari A, Yılmaz Y. Deep reinforcement learning for intelligent transportation systems: A survey[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(1): 11-32.

[6] Khoo S, Xie L H, Man Z H. Robust finite-time consensus tracking algorithm for multirobot systems[J]. IEEE/ASME Transactions on Mechatronics, 2009, 14(2): 219-228.

[7] Xiao F, Wang L, Chen J, et al. Finite-time formation control for multi-agent systems[J]. Automatica, 2009, 45(11): 2605-2611.

[8] Pavlin G, de Oude P, Maris M, et al. A multi-agent systems approach to distributed Bayesian information

fusion[J]. Information Fusion, 2010, 11(3): 267-282.

[9] Hanada K, Wada T, Masubuchi I, et al. Multi-agent consensus for distributed power dispatch with load balancing[J]. Asian Journal of Control, 2021, 23(2): 611-619.

[10] Qasim A, Ghouri A, Munawar A. An effective approach for reducing data redundancy in multi-agent system communication[J]. Multiagent and Grid Systems, 2024, 20(1): 69-88.

[11] Zhao N, Zhang H Y, Shi P. Observer-based sampled-data adaptive tracking control for heterogeneous nonlinear multi-agent systems under denial-of-service attacks[J]. IEEE Transactions on Automation Science and Engineering, 2024, 22: 4771-4779.

[12] Luzolo P H, Elrawashdeh Z, Tchappi I, et al. Combining multi-agent systems and artificial intelligence of things: Technical challenges and gains[J]. Internet of Things, 2024, 28: 101364.

[13] Feng Z, Hu G Q. Distributed secure average consensus for linear multi-agent systems under DoS attacks[C]. 2017 American Control Conference. Seattle, 2017: 2261-2266.

[14] Li Z C, Shi Y, Xu S Y, et al. Distributed model predictive consensus of MASs against false data injection attacks and denial-of-service attacks[J]. IEEE Transactions on Automatic Control, 2024, 69(8): 5538-5545.

[15] Yang Y, Li Y F, Yue D, et al. Distributed secure consensus control with event-triggering for multiagent systems under DoS attacks[J]. IEEE Transactions on Cybernetics, 2021, 51(6): 2916-2928.

[16] Lindsay J R. Stuxnet and the limits of cyber warfare[J]. Security Studies, 2013, 22(3): 365-404.

[17] McMullen D A, Sanchez M H, Reilly-Allen M O. Target security: A case study of how hackers hit the jackpot at the expense of customers[J]. Review of Business and Finance Studies, 2016, 7(2): 41-50.

[18] Case D U. Analysis of the cyber attack on the Ukrainian power grid[J]. Electricity Information Sharing and Analysis Center, 2016, 388(3): 1-29.

[19] Zhao B Y, Zhang Y. Secure encoding strategy for consensus of multi-agent systems in the presence of eavesdropper[J]. IEEE Transactions on Circuits and Systems II: Express Briefs, 2022, 69(8): 3420-3424.

[20] He W L, Gao X Y, Zhong W M, et al. Secure impulsive synchronization control of multi-agent systems under deception attacks[J]. Information Sciences, 2018, 459: 354-368.

[21] Zhang D, Ye Z H, Dong X W. Co-design of fault detection and consensus control protocol for multi-agent systems under hidden DoS attack[J]. IEEE Transactions on Circuits and Systems I: Regular Papers, 2021, 68(5): 2158-2170.

[22] Kapetanovic D, Zheng G, Rusek F. Physical layer security for massive MIMO: An overview on passive eavesdropping and active attacks[J]. IEEE Communications Magazine, 2015, 53(6): 21-27.

[23] Tahoun A H, Arafa M. Cooperative control for cyber–physical multi-agent networked control systems with unknown false data-injection and replay cyber-attacks[J]. ISA Transactions, 2021, 110: 1-14.

[24] Wang C Y, Huang J H, Wang D, et al. A secure strategy for a cyber physical system with multi-sensor under linear deception attack[J]. Journal of the Franklin Institute, 2021, 358(13): 6666-6683.

[25] Zhao R, Zuo Z Q, Shi Y, et al. DoS and stealthy deception attacks for switched systems: A cooperative approach[J]. IEEE Transactions on Automatic Control, 2024, 69(7): 4396-4410.

[26] Luo X Y, Zhao C C, Fang C R, et al. Submodularity-based false data injection attack scheme in multi-agent dynamical systems[J]. Automatica, 2024, 160: 111426.

[27] Jahangiri-Heidari M, Shahriari-Kahkeshi M, Shi P. Resilient consensus of nonlinear multiagent systems under false data injection attack on communication channels: An attack detection and isolation-based approach[J]. IEEE Internet of Things Journal, 2025, 12(7): 8219-8230.

[28] Sun L C, Wu T J, Zhang Y. A defense strategy for false data injection attacks in multi-agent systems[J]. International Journal of Systems Science, 2023, 54(16): 3071-3084.

[29] Li X M, Zhou Q, Li P S, et al. Event-triggered consensus control for multi-agent systems against false data-injection attacks[J]. IEEE Transactions on Cybernetics, 2020, 50(5): 1856-1866.

[30] Nasiri S, Seifi H, Delkhosh H. A secure power system distributed state estimation via a consensus-based mechanism and a cooperative trust management strategy[J]. IEEE Transactions on Industrial Informatics, 2024, 20(2): 3002-3014.

[31] Wang P Y, Govindarasu M. Multi-agent based attack-resilient system integrity protection for smart grid[J]. IEEE Transactions on Smart Grid, 2020, 11(4): 3447-3456.

[32] Amullen E, Lin H, Kalbarczyk Z, et al. Multi-agent system for detecting false data injection attacks against the power grid[C]. Proceedings of the 2nd Annual Industrial Control System Security Workshop. New York: ACM, 2016: 38-44.

[33] Xie Z K, Wu Z Q. Distributed fault-tolerant secondary control for DC microgrids against false data injection attacks[J]. International Journal of Electrical Power & Energy Systems, 2023, 144: 108599.

[34] Shisheh Foroush H, Martínez S. On triggering control of single-input linear systems under pulse-width modulated DoS signals[J]. SIAM Journal on Control and Optimization, 2016, 54(6): 3084-3105.

[35] Guo X G, Liu P M, Wang J L, et al. Event-triggered adaptive fault-tolerant pinning control for cluster consensus of heterogeneous nonlinear multi-agent systems under aperiodic DoS attacks[J]. IEEE

Transactions on Network Science and Engineering, 2021, 8(2): 1941-1956.

[36] Chen R Z, Li Y X, Hou Z S. Distributed model-free adaptive control for multi-agent systems with external disturbances and DoS attacks[J]. Information Sciences, 2022, 613: 309-323.

[37] Befekadu G K, Gupta V, Antsaklis P J. Risk-sensitive control under Markov modulated denial-of-service (DoS) attack strategies[J]. IEEE Transactions on Automatic Control, 2015, 60(12): 3299-3304.

[38] Dolk V S, Tesi P, De Persis C, et al. Event-triggered control systems under denial-of-service attacks[J]. IEEE Transactions on Control of Network Systems, 2017, 4(1): 93-105.

[39] Yan C, Yan L, Lv Y, et al. Adaptive event-triggered resilient control of heterogeneous multiagent under DoS attacks and link faults on directed graphs[J]. IEEE Transactions on Network Science and Engineering, 2024, 11(1): 834-847.

[40] Cetinkaya A, Ishii H, Hayakawa T. An overview on denial-of-service attacks in control systems: Attack models and security analyses[J]. Entropy, 2019, 21(2): 210.

[41] Wang J Y, Deng X M, Guo J H, et al. Resilient consensus control for multi-agent systems: A comparative survey[J]. Sensors, 2023, 23(6): 2904.

[42] Ishii H, Wang Y, Feng S. An overview on multi-agent consensus under adversarial attacks[J]. Annual Reviews in Control, 2022, 53: 252-272.

[43] He W L, Xu W Y, Ge X H, et al. Secure control of multiagent systems against malicious attacks: A brief survey[J]. IEEE Transactions on Industrial Informatics, 2022, 18(6): 3595-3608.

[44] Aslam M M, Ahmed Z, Du L P, et al. An overview of recent advances of resilient consensus for multiagent systems under attacks[J]. Computational Intelligence and Neuroscience, 2022, 2022: 6732343.

[45] Ruan M H, Gao H, Wang Y Q. Secure and privacy-preserving consensus[J]. IEEE Transactions on Automatic Control, 2019, 64(10): 4035-4049.

[46] Huang Z Q, Mitra S, Dullerud G. Differentially private iterative synchronous consensus[C]. Proceedings of the 2012 ACM Workshop on Privacy in the Electronic Society. New York: ACM, 2012: 81-90.

[47] Wang Y Q. Privacy-preserving average consensus *via* state decomposition[J]. IEEE Transactions on Automatic Control, 2019, 64(11): 4711-4716.

[48] Liu X W, Ma H L. Privacy preserving finite-time consensus in networks with time-varying topology[C]. 2019 34rd Youth Academic Annual Conference of Chinese Association of Automation. Jinzhou, 2019: 312-316.

[49] Zhang J, Lu J Q, Lou J G. Privacy-preserving average consensus *via* finite time-varying transformation[J]. IEEE Transactions on Network Science and Engineering, 2022, 9(3): 1756-1764.

[50] Yue J F, Li W H, Shi M J, et al. Finite-time privacy-preserving average consensus control of multi-agent systems via output mask approach[C]. 2023 35th Chinese Control and Decision Conference. Yichang, 2023: 2670-2675.

[51] Kishida M. Encrypted average consensus with quantized control law[C]. 2018 IEEE Conference on Decision and Control. Miami, 2018: 5850-5856.

[52] Gao C, Zhao D, Li J H, et al. Private bipartite consensus control for multi-agent systems: A hierarchical differential privacy scheme[J]. Information Fusion, 2024, 105: 102259.

[53] Yang Z W, Yu L Y, Liu Y R, et al. Event-triggered privacy-preserving bipartite consensus for multi-agent systems based on encryption[J]. Neurocomputing, 2022, 503: 162-172.

[54] Olfati-Saber R, Murray R M. Consensus problems in networks of agents with switching topology and time-delays[J]. IEEE Transactions on Automatic Control, 2004, 49(9): 1520-1533.

[55] Katz J, Lindell Y. Introduction to modern cryptography[M]. Chapman and Hall, 2014.

[56] Gao C, Wang Z D, He X, et al. Encryption-decryption-based consensus control for multi-agent systems: Handling actuator faults[J]. Automatica, 2021, 134: 109908.

[57] Navarro-Arribas G, Torra V. Information fusion in data privacy: A survey[J]. Information Fusion, 2012, 13(4): 235-244.

[58] López-Aguilar P, Batista E, Martínez-Ballesté A, et al. Information security and privacy in railway transportation: A systematic review[J]. Sensors, 2022, 22(20): 7698.

[59] Ye M B, Qin Y Z, Govaert A, et al. An influence network model to study discrepancies in expressed and private opinions[J]. Automatica, 2019, 107: 371-381.

[60] Lin P, Ren W, Wang H, et al. Multiagent rendezvous with shortest distance to convex regions with empty intersection: Algorithms and experiments[J]. IEEE Transactions on Cybernetics, 2019, 49(3): 1026-1034.

[61] Anwar M R, Apriani D, Adianita I R. Hash algorithm in verification of certificate data integrity and security[J]. Aptisi Transactions on Technopreneurship: ATT, 2021, 3(2): 65-72.

[62] Romney G W, Parry D W. A digital signature signing engine to protect the integrity of digital assets[C]. 2006 7th International Conference on Information Technology Based Higher Education and Training. Ultimo, 2006: 800-805.

[63] Ding D R, Han Q L, Ge X H, et al. Secure state estimation and control of cyber-physical systems: A survey[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2021, 51(1): 176-190.

[64] Liang G Q, Zhao J H, Luo F J, et al. A review of false data injection attacks against modern power systems[J]. IEEE Transactions on Smart Grid, 2017, 8(4): 1630-

1638.

[65] Peng C, Sun H T. Switching-like event-triggered control for networked control systems under malicious denial of service attacks[J]. IEEE Transactions on Automatic Control, 2020, 65(9): 3943-3949.

[66] Deng Y R, Yin X X, Hu S L. Event-triggered predictive control for networked control systems with DoS attacks[J]. Information Sciences, 2021, 542: 71-91.

[67] Murugesan S, Liu Y C. Resilient finite-time distributed event-triggered consensus of multi-agent systems with multiple cyber-attacks[J]. Communications in Nonlinear Science and Numerical Simulation, 2023, 116: 106876.

[68] Guo Z Y, Shi D W, Johansson K H, et al. Optimal linear cyber-attack on remote state estimation[J]. IEEE Transactions on Control of Network Systems, 2017, 4(1): 4-13.

[69] Yu S. Distributed denial of service attack and defense[M]. New York: Springer, 2014.

[70] Feng Z, Wen G H, Hu G Q. Distributed secure coordinated control for multiagent systems under strategic attacks[J]. IEEE Transactions on Cybernetics, 2017, 47(5): 1273-1284.

[71] Hespanha J P, Morse A S. Stability of switched systems with average dwell-time[C]. Proceedings of the 38th IEEE Conference on Decision and Control. Phoenix, 2002: 2655-2660.

[72] Zhang D Y, Li X J. Secure consensus control of multiagent systems under DoS attacks: A switching-scheme-based active defense method[J]. IEEE Transactions on Cybernetics, 2024, 54(12): 7404-7415.

[73] Naehrig M, Lauter K, Vaikuntanathan V. Can homomorphic encryption be practical?[C]. Proceedings of the 3rd ACM Workshop on Cloud Computing Security Workshop. New York: ACM, 2011: 113-124.

[74] Wang Y Q, Lu J Q, Zheng W X, et al. Privacy-preserving consensus for multi-agent systems via node decomposition strategy[J]. IEEE Transactions on Circuits and Systems I: Regular Papers, 2021, 68(8): 3474-3484.

[75] Zhang J, Lu J Q, Chen X Y. Privacy-preserving average consensus via edge decomposition[J]. IEEE Control Systems Letters, 2022, 6: 2503-2508.

[76] Zhang Y W, Peng Z X, Wen G G, et al. Privacy preserving-based resilient consensus for multiagent systems *via* state decomposition[J]. IEEE Transactions on Control of Network Systems, 2023, 10(3): 1172-1183.

[77] Zhang J H, Ma D. A novel state decomposition-based privacy-preserving algorithm for distributed optimization over directed networks[C]. 2024 14th Asian Control Conference. Dalian, 2024: 1145-1150.

[78] Chen W, Liu G P. Privacy-preserving consensus-based distributed economic dispatch of smart grids via state decomposition[J]. IEEE/CAA Journal of Automatica Sinica, 2024, 11(5): 1250-1261.

[79] Huang J H, Ho D W C, Li F F, et al. Secure remote state estimation against linear man-in-the-middle attacks using watermarking[J]. Automatica, 2020, 121: 109182.

[80] Khazraei A, Kebriaei H, Salmasi F R. Replay attack detection in a multi agent system using stability analysis and loss effective watermarking[C]. 2017 American Control Conference. Seattle, 2017: 4778-4783.

[81] Liu H, Li Y Z, Han Q L, et al. Watermark-based proactive defense strategy design for cyber-physical systems with unknown-but-bounded noises[J]. IEEE Transactions on Automatic Control, 2023, 68(6): 3300-3315.

[82] Basseville M, Nikiforov I V. Detection of abrupt changes: Theory and application[M]. Englewood Cliffs: Prentice Hall, 1993.

[83] Thomas M, Joy A T. Elements of information theory[M]. Wiley-Interscience, 2006.

[84] Mustafa A, Modares H, Moghadam R. Resilient synchronization of distributed multi-agent systems under attacks[J]. Automatica, 2020, 115: 108869.

[85] Zhang D, Feng G, Shi Y, et al. Physical safety and cyber security analysis of multi-agent systems: A survey of recent advances[J]. IEEE/CAA Journal of Automatica Sinica, 2021, 8(2): 319-333.

[86] Yao A C. Protocols for secure computations[C]. The 23rd Annual Symposium on Foundations of Computer Science. Chicago, 1982: 160-164.

[87] Shamir A. How to share a secret[J]. Communications of the ACM, 1979, 22(11): 612-613.

[88] Zhang S L, Ohlson Timoudas T, Dahleh M A. Consensus with preserved privacy against neighbor collusion[J]. Control Theory and Technology, 2020, 18(4): 409-418.

[89] Hadjicostis C N, Domínguez-García A D. Privacy-preserving distributed averaging via homomorphically encrypted ratio consensus[J]. IEEE Transactions on Automatic Control, 2020, 65(9): 3887-3894.

[90] Liu C C, Chen C C, Chen Y Y, et al. Consensus control of discrete-time multi-agent systems with privacy preservation[J]. Asian Journal of Control, 2023, 25(5): 3431-3442.

[91] Wang Y M, Lam J, Lin H. Differentially private average consensus for networks with positive agents[J]. IEEE Transactions on Cybernetics, 2024, 54(6): 3454-3467.

[92] Wang X, He J P, Cheng P, et al. Differentially private maximum consensus: Design, analysis and impossibility result[J]. IEEE Transactions on Network Science and Engineering, 2019, 6(4): 928-939.

[93] Zhang W J, Wang B C, Liang Y. Differentially private consensus for second-order multiagent systems with quantized communication[J]. IEEE Transactions on Neural Networks and Learning Systems, 2024, 35(4): 5523-5535.

[94] Gao L, Deng S J, Ren W, et al. Differentially private consensus with quantized communication[J]. IEEE Transactions on Cybernetics, 2021, 51(8): 4075-4088.
[95] Zuo Z Q, Tian R, Han Q N, et al. Differential privacy for bipartite consensus over signed digraph[J]. Neurocomputing, 2022, 468: 11-21.
[96] Fiore D, Russo G. Resilient consensus for multi-agent systems subject to differential privacy requirements[J]. Automatica, 2019, 106: 18-26.
[97] Nozari E, Tallapragada P, Cortés J. Differentially private average consensus: Obstructions, trade-offs, and optimal algorithm design[J]. Automatica, 2017, 81: 221-231.
[98] Mo Y L, Murray R M. Privacy preserving average consensus[J]. IEEE Transactions on Automatic Control, 2017, 62(2): 753-765.
[99] Manitara N E, Hadjicostis C N. Privacy-preserving asymptotic average consensus[C]. 2013 European Control Conference. Zurich, 2013: 760-765.
[100] Charalambous T, Manitara N E, Hadjicostis C N. Privacy-preserving average consensus over digraphs in the presence of time delays[C]. 2019 57th Annual Allerton Conference on Communication, Control, and Computing. Monticello, 2019: 238-245.
[101] Gupta N, Kat J, Chopra N. Statistical privacy in distributed average consensus on bounded real inputs[C]. 2019 American Control Conference. Philadelphia, 2019: 1836-1841.
[102] Yin T J, Lv Y Z, Yu W W. Accurate privacy preserving average consensus[J]. IEEE Transactions on Circuits and Systems II: Express Briefs, 2020, 67(4): 690-694.
[103] Altafini C. A dynamical approach to privacy preserving average consensus[C]. 2019 IEEE 58th Conference on Decision and Control. Nice, 2019: 4501-4506.
[104] Altafini C. A system-theoretic framework for privacy preservation in continuous-time multiagent dynamics[J]. Automatica, 2020, 122: 109253.
[105] Xiong Y, Li Z K. Privacy-preserved average consensus algorithms with edge-based additive perturbations[J]. Automatica, 2022, 140: 110223.
[106] Wang Z Q, Ma M L, Zhou Q, et al. A privacy-preserving distributed control strategy in islanded AC microgrids[J]. IEEE Transactions on Smart Grid, 2022, 13(5): 3369-3382.
[107] Yan X, Qin Y M, Hu X D, et al. Distributed consensus problem with caching on federated learning framework[J]. International Journal of Distributed Sensor Networks, 2022, 18(4): 155013292210929.
[108] Hui Q, Haddad W M, Bhat S P. Finite-time semistability and consensus for nonlinear dynamical networks[J]. IEEE Transactions on Automatic Control, 2008, 53(8): 1887-1900.
[109] Maity D, Tsiotras P. Multiagent consensus subject to communication and privacy constraints[J]. IEEE Transactions on Control of Network Systems, 2022, 9(2): 943-955.
[110] Rikos A I, Charalambous T, Johansson K H, et al. Distributed event-triggered algorithms for finite-time privacy-preserving quantized average consensus[J]. IEEE Transactions on Control of Network Systems, 2023, 10(1): 38-50.
[111] Altafini C. Consensus problems on networks with antagonistic interactions[J]. IEEE Transactions on Automatic Control, 2013, 58(4): 935-946.
[112] Flanigan W H, Riker W H. The theory of political coalitions[J]. Administrative Science Quarterly, 1965, 9(4): 454.
[113] Zhang J, Lu J Q, Chen X Y, et al. Privacy-preserving bipartite consensus on signed networks[J]. IEEE Transactions on Control of Network Systems, 2024, 11(2): 696-704.
[114] Venkategowda N K D, Werner S. Privacy-preserving distributed maximum consensus[J]. IEEE Signal Processing Letters, 2020, 27: 1839-1843.
[115] Tian R, Mei J, Ma G F. Privacy-preserving resilient bipartite consensus of multi-agent systems: A differential privacy scheme[J]. Nonlinear Analysis: Hybrid Systems, 2025, 56: 101579.
[116] Zhang K X, Li Z J, Wang Y Q, et al. Privacy-preserving dynamic average consensus *via* state decomposition: Case study on multi-robot formation control[J]. Automatica, 2022, 139: 110182.
[117] Olfati-Saber R, Fax J A, Murray R M. Consensus and cooperation in networked multi-agent systems[J]. Proceedings of the IEEE, 2007, 95(1): 215-233.
[118] Bao G Y, Ma L F, Yi X J. Recent advances on cooperative control of heterogeneous multi-agent systems subject to constraints: A survey[J]. Systems Science & Control Engineering, 2022, 10(1): 539-551.
[119] Ju Y M, Ding D R, He X, et al. Consensus control of multi-agent systems using fault-estimation-in-the-loop: Dynamic event-triggered case[J]. IEEE/CAA Journal of Automatica Sinica, 2022, 9(8): 1440-1451.
[120] Xing M P, Lu J Q, Lou J G, et al. Event-based fixed-time synchronization of neural networks under DoS attack and its applications[J]. Neural Networks, 2023, 166: 622-633.
[121] Zhang D, Liu L, Feng G. Consensus of heterogeneous linear multiagent systems subject to aperiodic sampled-data and DoS attack[J]. IEEE Transactions on Cybernetics, 2019, 49(4): 1501-1511.
[122] Li B, Wang Z D, Han Q L, et al. Distributed quasiconsensus control for stochastic multiagent systems under round-robin protocol and uniform quantization[J]. IEEE Transactions on Cybernetics, 2022, 52(7): 6721-6732.
[123] Xu J C, Niu Y G, Lv X Y, et al. Sliding mode consensus control for multi-agent systems under component-based weighted try-once-discard protocol[J]. International Journal of Systems Science, 2023, 54(12): 2566-2578.
[124] Tang Y, Zhang D D, Shi P, et al. Event-based

formation control for nonlinear multiagent systems under DoS attacks[J]. IEEE Transactions on Automatic Control, 2021, 66(1): 452-459.

[125] Feng Z, Hu G Q. Secure cooperative event-triggered control of linear multiagent systems under DoS attacks[J]. IEEE Transactions on Control Systems Technology, 2020, 28(3): 741-752.

[126] Åarzén K E. A simple event-based PID controller[J]. IFAC Proceedings Volumes, 1999, 32(2): 8687-8692.

[127] Liu H. Event-triggering-based leader-following bounded consensus of multi-agent systems under DoS attacks[J]. Communications in Nonlinear Science and Numerical Simulation, 2020, 89: 105342.

[128] He W L, Mo Z K. Secure event-triggered consensus control of linear multiagent systems subject to sequential scaling attacks[J]. IEEE Transactions on Cybernetics, 2022, 52(10): 10314-10327.

[129] Yang Y, Li Y F, Yue D. Event-trigger-based consensus secure control of linear multi-agent systems under DoS attacks over multiple transmission channels[J]. Science China Information Sciences, 2020, 63(5): 150208.

[130] Zhang T Y, Ye D. Distributed event-triggered control for multi-agent systems under intermittently random denial-of-service attacks[J]. Information Sciences, 2021, 542: 380-390.

[131] Xu W Y, Ho D W C, Zhong J, et al. Event/self-triggered control for leader-following consensus over unreliable network with DoS attacks[J]. IEEE Transactions on Neural Networks and Learning Systems, 2019, 30(10): 3137-3149.

[132] Du S L, Xu W Y, Qiao J F, et al. Resilient output synchronization of heterogeneous multiagent systems with DoS attacks under distributed event-/self-triggered control[J]. IEEE Transactions on Neural Networks and Learning Systems, 2023, 34(3): 1169-1178.

[133] Li W H, Zhang H G, Zhang J, et al. Fully distributed event/self-triggered formation tracking for multiagent systems under denial-of-service attacks[J]. IEEE Systems Journal, 2023, 17(4): 5706-5713.

[134] Girard A. Dynamic triggering mechanisms for event-triggered control[J]. IEEE Transactions on Automatic Control, 2015, 60(7): 1992-1997.

[135] Du S L, Sheng H, Ho D W C, et al. Secure consensus of multiagent systems with DoS attacks via fully distributed dynamic event-triggered control[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2023, 53(10): 6588-6597.

[136] Xing M P, Lu J Q, Liu Y, et al. Event-based bipartite consensus of multi-agent systems subject to DoS attacks[J]. IEEE Transactions on Network Science and Engineering, 2023, 10(1): 68-80.

[137] Zhang Y F, Wu Z G, Shi P. Resilient event-/self-triggering leader-following consensus control of multiagent systems against DoS attacks[J]. IEEE Transactions on Industrial Informatics, 2023, 19(4): 5925-5934.

[138] Yang H J, Ye D. Observer-based fixed-time secure tracking consensus for networked high-order multiagent systems against DoS attacks[J]. IEEE Transactions on Cybernetics, 2022, 52(4): 2018-2031.

[139] Ni J K, Duan F Y, Shi P. Fixed-time consensus tracking of multiagent system under DOS attack with event-triggered mechanism[J]. IEEE Transactions on Circuits and Systems I: Regular Papers, 2022, 69(12): 5286-5299.

[140] Hui M H, Liu X Y, Cao J D. Improved fixed-time event-triggered average consensus of multi-agent systems under DoS attacks[J]. IEEE Transactions on Circuits and Systems II: Express Briefs, 2024, 71(8): 3815-3819.

[141] Liu Y, Zhou S W, Long J, et al. Event-triggered bipartite consensus of linear multiagent systems under DoS attacks[J]. IEEE Transactions on Control of Network Systems, 2024, 11(3): 1502-1513.

[142] Wang X, Liu J, Wu Y B, et al. Dynamic event-triggered leader-following bipartite consensus of second-order multi-agent systems under DoS attacks[J]. International Journal of Robust and Nonlinear Control, 2024, 34(15): 10731-10749.

[143] Wang X M, Na J, Niu B, et al. Event-triggered adaptive bipartite secure consensus asymptotic tracking control for nonlinear MASs subject to DoS attacks[J]. IEEE Transactions on Automation Science and Engineering, 2024, 21(3): 3816-3825.

[144] Zhao H R, Shan J J, Peng L, et al. Data-driven event-triggered bipartite consensus for multi-agent systems preventing DoS attacks[J]. IEEE Control Systems Letters, 2023, 7: 1915-1920.

[145] Chen Y J, Zhang G, Liu C K, et al. Privacy-preserving modeling of trajectory data: Secure sharing solutions for trajectory data based on granular computing[J]. Mathematics, 2024, 12(23): 3681.

[146] Li Y J, Tao X F, Zhang X F, et al. Privacy-preserved federated learning for autonomous driving[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(7): 8423-8434.

[147] Tran H Y, Hu J K, Pota H R. A privacy-preserving state estimation scheme for smart grids[J]. IEEE Transactions on Dependable and Secure Computing, 2023, 20(5): 3940-3956.

[148] Zhu D, Zhu H, Wang X Y, et al. An accurate and privacy-preserving retrieval scheme over outsourced medical images[J]. IEEE Transactions on Services Computing, 2023, 16(2): 913-926.

[149] Li T, Zhang J W, Obaidat M S, et al. Energy-efficient and secure communication toward UAV networks[J]. IEEE Internet of Things Journal, 2022, 9(12): 10061-10076.

[150] Bhosale P, Kastner W, Sauter T. AutomationML meets Bayesian networks: A comprehensive safety-security risk assessment in industrial control systems[J]. IEEE Open Journal of the Industrial Electronics Society,

2024, 5: 823-835.

[151] Wang X Y, Gursoy M C. Resilient path planning for UAVs in data collection under adversarial attacks[J]. IEEE Transactions on Information Forensics and Security, 2023, 18: 2766-2779.

[152] Luo J C, Liu Z Q, Zhou M C, et al. Deadlock-free scheduling of flexible assembly systems based on Petri nets and local search[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2020, 50(10): 3658-3669.

[153] Lian Y D, Xie W, Zhang L W. A probabilistic time-constrained based heuristic path planning algorithm in warehouse multi-AGV systems[J]. IFAC-PapersOnLine, 2020, 53(2): 2538-2543.

作者简介

卢剑权 (1981−), 男, 教授, 博士, 主要研究方向为网络耦合系统的动态分析、协同控制、安全及相关应用, E-mail: jqluma@seu.edu.cn;

邢梦平 (1993−), 女, 博士后, 主要研究方向为网络化系统安全滤波与控制、多智能体系统一致性分析, E-mail: 13856079203@163.com;

张晶 (1994−), 女, 讲师, 博士, 主要研究方向为多智能体系统一致性控制、安全与隐私保护, E-mail: jingzhangrz@126.com.

科研团队简介

卢剑权教授研究团队依托东南大学复杂系统与网络科学研究中心、江苏省网络群体智能重点实验室、江苏国家应用数学 (东南大学) 中心, 长期专注复杂网络系统建模与控制、网络耦合系统的动态分析、协同控制、安全及相关应用的科学研究工作. 团队主持承担江苏省攀登项目, 江苏省杰出青年基金, 教育部新世纪优秀人才支持计划, 江苏省“333 高层次人才培养工程”领军人才培养对象. 主持国家自然科学基金 5 项, 霍英东基金等省部级项目 7 项. 团队负责人卢剑权教授为德国洪堡学者, 复杂系统与网络科学研究中心副主任, IEEE 高级会员. 获江苏省科学技术奖二等奖 (排一) 和一等奖 (排二), 江苏省基础研究领域十大科技进展 (排一), 江苏省数学成就奖. 入选科睿唯安全球高被引科学家和 Elsevier 中国高被引学者榜单. 担任 3 个 SCI 期刊的编委 (JFI、NEPL、NCA), 期刊《系统科学与数学》和《控制与决策》的编委和 4 个 SCI 期刊的客座编辑. 作为第一导师, 指导学生获江苏省优博、江苏省优硕 (4 位)、自动化学会优硕、华人数学家大会 ICCM 硕士毕业论文银奖; 2020 年至今, 作为一作/通讯在 SIAM Journal on Control and Optimization、Automatica、IEEE TAC 三个期刊发表论文 20 余篇; SCI 的 H 指数为 68.