基于全过程隐私保护的多智能体系统平均一致性

纪良浩 [1] 唐少洪 [1] 郭 兴 [1] 解 燕 [1]

摘 要 针对通信网络可能遭受多邻居联合窃听的多智能体系统, 研究其基于全过程隐私保护的平均一致性问题, 具体包括保护智能体的初始状态以及智能体在实现平均一致性整个过程中的实时状态. 不同于现有的隐私保护平均一致性算法仅能保护智能体的初始状态且无法抵御联合窃听, 提出基于虚拟子网和非消失扰动的全过程隐私保护平均一致性算法. 在所提算法下, 即使智能体的所有信道都被窃听, 仍然可以实现多智能体系统的平均一致性且智能体的状态可以得到全过程保护. 最后, 通过几个数值仿真实验验证了算法的有效性.

关键词 多智能体系统, 平均一致性, 隐私保护, 全过程隐私, 联合窃听

引用格式 纪良浩, 唐少洪, 郭兴, 解燕. 基于全过程隐私保护的多智能体系统平均一致性. 自动化学报, 2025, **51**(6): 1359−1370

DOI 10.16383/j.aas.c240471 **CSTR** 32138.14.j.aas.c240471

Average Consensus in Multi-agent Systems Based on Whole-process Privacy Protection

JI Liang-Hao[1] TANG Shao-Hong[1] GUO Xing[1] XIE Yan[1]

Abstract This paper investigates the average consensus problem with whole-process privacy protection for multi-agent systems facing potential collaborative eavesdropping from multiple neighbors. The research focuses on protecting both the initial states of agents and their real-time states throughout the entire process of achieving average consensus. Different from existing privacy-preserving average consensus algorithms that only safeguard initial states and cannot resist collaborative eavesdropping, a novel whole-process privacy-preserving average consensus algorithm based on virtual subnetworks and non-vanishing perturbations is proposed. Under the proposed algorithm, even if all communication channels of agents are eavesdropped, the average consensus of the multi-agent system can still be achieved while ensuring whole-process protection of agent states. Finally, several numerical simulation experiments verify the effectiveness of the algorithm.

Key words Multi-agent systems, average consensus, privacy protection, whole-process privacy, collaborative eavesdropping

Citation Ji Liang-Hao, Tang Shao-Hong, Guo Xing, Xie Yan. Average consensus in multi-agent systems based on whole-process privacy protection. *Acta Automatica Sinica*, 2025, **51**(6): 1359−1370

近年来, 研究者们对多智能体系统的平均一致性算法进行了广泛而深入的研究[1]. 这一领域因其广泛的应用前景而备受瞩目: 从分布式决策[2] 到智能电网[3], 再到信息融合[4] 和机器人协同控制[5], 平均一致性算法都起着重要的作用. 然而, 传统的平均一致性算法主要聚焦于实现初始状态的平均值计算, 却往往忽视在实现这一过程中可能存在的隐私泄露问题. 鉴于此, 本文将讨论在一致性问题中引入隐私保护机制的必要性, 并进一步探究基于隐私保护的平均一致性相关问题.

1) 保护初始状态. 初始状态作为信息处理和决策的重要来源, 具有重要的保护价值. 在社交网络中, 关于特定话题的意见是集体形成的, 但个人的初始意见不应公开[6]. 同样地, 在交会问题上, 尽管参与者的共同目标是在某一特定地点实现会面, 但个体的初始位置信息可能隐含着家庭住址等敏感信息[7], 因此理应受到保护.

2) 保护实时状态. 在保护初始状态的基础上, 保护智能体的实时状态也同样重要. 例如, 在军事和国防应用中, 未经授权获取实时位置可能会导致在前往会合点的途中或轨迹上遭到伏击[8]. 在隔离交流微电网领域, 获取有功功率等实时运行参数可

收稿日期 2024-07-03 录用日期 2025-01-17
Manuscript received July 3, 2024; accepted January 17, 2025
国家自然科学基金 (62276036), 重庆市自然科学基金创新发展联合基金重点项目 (CSTB2024NSCQ-LZX0118), 重庆市教委科技重大项目 (KJZD-M202100602), 重庆市教委科学技术研究项目 (KJQN202400627) 资助
Supported by National Natural Science Foundation of China (62276036), Innovation and Development Joint Fund Project of Chongqing Natural Science Foundation (CSTB2024NSCQ-LZX0118), Major Project of Scientific and Technological Research Program of Chongqing Municipal Education Commission (KJZD-M202100602), and Science and Technology Research Project of the Chongqing Education Commission (KJQN202400627)
本文责任编委 诸兵
Recommended by Associate Editor ZHU Bing
1. 重庆邮电大学图像认知重庆市重点实验室 重庆 400065
1. Chongqing Key Laboratory of Image Cognition, Chongqing University of Posts and Telecommunications, Chongqing 400065

能被用于定向网络攻击, 进而引发断电事故与重大经济损失[9]. 然而, 现有多数隐私保护一致性研究聚焦于初始状态保护, 忽视实时状态值的保护.

长期以来, 加密方案一直应用于保护多方通信或信息共享系统中的数据隐私, 如 Yao 的乱码电路[10]、Shamir 的秘密共享[11]、安全多方计算[12] 等. 本文聚焦于分布式平均一致性过程中的数据隐私保护问题, 要求智能体在仅通过局部信息交互的条件下, 准确计算整个网络初始状态的均值, 同时确保其初始状态和实时状态不为邻居所知. 现有的方法主要包括: 1) 差分隐私方法[13]; 2) 同态加密方法[14]; 3) 关联扰动方法[15]; 4) 虚拟节点方法[16]; 5) 网络重构方法[17].

差分隐私最早应用于数据库系统中, 之后被用于解决多智能体系统一致性中的隐私保护问题. 差分隐私的核心思想是在共享信息中添加噪声[13, 18−19]. 平均一致性算法旨在计算初始状态的平均值, 但引入的噪声使得整个系统不能达成准确的平均值. 此外, 该方法在隐私性和一致性结果之间存在一个权衡的问题. 当添加的噪声较大时, 无法保证一致性结果的准确性; 当添加的噪声较小时, 隐私保护效果则不理想.

为使多智能体系统能够实现准确的平均一致性, 研究者们提出同态加密、关联扰动、虚拟节点和网络重构等方法. 文献 [14, 20] 提出基于同态加密的平均一致性方法. 文献 [20] 证明当有至少一个可信中心时, 所有智能体的隐私都能得到保护. 文献 [14] 以去中心化的方式实现同态加密在平均一致性下的应用. 然而, 其要求每个智能体至少连接一个合法邻居, 无法抵御联合好奇邻居的推断. 值得一提的是, 文献 [20] 存在智能体的初始状态为整数的限制, 而本文智能体的初始状态可以为任意实数, 适用范围更广. 另外, 同态加密方法显著提高智能体的计算和通信开销, 这使得其在计算能力和通信受限的实际应用中受到限制.

为实现不干扰一致性结果的准确性和以轻量的计算方式实现平均一致性, 研究者们提出关联的扰动信号. 大致可分为以下三类: 加性消失扰动、加性零和扰动、乘法扰动. 加性消失扰动[15, 21−23] 允许智能体传输真实状态与扰动信号之和, 但这些扰动需要随时间衰减为零. 加性零和扰动[24−25] 允许智能体添加任意扰动, 但添加的扰动之和在特定时刻需要归零. 与零和扰动、消失扰动不同, 文献 [26] 指出, 扰动信号不必消失为零, 而是可以收敛到智能体之间事先约定的值. 例如, 添加的扰动可以衰减为相同的常数. 此外, 文献 [27] 设计乘法扰动作为保护初始状态的措施, 然而其为确保收敛性, 只通过添加一段时间的乘法扰动而之后则传递智能体的真实状态, 扰动最后同样消失为零. 上述这些扰动方法忽略对智能体实时状态的保护, 实时状态的泄露导致隐私易于泄露. 当智能体的所有邻居相互串通时, 智能体的初始状态和实时状态都无法得到保护.

与关联扰动策略有所不同, 文献 [16, 28] 提出基于虚拟节点的方法, 该方法具备操作简单和准确性的优势. 然而, 在文献 [16, 28] 中, 每个子网络具有相同结构的同时, 随着时间的推移, 虚拟节点状态值会逐渐趋近于真实节点状态值. 这一特性意味着, 为保障智能体初始值的隐私, 智能体至少需要保证一条边的输入不被窃听. 另一方面, 文献 [17, 29−30] 讨论基于网络重构的方法, 以保护智能体初始状态的隐私性. 其基本思想是通过设计网络拓扑, 使特定节点的可观测性最小化. 然而, 该方法存在以下两点不足: 1) 在实际应用中, 重构整个网络可能会受到限制, 例如需要重新布线或建立新的连接; 2) 重构整个网络较难完成, 例如文献 [30] 需要集中计算整个网络的左特征向量, 求解高次多项式的根也是一个复杂的问题.

值得注意的是, 上述相关工作主要聚焦于保护智能体的初始状态而忽略对智能体实时状态的保护. 如前文所述, 保护智能体的实时状态同样具有重要价值. 因此文献 [31−32] 提出全过程隐私保护的平均一致性算法. 在文献 [31] 中, 通过为每个智能体分配不同的比例系数实现全过程的隐私保护. 此外, 文献 [32] 将多频正弦信号集成到传递信息中, 用以掩盖智能体的真实状态. 然而, 遗憾的是, 即便采用上述两种方法, 在邻居节点联合窃听的情境下, 智能体的初始状态和实时状态依然难以得到有效保护. 鉴于此, 需要设计能抵御联合窃听并实现全过程隐私保护的平均一致性算法.

基于上述讨论, 本文的研究动机如下: 1) 针对上述文献需要至少一条边的输入不被窃听, 即无法抵御联合窃听这一关键问题; 2) 大多数隐私保护方法仅考虑对智能体初始状态值的保护而忽略对实时状态值的保护; 3) 隐私保护方案应易于实现, 不影响一致性结果的准确性, 并且具有计算简单等优势.

受传统虚拟节点方法[16] 构建相同结构虚拟子网的启发, 本文设计一种基于不同结构的虚拟子网和非消失扰动 (由每个智能体根据初始条件独立定义) 的全过程隐私保护平均一致性算法以解决上述问题. 本文的主要贡献总结如下:

1) 针对文献 [14−17, 21−32] 旨在抵御单邻居窃听而不能抵御邻居节点联合窃听这一问题, 本文提出一种具有更强隐私性的分布式平均一致性算法. 具体地, 该算法通过为每个智能体构建不同结构的虚拟子网络, 并让智能体在传递信息中添加独立定义并且非消失的扰动信号, 实现全过程隐私保护的同时也能抵御联合窃听. 此外, 本文给出独立定义

的扰动信号不会干扰一致性结果的充分条件, 解决了邻居节点联合窃听下初始状态和实时状态无法保护的问题.

2) 本文在平均一致性问题中引入更全面的隐私保护和隐私性定义, 除保护智能体的初始状态以外, 同时保护智能体的实时状态. 与文献 [13] 基于差分隐私和文献 [30] 基于可观测性的隐私性定义不同, 在本文所给出的隐私性定义下, 好奇的智能体根据收集到的信息, 既不能推测其他智能体的初始状态和实时状态, 也不能推测其他智能体的初始状态和实时状态所属的有界集合.

3) 与基于同态加密的方法[20] 相比, 本文所提出的算法具有易于实施、计算量小等优势以及更好的适用性.

本文的内容结构如下: 第 1 节介绍本文用到的图论、窃听攻击模型和隐私性定义等相关预备知识; 第 2 节对联合窃听问题进行分类并对算法进行详细说明, 分别对算法的准确性以及隐私性进行分析; 第 3 节通过四个数值仿真实验验证所提算法的有效性; 第 4 节对全文进行总结, 指出本文算法的不足并对未来工作提出设想.

1 预备知识

$\mathbf{R}$ 表示实数域. $\mathbf{R}^N$ 表示 N 维列向量集合, $\mathbf{R}^{N\times N}$ 表示所有 $N\times N$ 实矩阵的集合. $\boldsymbol{c}_N\in\mathbf{R}^N$ 表示 N 维的全 c 列向量. $\mathrm{diag}\{d_1,\,d_2,\,\cdots,\,d_n\}$ 表示具有对角元素 $d_1,\,d_2,\,\cdots,\,d_n$ 的对角矩阵.

1.1 图论知识

网络的通信拓扑图定义为 $\mathcal{G}=\{\mathcal{V},\,\mathcal{E},\,A\}$. 其中, 符号 $\mathcal{V}=\{1,\,2,\,\cdots,\,N\}$ 表示节点集, $\mathcal{E}$ 表示边集, $A=[a_{ij}]\in\mathbf{R}^{N\times N}$ 对应于加权邻接矩阵. 对于所有 $i=1,\,2,\,\cdots,\,N$, $a_{ii}=0$. $a_{ij}>0$ 表示节点 j 可以向节点 i 传递信息. 若节点 j 没有信息传向节点 i 则 $a_{ij}=0$. 拉普拉斯矩阵 $L=[l_{ij}]\in\mathbf{R}^{N\times N}$, 其中, $l_{ij}=-a_{ij},\,i\neq j$, $l_{ii}=\sum_{j=1}^{N}a_{ij}$.

在图 $\mathcal{G}$ 中, 从 i_1 到 i_m 的路径是一连串具有不同节点的边 $(i_1,\,i_2),\,\cdots,\,(i_{m-1},\,i_m)$.

假设 1. 网络拓扑图 $\mathcal{G}$ 是强连通且权重平衡的, 即对于任意节点 i 和 j, 至少存在一条从 i 到 j 的路径, 且 $\sum_{j=1}^{N}a_{ji}=\sum_{j=1}^{N}a_{ij}$.

1.2 连续时间平均一致性算法

传统的连续时间平均一致性算法[33] 如下:

$$\dot{x}_i(t)=\sum_{j=1}^{N}a_{ij}(x_j(t)-x_i(t)) \tag{1}$$

其中, $i\in\mathcal{V}$. 基于假设 1, 式 (1) 能实现准确的平均一致性.

1.3 窃听攻击模型

定义 1 (好奇智能体)[16]. 好奇智能体 (节点) 是指正确遵循算法所有步骤的智能体, 但试图根据收集到的信息获取其他智能体的初始状态和实时状态. 好奇智能体之间可以相互串通, 即共享已知信息.

1.4 隐私性定义

定义 I 为智能体 i 在初始状态 $x_i(0)$ 下达成一致, 其邻居可以获得的关于 i 的信息集.

考虑智能体 i 的另一种初始状态 $\bar{x}_i(0)\neq x_i(0)$, 在 $\bar{x}_i(0)$ 下与邻居进行信息交互并达成一致, 智能体 i 的邻居可获得的所有信息定义为 $\bar{I}$. 下面给出隐私性定义.

定义 2 (隐私性). 如果存在无穷多个 $\bar{x}_i(0)\neq x_i(0)$(或实时状态 $\bar{x}_i(t)\neq x_i(t),\,t>0$) 和任意大的实数 c, 使得 $|\bar{x}_i(0)-x_i(0)|>c$ (或 $|\bar{x}_i(t)-x_i(t)|>c$) 和 $\bar{I}\equiv I$ 同时成立, 则称智能体 i 的初始状态 (或实时状态) 是隐私的.

注 1. 遵守这一隐私性定义意味着, 无法根据完全相同的信息 $\bar{I}$ 和 I 判断智能体 i 的初始状态更有可能为 $\bar{x}_i(0)$ 还是 $x_i(0)$. 另外, 智能体 i 的实时状态同样如此. 注意到 c 可以任意大, 因此无穷多个 $\bar{x}_i(0)$ (或 $\bar{x}_i(t)$) 构成的集合是无界的.

值得注意的是, 例如式 (1) 中实时状态趋于相同的方式无法满足定义 2. 因此, 为实现全过程的隐私保护, 本文采用文献 [31] 给出的输出平均一致性定义.

1.5 平均一致性定义

令 $y_i(t)$ 表示智能体 i 在 t 时刻的输出, 下面给出文献 [31] 的平均一致性定义.

定义 3 (输出平均一致性)[31]. 对于一个由 N 个智能体 (节点) 组成的拓扑图 $\mathcal{G}$ 所代表的多智能体分布式网络, 如果 $\forall i\in\mathcal{V}$, 有 $\lim_{t\to\infty}y_i(t)=\frac{1}{N}\sum_{i=1}^{N}x_i(0)$, 则称整个多智能体系统实现准确的平均一致性.

1.6 问题描述

基于上述描述, 每个智能体都将贡献一个初始值 $x_i(0)$, 本文的主要目标是提出一个新的隐私保护平均一致性算法, 从而实现以下三个目标:

1) 隐私性. 即使智能体在一致性过程中的所有信道都被窃听, 无论是其初始状态还是实时状态都能得到保护.

2) 准确性. 即使在隐私保护算法的操作下, 仍然确保网络中的每个智能体都能实现准确的平均一致性 ($\lim_{t\to\infty} y_i(t) = \frac{1}{N}\sum_{i=1}^{N} x_i(0)$).

3) 算法简单, 易于实现且计算量级较轻.

2 主要结论

本文考虑的联合窃听问题如图 1 ~ 4 所示. 图中绿色节点表示中立的智能体, 橙色节点表示好奇智能体.

在图 1 中, 智能体 i 存在一个好奇的邻居, 表示为非联合窃听; 图 2 表示智能体 i 遭受联合窃听但还存在至少一个中立邻居的情况 (弱联合窃听); 图 3 表示智能体 i 的所有邻居都是好奇智能体并且相互串通, 但整个网络中还存在中立的节点 (强联合窃听); 图 4 则表示除智能体 i 以外, 整个网络中

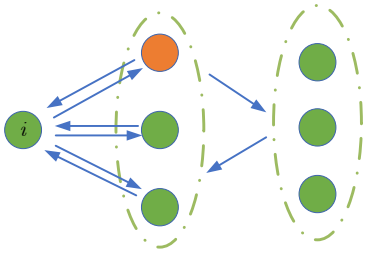

图 1 非联合窃听下的网络示意图

Fig. 1 Schematic of the network under non-collaborative eavesdropping

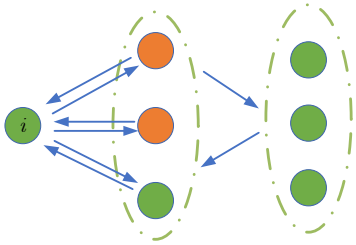

图 2 弱联合窃听下的网络示意图

Fig. 2 Schematic of the network under weak collaborative eavesdropping

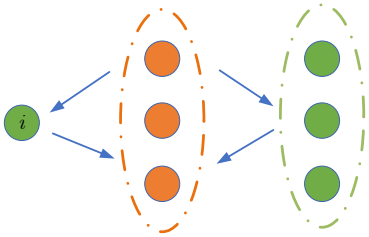

图 3 强联合窃听下的网络示意图

Fig. 3 Schematic of the network under strong collaborative eavesdropping

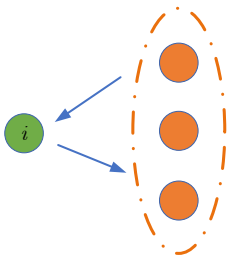

图 4 完全联合窃听下的网络示意图

Fig. 4 Schematic of the network under full collaborative eavesdropping

的其他所有智能体都是好奇智能体并且相互串通的情况 (完全联合窃听).

当智能体 i 的邻居中存在好奇智能体时, 需要保护智能体 i 的初始状态并且确保智能体 i 实时状态的隐私性.

2.1 隐私保护方案设计

为解决上述问题, 本文所提算法需要一个预准备阶段, 用于为每个智能体构建不同的虚拟子网络. 该准备阶段如下:

第三方随机产生 N 个正数 $m_1, m_2, \cdots, m_N$, 并计算 $S = (N + \sum_{i=1}^{N} m_i)/N$, 然后将 S 和这些参数随机地分配给 N 个智能体.

注 2. 虽然需要第三方来生成这些参数, 但该过程只需执行一次. 此外, 该操作只是为保护隐私, 与后续的一致性过程无关. 这并不影响多智能体系统的分布式设置, 平均一致性的实现仍然不需要中心节点. 在文献 [31] 中, 需要引入第三方来分配比例系数. 另外, 在文献 [30] 中, 有向网络下的网络参数重构也需要第三方实现. 除此之外, 参数 m_i 同样可以由网络的构建者分配. 对比文献 [20] 分配公钥和私钥, 这一要求更简单. 因此本文引入第三方分配参数的这一需求是合理的.

为便于对本文方法的理解以及后续工作的展开, 以图 5 为例, 阐述所提算法的核心思想.

从图 5 可以看出, 每个节点运行着一个自身的虚拟子网络. 对节点 i 的邻居而言, 唯一可见的为节点 i 的输出信息 $y_i^{\alpha}(t)$ 和原始网络拓扑权重 a_{ij}. 将真实的实时状态 $x_i^{\alpha}(t)$ 与虚拟状态 $x_i^{\beta}(t)$ 之间的权重 $a_{i,\,\alpha\beta}$ 和 $a_{i,\,\beta\alpha}$ 称为虚拟耦合权重. 虚拟耦合权重 $a_{i,\,\alpha\beta}$ 和 $a_{i,\,\beta\alpha}$ 由节点 i 自身任意选择, 满足 $a_{i,\,\alpha\beta} > 0$, $a_{i,\,\beta\alpha} > 0$ 且 $a_{i,\,\alpha\beta}/a_{i,\,\beta\alpha} = m_i > 0$. 为避免实时状态信息 $x_i^{\alpha}(t)$ 的泄露, 节点 i 的输出信息为 $y_i^{\alpha}(t) = x_i^{\alpha}(t) + \theta_i^{\alpha}(t)$, 其中 $\theta_i^{\alpha}(t)$ 为节点 i 在传递信息中添加的扰动信号. 将节点 i 的虚拟节点的输出和扰动信号分别定义为 $y_i^{\beta}(t)$, $\theta_i^{\beta}(t)$, 满足 $y_i^{\beta}(t) = x_i^{\beta}(t) +$

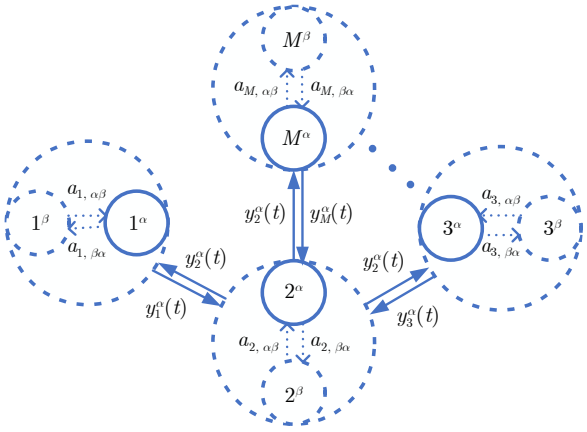

图 5　算法示例图

Fig. 5　Schematic diagram of the algorithm example

$\theta_i^\beta(t)$. 扰动信号 $\theta_i^\alpha(t)$ 和 $\theta_i^\beta(t)$ 由节点 i 独立定义且无需与其他智能体事先约定. 由于扰动信号存在一些限制, 下面给出如何设计有效扰动的一般形式.

定义 4 (有效扰动). 当扰动信号 $\theta_i(t)$ 同时满足条件 1) $\forall t\in[0,\,\infty)$, $\theta_i(t)$ 有界且连续及 2) $\lim_{t\to\infty}\theta_i(t)=\theta_i\in\mathbf{R}/\{0\}$ 时, 称扰动信号 $\theta_i(t)$ 为有效扰动.

注 3. 不同于文献 [15, 21−27] 中扰动信号存在共同消失为零或收敛为约定值的限制, 本文的扰动信号无这一限制以确保智能体的实时状态具有更好的隐私性.

为确保整个网络能够实现准确的平均一致性, 每个节点进行如下的状态分配:

$$x_i^\alpha(0)+\theta_i^\alpha+m_i(x_i^\beta(0)+\theta_i^\beta)=Sx_i(0)\tag{2}$$

由式 (2) 可知, 初始状态不仅分布在初始实时状态 $x_i^\alpha(0)$ 和虚拟节点初始状态 $x_i^\beta(0)$ 上, 还分布在扰动信号 $\theta_i^\alpha(t)$ 和 $\theta_i^\beta(t)$ 最终的收敛值 θ_i^α 和 θ_i^β 上. $x_i^\alpha(0)$, $x_i^\beta(0)$, $\theta_i^\alpha(t)$, $\theta_i^\beta(t)$ 在满足式 (2) 的条件下由节点 i 任意选择并完成分配.

考虑每个智能体的动力学方程如下:

$$\begin{cases}\dot{x}_i^\alpha(t)=\displaystyle\sum_{j=1}^{N}a_{ij}(y_j^\alpha(t)-y_i^\alpha(t))+\\ \qquad\qquad a_{i,\,\alpha\beta}(y_i^\beta(t)-y_i^\alpha(t))\\ \dot{x}_i^\beta(t)=a_{i,\,\beta\alpha}(y_i^\alpha(t)-y_i^\beta(t))\end{cases}\tag{3}$$

本文所提出的连续时间下的平均一致性算法如算法 1 所示.

算法 1. 全过程隐私的分布式平均一致性算法

输入. 节点 i 的初始状态 $x_i(0)$.

输出. 节点 i 的输出 $y_i^\alpha(t)$.

步骤 1. 权重生成: 智能体 i 生成虚拟耦合权重 $a_{i,\,\alpha\beta}$ 和 $a_{i,\,\beta\alpha}>0$, 满足 $a_{i,\,\alpha\beta}/a_{i,\,\beta\alpha}=m_i>0$.

步骤 2. 扰动生成: 智能体 i 生成一对有效扰动 $\theta_i^\alpha(t)$ 和 $\theta_i^\beta(t)$.

步骤 3. 状态分配: 智能体 i 在满足式 (2) 的条件下, 分配其初始状态为 $x_i^\alpha(0)$, $x_i^\beta(0)$, θ_i^α, θ_i^β.

步骤 4. 本地计算: 智能体 i 传递 $y_i^\alpha(t)$ 给满足 $a_{ji}>0$ $(j=1,\,2,\,\cdots,\,N)$ 的邻居. 每个智能体根据式 (3) 更新实时状态 $x_i^\alpha(t)$ 和虚拟状态 $x_i^\beta(t)$.

注 4. 算法中的扰动信号 $\theta_i^\alpha(t)$, $\theta_i^\beta(t)$ 由每个智能体根据式 (2) 独立分配, 这意味着每个节点的扰动信号对于邻居智能体而言是未知的. 扰动可以是满足定义 4 的任意轨迹的连续函数. 例如, 智能体 i 可以将扰动 $\theta_i^\alpha(t)$ 选择为如下两种形式:

$$\theta_i^\alpha(t)=b_i\mathrm{e}^{-\sigma_i t}+\theta_i^\alpha$$

$$\theta_i^\alpha(t)=\begin{cases}(\sigma_i-t)b_i+\theta_i^\alpha, & t\in[0,\,\sigma_i)\\ \theta_i^\alpha, & t\in[\sigma_i,\,\infty)\end{cases}$$

其中, $b_i\neq0$, $\sigma_i>0$, $\theta_i^\alpha\neq0$. 此外, 其他满足定义 4 的扰动同样是允许的.

注 5. 即便各个智能体的虚拟权重比 m_i 均不相同, 且每个智能体的扰动信号是独立生成的, 式 (2) 的条件仍然是使得整个多智能体系统 (式 (3)) 能够实现准确平均一致性的充分条件.

注 6. 在智能体传递给邻居的信息中添加独立定义的非消失扰动的目的是让实时状态 $x_i^\alpha(t)$ 在任意联合窃听下是隐私的. 而不同结构的虚拟子网络和独立定义的非消失扰动信号使得智能体的初始状态 $x_i(0)$ 在强联合窃听下是隐私的. 具体地, 智能体 i 的初始状态 $x_i(0)$ 和实时状态 $x_i^\alpha(t)$ 的变化总是可以通过实时初始状态 $x_i^\alpha(0)$、虚拟状态 $x_i^\beta(0)$、扰动 $\theta_i^\alpha(t)$, $\theta_i^\beta(t)$ 和虚拟权重比 m_i 的变化完全补偿, 从而使得联合好奇节点获得的信息完全相同.

接下来给出算法 1 的准确性分析.

2.2　算法的准确性分析

定理 1. 基于假设 1, 在算法 1 的作用下, 多智能体系统 (3) 能实现准确的平均一致性, 即网络中的每个智能体都能获得准确的平均一致性值 $(\lim_{t\to\infty}y_i^\alpha(t)=\frac{1}{N}\sum_{i=1}^{N}x_i(0),\,\forall i\in\mathcal{V})$.

证明. 令

$$A_1=\mathrm{diag}\{a_{1,\,\alpha\beta},\,a_{2,\,\alpha\beta},\,\cdots,\,a_{N,\,\alpha\beta}\}$$

$$A_2=\mathrm{diag}\{a_{1,\,\beta\alpha},\,a_{2,\,\beta\alpha},\,\cdots,\,a_{N,\,\beta\alpha}\}$$

$\tilde{L}$ 表示为

$$\tilde{L}=\begin{bmatrix}L+A_1 & -A_1\\ -A_2 & A_2\end{bmatrix}\tag{4}$$

由式 (4) 给出整个系统 (3) 的紧凑形式如下:

$$\dot{\tilde{\boldsymbol{x}}}(t) = -\tilde{L}\tilde{\boldsymbol{y}}(t) \tag{5}$$

其中,

$$\begin{aligned}\tilde{\boldsymbol{y}}(t) = [&y_1^{\alpha}(t),\, y_2^{\alpha}(t),\, \cdots,\, y_N^{\alpha}(t),\\ &y_1^{\beta}(t),\, y_2^{\beta}(t),\, \cdots,\, y_N^{\beta}(t)]^{\mathrm{T}}\\ \tilde{\boldsymbol{x}}(t) = [&x_1^{\alpha}(t),\, x_2^{\alpha}(t),\, \cdots,\, x_N^{\alpha}(t),\\ &x_1^{\beta}(t),\, x_2^{\beta}(t),\, \cdots,\, x_N^{\beta}(t)]^{\mathrm{T}}\end{aligned}$$

由假设 1 可知, $\underbrace{[\frac{1}{N},\, \frac{1}{N},\, \cdots,\, \frac{1}{N}]}_{N}$ 为拉普拉斯矩阵 L 的零特征值对应的归一化的左特征向量. 而 $A_1 = \mathrm{diag}\{m_1,\, \cdots,\, m_N\}A_2$, 因此有如下等式成立:

$$\underbrace{\left[\frac{1}{N}, \frac{1}{N}, \cdots, \frac{1}{N}, \frac{m_1}{N}, \frac{m_2}{N}, \cdots, \frac{m_N}{N}\right]}_{2N}\tilde{L} = \underbrace{[0, \cdots, 0]}_{2N} \tag{6}$$

由式 (6) 可得

$$\boldsymbol{u} = \underbrace{\left[\frac{1}{N}, \frac{1}{N}, \cdots, \frac{1}{N}, \frac{m_1}{N}, \frac{m_2}{N}, \cdots, \frac{m_N}{N}\right]}_{2N}$$

为 $\tilde{L}$ 的零特征值的一个左特征向量. 注意到 $\tilde{L}$ 矩阵的零特征值的左特征向量只需满足前 N 个元素相同、后 N 个元素比前 N 个元素分别是 m_i 倍数即可. 因此可将向量 $\boldsymbol{u}$ 改写为

$$\boldsymbol{v} = \left[\frac{1}{N + \sum\limits_{i=1}^{N} m_i}, \cdots, \frac{1}{N + \sum\limits_{i=1}^{N} m_i}, \frac{m_1}{N + \sum\limits_{i=1}^{N} m_i}, \cdots, \frac{m_N}{N + \sum\limits_{i=1}^{N} m_i}\right]$$

$\boldsymbol{v}$ 仍然为 $\tilde{L}$ 的零特征值的左特征向量并且是归一化的 (向量中每一个元素相加之后的和为 1).

此外, $\tilde{L}$ 的零特征值的一个右特征向量为 $\boldsymbol{w}_r = [\frac{1}{\sqrt{N+\sum_{i=1}^{N} m_i}}, \cdots, \frac{1}{\sqrt{N+\sum_{i=1}^{N} m_i}}]^{\mathrm{T}}$, 一个左特征向量为

$$\boldsymbol{w}_l = \left[\frac{1}{\sqrt{N + \sum\limits_{i=1}^{N} m_i}}, \cdots, \frac{1}{\sqrt{N + \sum\limits_{i=1}^{N} m_i}}, \frac{m_1}{\sqrt{N + \sum\limits_{i=1}^{N} m_i}}, \cdots, \frac{m_N}{\sqrt{N + \sum\limits_{i=1}^{N} m_i}}\right]$$

满足 $\boldsymbol{w}_l\boldsymbol{w}_r = 1$ 且 $\boldsymbol{w}_r\boldsymbol{w}_l = \boldsymbol{1}_{2N}\boldsymbol{v}$.

令 $\tilde{\boldsymbol{\theta}}(t) = [\theta_1^{\alpha}(t),\, \cdots,\, \theta_N^{\alpha}(t),\, \theta_1^{\beta}(t),\, \cdots,\, \theta_N^{\beta}(t)]^{\mathrm{T}} \in \mathbf{R}^{2N}$, $\tilde{\boldsymbol{\theta}} = [\theta_1^{\alpha},\, \cdots,\, \theta_N^{\alpha},\, \theta_1^{\beta},\, \cdots,\, \theta_N^{\beta}]^{\mathrm{T}} \in \mathbf{R}^{2N}$. 根据定义 4, 可知:

$$\begin{aligned}\lim_{t\to\infty}\dot{\tilde{\boldsymbol{y}}}(t) = \lim_{t\to\infty}\dot{\tilde{\boldsymbol{x}}}(t) = \lim_{t\to\infty} -\tilde{L}\tilde{\boldsymbol{y}}(t) =\\ \lim_{t\to\infty} -\tilde{L}(\tilde{\boldsymbol{x}}(t) + \tilde{\boldsymbol{\theta}})\end{aligned} \tag{7}$$

由式 (7) 和 $\tilde{L}$ 为强连通有向图的拉普拉斯矩阵可得

$$\begin{aligned}\lim_{t\to\infty}\tilde{\boldsymbol{y}}(t) = \lim_{t\to\infty}\mathrm{e}^{-\tilde{L}t}\tilde{\boldsymbol{y}}(t) = \lim_{t\to\infty}\boldsymbol{w}_r\boldsymbol{w}_l\tilde{\boldsymbol{y}}(t) =\\ \lim_{t\to\infty}\boldsymbol{1}_{2N}\boldsymbol{v}\tilde{\boldsymbol{y}}(t) = \lim_{t\to\infty}\boldsymbol{v}(\tilde{\boldsymbol{x}}(t) + \tilde{\boldsymbol{\theta}})\boldsymbol{1}_{2N}\end{aligned} \tag{8}$$

同时有下式成立

$$\boldsymbol{v}\dot{\tilde{\boldsymbol{x}}}(t) = -\boldsymbol{v}\tilde{L}\tilde{\boldsymbol{y}}(t) = 0 \tag{9}$$

联立式 (2)、(8) 和 (9) 可得

$$\lim_{t\to\infty}\tilde{\boldsymbol{y}}(t) = \lim_{t\to\infty}\boldsymbol{v}(\tilde{\boldsymbol{x}}(0) + \tilde{\boldsymbol{\theta}})\boldsymbol{1}_{2N} = \frac{1}{N}\sum_{i=1}^{N} x_i(0)\boldsymbol{1}_{2N} \tag{10}$$

由式 (10) 可知, 每个智能体的输出达成平均一致性, 即网络中的所有节点都能获得准确的平均一致性值. □

2.3 联合窃听下的隐私性分析

假定智能体 i 由第三方分配的参数 m_i 是对其邻居未知的, 网络中的所有公开参数都是其邻居已知的. 将智能体 j 的知识集合定义为 I_j, 具体形式如下:

$$\begin{aligned}I_j = \Big\{&x_j^{\alpha}(t),\, x_j^{\beta}(t),\, \theta_j^{\alpha}(t),\, \theta_j^{\beta}(t),\, \forall t > 0;\\ &y_i^{\alpha}(t),\, i \in N_j^{\mathrm{in}},\, \forall t > 0;\\ &a_{ij},\, i \in N_j^{\mathrm{out}}; a_{ji},\, i \in N_j^{\mathrm{in}}; m_j,\, S\Big\}\end{aligned} \tag{11}$$

其中, N_j^{in} 表示给智能体 j 发送信息的邻居集合, N_j^{out} 表示接收智能体 j 信息的邻居集合.

接下来讨论在不同程度的联合窃听下, 算法 1 对于智能体 i 的初始状态 $x_i(0)$ 和实时状态 $x_i^{\alpha}(t)$ 的保护情况.

定理 2. 基于假设 1, 在算法 1 的作用下, 如下结论成立:

1) 非、弱和强联合窃听下智能体 i 的初始状态 $x_i(0)$ 是隐私的;

2) 完全联合窃听下智能体 i 的初始状态 $x_i(0)$ 无法得到保护;

3) 完全联合窃听下智能体 i 的实时状态 $x_i^{\alpha}(t)$, $\forall t \in [0,\, \infty)$ 是隐私的.

证明. 首先证明 1). 在非联合窃听和弱联合窃听下, 好奇的联合节点不能获得节点 i 所有的输入信息 $y_j^\alpha(t),\ \forall j \in N_i^{\text{in}}$, 而强联合窃听下可以获得所有的 $y_j^\alpha(t),\ \forall j \in N_i^{\text{in}}$. 这意味着强联合窃听包含比非联合窃听和弱联合窃听更多的关于节点 i 的信息. 因此只需要证明在包含节点 i 更多信息的强联合窃听下, 节点 i 的初始状态 $x_i(0)$ 是隐私的, 即证明在非联合窃听和弱联合窃听下, 节点 i 的初始状态 $x_i(0)$ 是隐私的. 而强联合窃听可以简化为如图 6 所示的网络拓扑. 不妨假设节点 2 为好奇节点, 好奇节点 2 试图根据其收集到的信息推测节点 1 和节点 3 的初始状态.

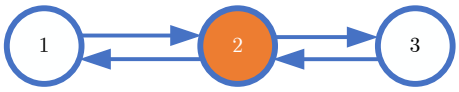

图 6　强联合窃听下的简化示意图

Fig. 6　Simplified schematic under strong collaborative eavesdropping

考虑节点 1 和节点 3 在算法 1 下的两种不同执行情形. 节点 1 和节点 3 的初始状态在第一种情形下分别为 $x_1(0)$ 和 $x_3(0)$, 在第二种情形下分别为 $\bar{x}_1(0)$ 和 $\bar{x}_3(0)$. 第二种情形下的算法参数也通过在字母上方添加横线与第一种情形进行区分, 例如定义 $\bar{x}_i^\alpha(t)$ 为不同于 $x_i^\alpha(t)$ 的另一种情形, 但都表示智能体 i 的实时状态. 令 $\Delta x_i^\alpha(t) = \bar{x}_i^\alpha(t) - x_i^\alpha(t)$, 具体地, 两种不同情形的初始状态分别如下:

$$\bar{x}_1(0) = x_1(0) + \frac{1}{S}(c_1^\alpha - \lim_{t\to\infty}\Delta x_1^\alpha(t) + \delta(x_1^\beta(0) + \theta_1^\beta) + (m_1+\delta)(c_1^\beta + \lim_{t\to\infty} f_1(t))) \tag{12a}$$

$$\bar{x}_3(0) = x_3(0) + \frac{1}{S}(c_3^\alpha - \lim_{t\to\infty}\Delta x_3^\alpha(t) - \delta(x_3^\beta(0) + \theta_3^\beta) + (m_3-\delta)(c_3^\beta + \lim_{t\to\infty} f_3(t))) \tag{12b}$$

其中, $0 < \delta < m_3$; c_1^α, c_1^β, c_3^α, c_3^β 表示不为零的实数; $f_1(t)$, $f_3(t)$ 表示满足定义 4 的任意函数. 此外, 两种不同情形下的算法参数分别为

$$\bar{m}_1 = m_1 + \delta \tag{13a}$$

$$\bar{a}_{1,\,\alpha\beta} = (m_1+\delta)a_{1,\,\beta\alpha} \tag{13b}$$

$$\bar{m}_3 = m_3 - \delta \tag{13c}$$

$$\bar{a}_{3,\,\alpha\beta} = (m_3-\delta)a_{3,\,\beta\alpha} \tag{13d}$$

$$\bar{x}_1^\alpha(0) = x_1^\alpha(0) + c_1^\alpha \tag{13e}$$

$$\bar{\theta}_1^\alpha(t) = \theta_1^\alpha(t) - \Delta x_1^\alpha(t) \tag{13f}$$

$$\bar{x}_1^\beta(0) = x_1^\beta(0) + c_1^\beta \tag{13g}$$

$$\bar{\theta}_1^\beta(t) = \theta_1^\beta(t) + f_1(t) \tag{13h}$$

$$\bar{x}_3^\alpha(0) = x_3^\alpha(0) + c_3^\alpha \tag{13i}$$

$$\bar{\theta}_3^\alpha(t) = \theta_3^\alpha(t) - \Delta x_3^\alpha(t) \tag{13j}$$

$$\bar{x}_3^\beta(0) = x_3^\beta(0) + c_3^\beta \tag{13k}$$

$$\bar{\theta}_3^\beta(t) = \theta_3^\beta(t) + f_3(t) \tag{13l}$$

其中

$$\Delta x_1^\alpha(t) = \Delta x_1^\alpha(0) + \int_0^t \bar{a}_{1,\,\alpha\beta}(\bar{y}_1^\beta(\tau) - \bar{y}_1^\alpha(\tau))\mathrm{d}\tau - \int_0^t a_{1,\,\alpha\beta}(y_1^\beta(\tau) - y_1^\alpha(\tau))\mathrm{d}\tau \tag{14a}$$

$$\Delta x_3^\alpha(t) = \Delta x_3^\alpha(0) + \int_0^t \bar{a}_{3,\,\alpha\beta}(\bar{y}_3^\beta(\tau) - \bar{y}_3^\alpha(\tau))\mathrm{d}\tau - \int_0^t a_{3,\,\alpha\beta}(y_3^\beta(\tau) - y_3^\alpha(\tau))\mathrm{d}\tau \tag{14b}$$

好奇节点 2 完整的信息集合由式 (15) 给出

$$\begin{aligned} I_2 = \{&x_2^\alpha(t),\ x_2^\beta(t),\ \theta_2^\alpha(t),\ \theta_2^\beta(t),\ \forall t>0;\\ &y_1^\alpha(t),\ y_3^\alpha(t),\ \forall t>0;\\ &a_{12},\ a_{21},\ a_{23},\ a_{32};m_2,\ S\} \end{aligned} \tag{15}$$

联立式 (12) ~ (14) 和定理 1, 有如下等式成立:

$$\bar{m}_1 + \bar{m}_3 = m_1 + m_3 \tag{16}$$

$$\bar{y}_1^\alpha(t) = y_1^\alpha(t),\ \bar{y}_3^\alpha(t) = y_3^\alpha(t),\ \forall t \in [0,\ \infty) \tag{17}$$

$$\bar{x}_1(0) + \bar{x}_3(0) = x_1(0) + x_3(0) \tag{18}$$

式 (16) 表明 $\bar{S} = S$. 此外, 式 (18) 表明在 $\bar{x}_1(0)$, $\bar{x}_3(0)$ 和 $x_1(0)$, $x_3(0)$ 这两种不同情形下, 同样实现准确的平均一致性. 由式 (16) ~ (18) 可得, 好奇节点 2 在两种不同情形下获得的信息完全相同. 即对于给定的信息集合 I_2, 存在节点 1 和节点 3 初始状态和实时状态的无穷多种组合, 使得 $\bar{I}_2 = I_2$. 因此, 在强联合窃听下仍然满足定义 2. 即存在 $\bar{x}_1(0) = x_1(0) + c$, $\bar{x}_3(0) = x_3(0) - c$, c 任意大, 使得 $\bar{I}_2 = I_2$ 成立.

接下来证明 2). 在该情况下, 联合的好奇节点可以通过平均一致性结果乘以节点总数 N, 获得整个系统初始状态的总和 $\sum_{j=1}^N x_j(0)$. 因此, 联合的好奇节点可以通过等式 $x_i(0) = \sum_{j=1}^N x_j(0) - \sum_{j=1,\ j\neq i}^N x_j(0)$ 唯一确定节点 i 初始状态.

下面证明 3). 采取与证明 1) 相同的思路, 考虑节点 i 在算法 1 下的两种不同情形

$$\begin{cases}\bar{x}_i^{\alpha}(0)=x_i^{\alpha}(0)+c,\ c\in\mathbf{R}/\{0\}\\ \bar{\theta}_i^{\alpha}(t)=\theta_i^{\alpha}(t)-c\end{cases}\tag{19}$$

其余条件都相同. 同理, 根据式 (19) 不难验证 $\bar{I}_j=I_j$. 由于 c 可以为任意非零值, 即对于给定的节点 i 的输出轨迹, 存在节点 i 的实时状态值 $x_i^{\alpha}(t)$ 的无穷多种解. 此外, 完全相同的信息使得节点 i 的所有邻居无法判断 $\bar{x}_i^{\alpha}(t)$ 和 $x_i^{\alpha}(t)$ 这两种情况中的哪一种情况更有可能. 即存在 $\bar{x}_i^{\alpha}(t)=x_i^{\alpha}(t)+c$, c 任意大, 使得 $\bar{I}_j\equiv I_j$. □

3 数值仿真实验

考虑一个由 5 个节点组成的多智能体系统 (3), 其通信拓扑如图 7 所示. 图中的每个节点分别制定初始状态值 $x_1(0)=1$, $x_2(0)=3$, $x_3(0)=6$, $x_4(0)=10$, $x_5(0)=15$, 则初始状态的平均值为 $\frac{1}{5}\sum_{i=1}^{5}x_i(0)=7$.

注 7. 在图 7 所示的网络拓扑下, 采用文献 [14−17, 21−32] 所提方法, 好奇的节点 2 能唯一确定节点 1 的初始状态或实时状态. 而在本文所提算法下, 无论是节点 1 的初始状态, 还是节点 1 的实时状态, 对节点 2 而言都是隐私的.

根据算法的准备阶段, 生成 5 个参数 $m_1=1, m_2=2$, $m_3=3$, $m_4=4$, $m_5=5$ 和 $S=(5+\sum_{i=1}^{5}m_i)/5=4$. 然后将每个参数 m_i 和 S 分配给相应的节点, 除相互串通的节点外, 节点 i 只知道自己的参数 m_i.

示例 1 (准确性验证). 设定每个节点的算法参数如下:

$$\begin{cases}m_1=1,\ m_2=2,\ m_3=3,\ m_4=4,\ m_5=5\\ A_1=\mathrm{diag}\{1,\ 4,\ 9,\ 16,\ 25\}\\ A_2=\mathrm{diag}\{1,\ 2,\ 3,\ 4,\ 5\}\\ x^{\alpha}(0)=(1,\ 2,\ 3,\ 4,\ 5)^{\mathrm{T}}\\ x^{\beta}(0)=(1,\ 2,\ 3,\ 4,\ 5)^{\mathrm{T}}\\ \theta_i^{\alpha}(t)=\begin{cases}(i-t)i+i, & t<i\\ i, & t\geq i\end{cases}\\ \theta_i^{\beta}(t)=\begin{cases}(i-t)i+i, & t<i\\ i, & t\geq i\end{cases}\end{cases}\tag{20}$$

其中, $i=1,2,3,4,5$. 在式 (19) 下节点的输出轨迹如图 8 所示, 实线表示节点的输出, 虚线表示其邻居不可见的虚拟节点输出. 由图 8 可知, $\lim_{t\to\infty}y_i^{\alpha}(t)=\sum_{i=1}^{5}x_i(0)=7, \forall i\in\mathcal{V}$, 整个多智能体系统能实现准确的平均一致性.

下面验证所提算法的隐私性. 不妨假定节点 2、3、5 为联合的好奇节点, 试图根据获得的信息推测节点 1 和节点 4 的初始状态和实时状态. 此时对应的联合窃听为强联合窃听. 为验证算法 1 的隐私性, 以下提供算法 1 的一个不同于式 (20) 的执行情形, 这种情形确保节点 1 和节点 4 的初始状态不同, 但节点 1 和节点 4 传递给邻居的信息完全相同.

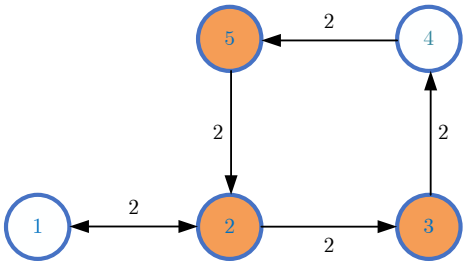

图 7 5 个节点组成的多智能体系统网络拓扑

Fig. 7 Network topology of multi-agent system with 5 nodes

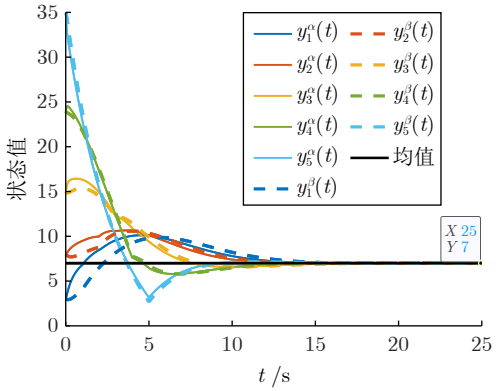

图 8 在式 (20) 下系统各节点输出值的变化轨迹

Fig. 8 Trajectories of changes in the output values of each node of the system under (20)

示例 2 (强联合窃听下初始状态的隐私性验证). 不同于式 (20) 的算法 1 的另一种情形如下:

$$\begin{cases}\bar{m}_1=m_1+1,\ \bar{m}_4=m_4-1\\ \bar{m}_2=m_2,\ \bar{m}_3=m_3,\ \bar{m}_5=m_5\\ \bar{A}_1=\mathrm{diag}\{2,\ 4,\ 9,\ 12,\ 25\}\\ \bar{A}_2=\mathrm{diag}\{1,\ 2,\ 3,\ 4,\ 5\}\\ \bar{x}^{\alpha}(0)=(2,\ 2,\ 3,\ 6,\ 5)^{\mathrm{T}}\\ \bar{x}^{\beta}(0)=(2,\ 2,\ 3,\ 6,\ 5)^{\mathrm{T}}\\ f_1(t)=1,\ f_4(t)=2\end{cases}\tag{21}$$

在该情形下, 根据式 (13) 给出 $\bar{\theta}_1^{\alpha}(t)$, $\bar{\theta}_1^{\beta}(t)$, $\bar{\theta}_4^{\alpha}(t)$, $\bar{\theta}_4^{\beta}(t)$. 此外, $\bar{\theta}_i^{\alpha}(t)=\theta_i^{\alpha}(t)$, $\bar{\theta}_i^{\beta}(t)=\theta_i^{\beta}(t)$, $i=2,\ 3,\ 5$. 如图 9 所示, 在式 (20) 和式 (21) 这两种不同的实现方式下, 每个节点传递给邻居的信息完全相同.

由图 10 ~ 12 可得

$$\bar{x}_1(0)=\frac{\bar{x}_1^{\alpha}(0)+\bar{\theta}_1^{\alpha}+\bar{m}_1(\bar{x}_1^{\beta}(0)+\bar{\theta}_1^{\beta})}{S}=\frac{2+1+2(2+2)}{4}=\frac{11}{4}$$

$$\bar{x}_4(0)=\frac{\bar{x}_4^{\alpha}(0)+\bar{\theta}_4^{\alpha}+\bar{m}_4(\bar{x}_4^{\beta}(0)+\bar{\theta}_4^{\beta})}{S}=\frac{6+(-9)+3(6+6)}{4}=\frac{33}{4}$$

$$\bar{x}_2(0)=x_2(0),\ \bar{x}_3(0)=x_3(0),\ \bar{x}_5(0)=x_5(0)$$

因此, 在实现式 (21) 下, 整个系统仍然实现准确的平均一致性.

在 $\bar{x}_1(0)=x_1(0)+\frac{7}{4}$ 和 $\bar{x}_4(0)=x_4(0)-\frac{7}{4}$ 这两种不同情形下, 完全相同的信息使得联合的好奇节点 2、3、5 无法判断 $\bar{x}_1(0)$, $\bar{x}_4(0)$ 和 $x_1(0)$, $x_4(0)$ 这两种情况中的哪一种情况更有可能. 这里虽然只给出一个例子, 但存在无穷多种与式 (21) 一样符合算法条件且初始状态差距任意大的情形可以使得联合的好奇节点 2、3、5 获得的信息与式 (20) 完全相同. 因此好奇的联合节点 2、3、5 既不能估计节点 1 和节点 4 的初始状态, 也不能估计节点 1 和节点 4 的初始状态所属的有界集合.

综上所述, 即使在强联合窃听下, 算法 1 仍然能够保护智能体的初始状态.

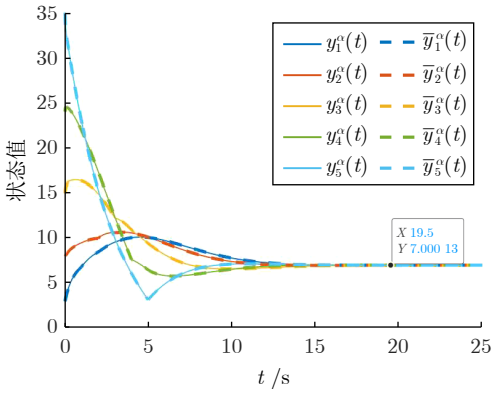

图 9　在实现方式 (20) 和 (21) 下, 智能体的输出轨迹 $y_i^{\alpha}(t)$ 和 $\bar{y}_i^{\alpha}(t)$

Fig. 9　The output trajectories of agents under realizations (20) and (21): $y_i^{\alpha}(t)$ and $\bar{y}_i^{\alpha}(t)$

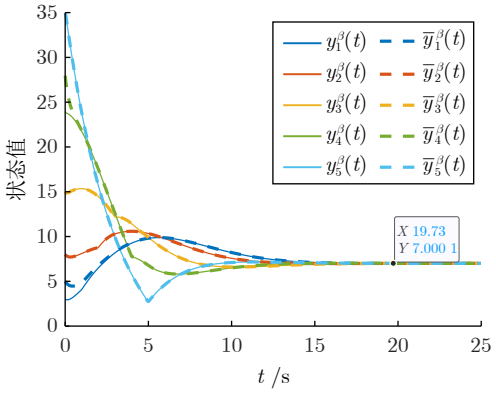

图 10　在实现方式 (20) 和 (21) 下, 虚拟节点的输出轨迹 $y_i^{\beta}(t)$ 和 $\bar{y}_i^{\beta}(t)$

Fig. 10　The output trajectories of virtual nodes under realizations (20) and (21): $y_i^{\beta}(t)$ and $\bar{y}_i^{\beta}(t)$

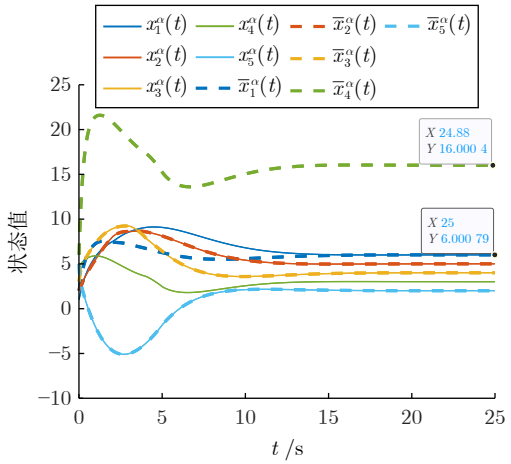

图 11　在实现方式 (20) 和 (21) 下, 智能体的实时状态轨迹 $x_i^{\alpha}(t)$ 和 $\bar{x}_i^{\alpha}(t)$

Fig. 11　The real-time state trajectories of agents under realizations (20) and (21): $x_i^{\alpha}(t)$ and $\bar{x}_i^{\alpha}(t)$

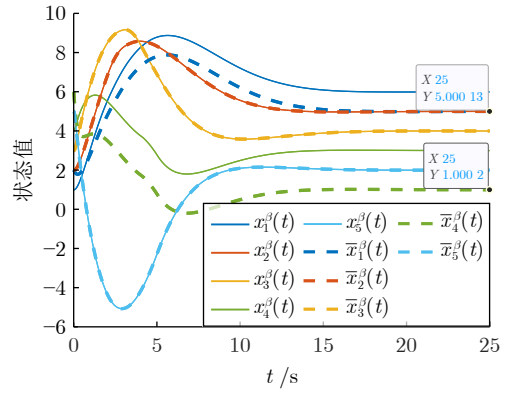

图 12　在实现方式 (20) 和 (21) 下, 虚拟节点的状态轨迹 $x_i^{\beta}(t)$ 和 $\bar{x}_i^{\beta}(t)$

Fig. 12　The state trajectories of virtual nodes under realizations (20) and (21): $x_i^{\beta}(t)$ and $\bar{x}_i^{\beta}(t)$

示例 3 (完全联合窃听下实时状态的隐私性验证). 下面给出智能体 1 不同于式 (20) 和式 (21) 的第三种情形.

$$\begin{cases}\bar{\bar{x}}_1^{\alpha}(0)=x_1^{\alpha}(0)+10\\ \bar{\bar{\theta}}_1^{\alpha}(t)=\theta_1^{\alpha}(t)-10\end{cases}\tag{22}$$

这里的 $\bar{\bar{x}}$ 表示是为与式 (20) 和式 (21) 中的表

示方式 $\bar{x}$ 和 x 进行区分, 其余节点的条件与式 (20) 相同.

不妨假定节点 2、3、4、5 为联合的好奇节点, 此时节点 1 的初始状态 $x_1(0)=1$ 和虚拟权重比 $m_1=1$ 泄露. 但从图 13 可知, 在 $x_1^{\alpha}(t)$ 和 $\bar{x}_1^{\alpha}(t)$ 这两种不同的实时状态轨迹下, 节点 1 的输出轨迹完全相同. 对联合好奇节点 2、3、4、5 而言, 节点 1 的实时状态 $x_1^{\alpha}(t)$ 仍然是不确定的, 并且 $x_1^{\alpha}(t)$ 由节点 1 自身任意决定. 综上, 即使在完全联合窃听下, 智能体的实时状态 $x_i^{\alpha}(t)$ 仍然是隐私的.

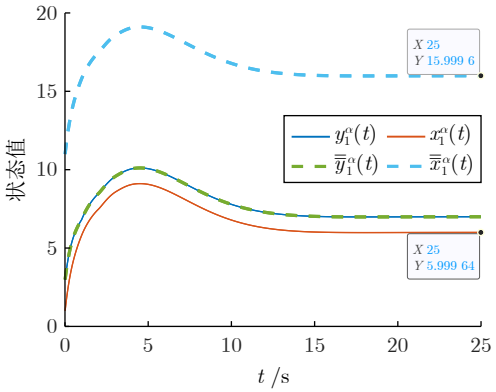

图 13 在实现方式 (20) 和 (22) 下, 节点 1 的实时状态轨迹 $x_i^{\alpha}(t)$ 和 $\bar{x}_i^{\alpha}(t)$

Fig. 13 The real-time state trajectories of node 1 under realizations (20) and (22): $x_i^{\alpha}(t)$ and $\bar{x}_i^{\alpha}(t)$

示例 4 (隐私性对比实验). 为说明本文所提出的全过程保护平均一致性算法的优越性, 与文献 [31] 提出的全过程隐私保护平均一致性算法在相同网络拓扑和相同初始条件下进行对比实验. 文献 [31] 中的平均一致性算法为

$$\dot{x}_i(t)=\sum_{j=1}^{N}a_{ij}(\alpha_j x_j(t)-\alpha_i x_i(t)) \tag{23}$$

其中, 比例系数 α_i 由第三方在满足 $\frac{1}{\alpha_1}+\frac{1}{\alpha_2}+\cdots+\frac{1}{\alpha_N}=N$ 和 $\alpha_i>0,\ \forall i\in\mathcal{V}$ 的条件下为每个智能体分配, 智能体 i 传递给邻居节点的信息为 $\alpha_i x_i(t)$. 不妨假定$\alpha_1=1/2,\ \alpha_2=2,\ \alpha_3=4,\ \alpha_4=8,\ \alpha_5=8/17$.节点 2 构建如下的两个观测器用以估计节点 1 的初始状态和实时状态:

$$Z_1(t)=\frac{y_2(t)-y_1(0)}{\int_0^t a_{ij}(y_2(\tau)-y_1(\tau))\mathrm{d}\tau}$$

$$Z_2(t)=\frac{y_1(t)}{\lim_{\tau\to\infty}Z_1(\tau)}$$

其中, $y_1(t)$ 和 $y_2(t)$ 分别表示节点 1 和节点 2 在 t 时刻的输出.

图 14 表明文献 [31] 中的算法可以实现准确的平均一致性, 观测器 $Z_1(t)$ 能够重构出节点 1 的隐私参数 α_1. 图 15 表明, 在文献 [31] 的方法下, 通过观测器 $Z_2(t)$ 可以重构节点 1 任意时刻的状态值 $x_1(t)$, 而在本文的方法 (式 (20)) 下则不能. 对比于全过程隐私保护平均一致性算法[31], 本文的全过程隐私保护平均一致性算法具备更好的隐私性.

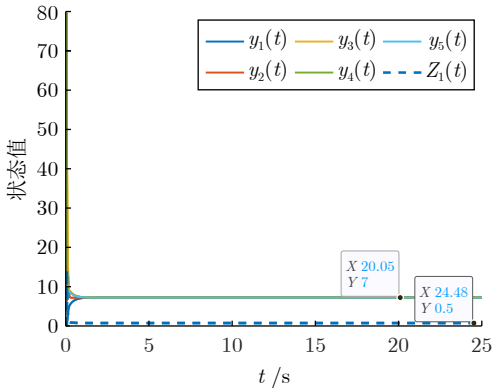

图 14 在文献 [31] 的方法下, 每个节点的输出轨迹 $y_i(t)$

Fig. 14 The output trajectory $y_i(t)$ of each node under the method of [31]

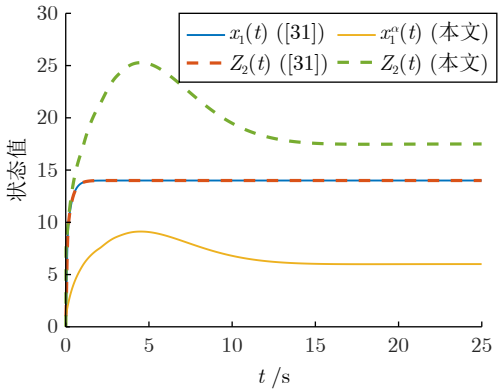

图 15 $Z_2(t)$ 在本文和文献 [31] 下的轨迹

Fig. 15 $Z_2(t)$ trajectories under the method of ours and [31]

4 结束语

针对联合窃听问题, 本文在强连通权重平衡网络下提出一种能够保护智能体的初始状态和实时状态的分布式平均一致性算法. 首先, 证明了该算法能够确保网络中的所有节点都能获得准确的平均一致性值; 其次, 从信息论的角度证明了只要整个网络中存在一个中立的节点, 则智能体的初始状态值无法被联合的好奇邻居估计. 此外, 即使在完全联合窃听下, 智能体的实时状态仍然可以得到保护. 令 $\hat{x}_i(0)$ 表示好奇节点对节点 i 初始状态值的估计

值. 在基于概率的差分隐私性定义下, 估计值和真实状态值的差位于有界集合 $|\hat{x}_i(0)-x_i(0)|<\kappa$ 中的概率随着 κ 的增大而增加, 其中 κ 为实数. 而在本文所给出的隐私性定义下, 联合的好奇节点既不能估计其他智能体的初始状态和实时状态, 也不能估计其他智能体初始状态和实时状态所属的有界集合, 即 $|\hat{x}_i(0)-x_i(0)|<\infty$. 最后, 通过 5 个节点的简单网络验证了所提算法的有效性.

然而, 本文所提出的平均一致性算法仍存在着不足:

1) 为确保每个节点在构建不同的虚拟子网络的同时, 仍然能够实现准确平均一致, 引入虚拟权重比参数 m_i, 这一参数依赖于第三方分配. 设计不需要第三方参与并能抵御联合窃听的分布式平均一致性算法将在未来的工作中进一步研究.

2) 网络中不仅存在窃听攻击这样的被动网络攻击, 也包含着重放攻击和欺骗攻击这样的主动网络攻击, 将在未来的工作中考虑混合网络攻击下的多智能体系统安全一致性.

References

1 Chung Y F, Kia S S. Dynamic active average consensus. *IEEE Control Systems Letters*, 2020, **5**(4): 1177−1182

2 Hassani H, Razavi-Far R, Saif M, Herrera-Viedma E. Consensus-based decision support model and fusion architecture for dynamic decision making. *Information Sciences*, 2022, **597**: 86−104

3 Du Y H, Hao T, Hui Y, Srdjan L. Accurate consensus-based distributed averaging with variable time delay in support of distributed secondary control algorithms. *IEEE Transactions on Smart Grid*, 2020, **11**(4): 2918−2928

4 Alsaadi F E, Wang Z D, Wang D, Alsaadi F E, Alsaade F W. Recursive fusion estimation for stochastic discrete time-varying complex networks under stochastic communication protocol: The state-saturated case. *Information Fusion*, 2020, **60**: 11−19

5 Zhang K, Li Z J, Wang Y Q, Louati A, Chen J. Privacy-preserving dynamic average consensus via state decomposition: Case study on multi-robot formation control. *Automatica*, 2022, **139**: Article No. 110182

6 Dong Y C, Zha Q B, Zhang H J, Kou G, Fujita H, Chiclana F, et al. Consensus reaching in social network group decision making: Research paradigms and challenges. *Knowledge-Based Systems*, 2018, **162**: 3−13

7 Yi D, Yu S Y. Rendezvous with connectivity preservation problem of linear multiagent systems via parallel event-triggered control strategies. *IEEE Transactions on Cybernetics*, 2020, **52**(5): 2725−2734

8 Maity D, Tsiotras P. Multiagent consensus subject to communication and privacy constraints. *IEEE Transactions on Control of Network Systems*, 2021, **9**(2): 943−955

9 Wang Z Q, Ma M L, Zhou Q, Xiong L Y, Wang L L, Wang J M, et al. A privacy-preserving distributed control strategy in islanded AC microgrids. *IEEE Transactions on Smart Grid*, 2022, **13**(5): 3369−3382

10 Yao A C. Protocols for secure computations. In: Proceedings of the 23rd Annual Symposium on Foundations of Computer Science. IEEE, 1982. 160−164

11 Shamir A. How to share a secret. *Communications of the ACM*, 1979, **22**(11): 612−613

12 Chaum D, Crépeau C, Damgard I. Multiparty unconditionally secure protocols. In: Proceedings of the 20th Annual ACM Symposium on Theory of Computing. Berlin, Germany: Springer, 1988. 11−19

13 Nozari E, Tallapragada P, Cortés J. Differentially private average consensus: Obstructions, trade-offs, and optimal algorithm design. *Automatica*, 2017, **81**: 221−231

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

15 Mo Y L, Murray R M. Privacy preserving average consensus. *IEEE Transactions on Automatic Control*, 2016, **62**(2): 753−765

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

17 Ramos G, Pequito S. Designing communication networks for discrete-time consensus for performance and privacy guarantees. *Systems Control Letters*, 2023, **180**: Article No. 105608

18 Gao L, Deng S J, Ren W. Differentially private consensus with an event-triggered mechanism. *IEEE Transactions on Control of Network Systems*, 2018, **6**(1): 60−71

19 Fiore D, Russo G. Resilient consensus for multi-agent systems subject to differential privacy requirements. *Automatica*, 2019, **106**: 18−26

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

21 Altafini C. A system-theoretic framework for privacy preservation in continuous-time multiagent dynamics. *Automatica*, 2020, **122**: Article No. 109253

22 Xiong Y, Li Z K. Privacy-preserved average consensus algorithms with edge-based additive perturbations. *Automatica*, 2022, **140**: Article No. 110223

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

24 Manitara N E, Hadjicostis C N. Privacy-preserving asymptotic average consensus. In: Proceedings of the European Control Conference. Zurich, Switzerland: IEEE, 2013. 760−765

25 Charalambous T, Manitara N E, Hadjicostis C N. Privacy-preserving average consensus over digraphs in the presence of time delays. In: Proceedings of the 57th Annual Allerton Conference on Communication, Control, and Computing. Monticello, USA: IEEE, 2019. 238−245

26 Rezazadeh N, Kia S S. A study of privacy preservation in average consensus algorithm via deterministic obfuscation signals. *IEEE Transactions on Control of Network Systems*, 2023, **11**(1): 534−546

27 Ye F, Cao X H, Chow M Y, Cai L. Privacy-preserving average consensus: Fundamental analysis and a generic framework design. *IEEE Transactions on Information Theory*, 2024, **70**(4): 2870−2885

28 Ying Chen-Duo, Wu Yi-Ming, Xu Ming, Zheng Ning, He Xiong-Xiong. Privacy-preserving average consensus control for multi-agent systems under deception attacks. *Acta Automatica Sinica*, 2023, **49**(2): 425−436
(应晨铎, 伍益明, 徐明, 郑宁, 何熊熊. 欺骗攻击下具备隐私保护的多智能体系统均值趋同控制. 自动化学报, 2023, **49**(2): 425−436)

29 Zhang J, Lu J, Liang J, Shi K. Privacy-preserving average consensus in multi-agent systems via partial information transmission. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2022, **53**(5): 2781−2791

30 Ramos G, Aguiarz A P, Karx S, Pequito S. Privacy preserving average consensus through network augmentation. *IEEE Transactions on Automatic Control*, 2024, **69**(10): 6907−6919

31 Zhang J, Lu J Q, Liang J L, Zhong J. Average consensus of whole-process privacy protection: A scale parameter method. *In-*

formation Fusion, 2024, **107**: Article No. 102312

32 Wang Z Q, Wang J, Scala M L, Xiong L Y. Real-time privacy-preserving average consensus and its application to secondary control for AC microgrid. *IEEE Transactions on Industrial Informatics*, 2024, **20**(7): 9655−9669

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

纪良浩 重庆邮电大学图像认知重庆市重点实验室教授. 2014 年获得重庆大学博士学位. 主要研究方向为多智能体系统和智能信息处理. 本文通信作者. E-mail: jilh@cqupt.edu.cn

(**JI Liang-Hao** Professor at Chongqing Key Laboratory of Image Cognition, Chongqing University of Posts and Telecommunications. He received his Ph.D. degree from Chongqing University in 2014. His research interest covers multi-agent systems and intelligent information processing. Corresponding author of this paper.)

唐少洪 重庆邮电大学图像认知重庆市重点实验室硕士研究生. 2022 年获得沈阳师范大学软件学院学士学位. 主要研究方向为多智能体系统的一致性控制和隐私保护.
E-mail: s220201088@stu.cqupt.edu.cn

(**TANG Shao-Hong** Master student at Chongqing Key Laboratory of Image Cognition, Chongqing University of Posts and Telecommunications. He received his bachelor degree from the School of Software, Shenyang Normal University in 2022. His research interest covers consensus control of multi-agent systems and privacy protection.)

郭 兴 重庆邮电大学图像认知重庆市重点实验室讲师. 2020 年获得东南大学博士学位. 主要研究方向为多智能体系统的分布式控制和网络安全控制. E-mail: guoxing@cqupt.edu.cn

(**GUO Xing** Lecturer at Chongqing Key Laboratory of Image Cognition, Chongqing University of Posts and Telecommunications. He received his Ph.D. degree from Southeast University in 2020. His research interest covers distributed control and networked secure control of multi-agent systems.)

解 燕 重庆邮电大学图像认知重庆市重点实验室博士研究生. 2022 年获得重庆科技大学硕士学位. 主要研究方向为多智能体系统弹性控制.
E-mail: d220201014@stu.cqupt.edu.cn

(**XIE Yan** Ph.D. candidate at Chongqing Key Laboratory of Image Cognition, Chongqing University of Posts and Telecommunications. She received her master degree from Chongqing University of Science and Technology in 2022. Her research interest covers the resilient control of multi-agent systems.)