Pure Mathematics 理论数学, **2025, 15(4), 338-350**
Published Online April 2025 in Hans. https://www.hanspub.org/journal/pm
https://doi.org/10.12677/pm.2025.154137

Hans汉斯

基于输出掩码的合作 竞争多智能体系统二分一致性隐私保护问题研究

陈永玲

上海理工大学理学院，上海

收稿日期：2025 年 2 月 19 日；录用日期：2025 年 4 月 15 日；发布日期：2025 年 4 月 29 日

摘 要

本文研究了连续时间合作 -竞争多智能体系统的隐私保护二分一致性问题。为了避免泄露网络节点的初始状态，同时实现具有合作 - 竞争的网络节点的二分一致性，本文提出了一种新的基于隐私保护二分一致性控制算法。本文所采用的隐私保护方法为构造一个输出掩码，使智能体的内部状态不被其他智能体察觉。这与现有的差分隐私以及同态加密的隐私保护方法不同，并且创新性地使用在合作 - 竞争多智能体系统中。基于所提出的隐私保护算法，本文对网络节点进行了详细的理论二分一致性和隐私保护性分析。最后，通过仿真实验验证了理论结果的有效性。

关键词

多智能体系统，合作 - 竞争网络，输出掩码，隐私保护，二分一致性

Privacy-Preserving Bipartite Consensus with Cooperative-Competitive Multi-Agent Interactions: An Output Mask Approach

Yongling Chen

College of Science, University of Shanghai for Science and Technology, Shanghai

文章引用: 陈永玲. 基于输出掩码的合作 -竞争多智能体系统二分一致性隐私保护问题研究 [J]. 理论数学, 2025, 15(4): 338-350. DOI: 10.12677/pm.2025.154137

Received: Feb. 19th, 2025; accepted: Apr. 15th, 2025; published: Apr. 29th, 2025

Abstract

This paper investigates the privacy-preserving bipartite consensus problem in continuous-time cooperative-competitive multi-agent systems. To prevent the leakage of initial states of network nodes while achieving bipartite consensus in networks with cooperative-competitive interactions, this paper proposes a novel privacy-preserving bipartite consensus control algorithm. This paper introduces an output mask mechanism to ensure the internal states of agents remain unobservable to other nodes. This method differs from existing privacy-preserving techniques such as differential privacy and homomorphic encryption, and this paper innovatively applies this method to cooperative-competitive multi-agent systems. Based on the proposed algorithm, a detailed theoretical analysis of bipartite consensus and privacy preservation is conducted. Finally, a numerical simulation is given to validate the effectiveness of the proposed privacy-preserving bipartite consensus algorithm.

Keywords

Multi-Agent Systems, Cooperative-Competitive Interactions, Output Mask, Privacy-Preserving, Bipartite Consensus

1. 引言

随着现代生活中科技不断发展与进步，控制理论的发展并与其它学科领域不断交叉融合，控制领域的专家学者基于群体行为的自然规律来设计并构造多智能体系统 [1] [2]，以期进一步来完成复杂的任务。近年来，多智能体系统的协调控制问题引起了各领域专家的广泛关注 [3]。主要原因是分布式协调控制 [4] 在执行结构性能较差的任务时，具有网络资源消耗少、运行速度快、容错能力强、可靠性高等优点。多智能体系统一致性问题作为多智能体系统合作与协调的基础，越来越多的文献对其进行了研究 [5] [6]。在对多智能体系统的研究中，学者们在研究过程中常常默认系统

中智能体之间仅存在“协作”关系，但实际存在连接权重为负值的“对抗”情况 [7]。在日常生活中随处可以看到事物之间存在竞争的情况。例如在生态学中经常可见到生物之间捕食和被捕食问题，同样在社会生活当中关于不同政党之间竞争问题。在文献 [8] 中证明非负网络上的标准一致性的一些性质对合作 - 竞争关系的网络也是有效的。文献 [9] 考虑到智能体之间的合作 - 竞争交互作用，设计了一种一致性协议。通过引入时间相关阈值函数，提出了一类不依赖任何全局信息的事件触发条件，最终实现合作 - 竞争多智能体系统的耦合群一致性。因此合作 - 竞争下多智能体系统二分一致性相关问题进行研究具有理论和实际价值。为了达到所有节点初始值的平均值，传统的平均一致性算法要求每个节点与其邻居交换和公开状态信息，智能体的真实初始值可以被其他智能体计算获得，这将导致敏感信息的泄露 [10]，这在隐私保护方面是不可取的。在实际生活中，类似多智能体隐私问题随处可见，例如个人住址，电话等身份信息的泄露等。我们需要在实现多智能体一致性的同时还保证了多智能体初始状态的隐私性。同样在合作 -竞争网络下多智能体隐私保护二分一致性相关问题的研究也越来越受到学者关注。例如文献 [11] 提出事件触发控制方法，并基于观测器计算出的估计状态和攻击信息，提出了分布式控制器以抵御攻击，实现了合作 -竞争多智能体系统的分布式安全控制，确保合作 - 竞争多智能体系统中的二分一致性。一般来说，隐私保护算法的关键是防止初始状态值泄露给他人，即敏感信息的匿名性 [12]。为了实现多智能体系统平均一致下的隐私保护，现有的研究大致有如下几种相关的方法。同态加密和差分隐私是目前应用最广泛的两种加密方法。由于差分隐私中的概率概念，所有的差分隐私方法都必须是随机设计的。一种普遍使用的差分私有机制是在迭代过程中给初始值附加一个随机偏移量来掩盖真实值 [13]- [15]。差分隐私机制可以有效地保证初始状态值的隐私性，但这种方法的一个主要缺点是只能达到随机意义上的平均一致性。此外，差分隐私方法必须在准确性和隐私性之间进行性能权衡。例如文献 [16]，从隐私保护下的信息共享量、隐私的控制理论成本以及隐私与性能之间的权衡这三个维度量化隐私的影响。同态加密技术 [17] [18] 支持对密文的直接操作，解密结果可以实现对明文操作的变化，这是不可缺少的优势。因此，同态加密已成为保护隐私的有效方法。现有的研究还采用了许多特殊技术来解决隐私保护的平均一致性问题。对于通过改变图的拓扑结构来实现隐私保护，在 [19] 中，在无向网络中，提出了一种基于状态分解的方法，可以保证所有参与节点在平均共识下的隐私性，而不影响准确性。本文旨在连续时间合作 -竞争多智能体系统上，基于输出掩码 [20] 提出一种新的实现二分一致性的隐私保护算法。因受到系统理论的启发，该方法关键在于将分布式计算问题（例如二分一致性问题）解释为动态系统，因此也叫做动态隐私。从技术上讲，全局吸引性可以用与标准一致性问题的相同李雅普诺夫函数来证明，如果李雅普诺夫函数的导数由渐近衰减到 0 的项所限定，则获得全局吸引性 [21]。本文中，我们定义输出映射为输出掩码，输出掩码在公开传输之前改变（或“屏蔽”）节点的内部状态 [22]。由于这些输出掩码对每个节点都是私有的，因此它们不允许对节点的状态进行任何形式的观察。我们的输出掩码是确定性的、时变的，对于每个节点都可以独立实现，并且渐近收敛到真实的内部状态。输出掩码的函数的输出值与真实值相互独立，由此产生的掩蔽系统也是一个时变系统，并且其极限系统对应于原始（未屏蔽）系统，以此起到隐私保护的作用 [23]。

注：定义 $\mathrm{diag}[x]$ 为主对角元素为 x 的对角矩阵,"$\|\cdot\|$ " 为'$\cdot$' 的欧几里德范数。$M^{\top}$ 代表矩阵 M 的转置。$\mathrm{sgn}(a)$ 定义一个符号函数，若 $a>0$，则 $\mathrm{sgn}(a)=1$，若 $a=0$, 则 $\mathrm{sgn}(a)=0$, 若 $a<0$, 则 $\mathrm{sgn}(a)=-1$. 令 $\mathcal{K}_{\infty}^{2}, \mathcal{L}^{2}$，和 $\mathcal{KL}_{\infty}^{1,e}$ 为特定的函数集。如果一个函数 α 以 0 为初始值严格单调递增，并且为 i 阶齐次多项式，则可以表示为 $\alpha(r)\in\mathcal{K}_{\infty}^{i}$，形式为 $\alpha(r)=ar^{i}, a>0$。如果一个函

数 ζ 严格指数递减，且收敛值到 0，那么它可以表示为 $\zeta(r) \in \mathcal{L}^e$，形式如下：$\zeta(t) = ae^{-\eta t}, a > 0$ 且 $\eta > 0$。如果函数 $\beta(r,t) \in \mathcal{K}_\infty^i$ 和 $\beta(r,t) \in \mathcal{L}^e$ 分别对应于每个固定的 t 和 r，那么它可以表示为 $\beta(r,t) \in \mathcal{KL}_\infty^{1,e}$，形式如下：$\beta(r,t) = ar^i e^{-\eta t}, a > 0$，$\eta > 0$.

2. 预备知识和模型

2.1. 图论

本文考虑一个具有 n 个节点的交互网络。$\mathcal{G} = \{\mathcal{V}, \mathcal{E}, A\}$ 表示一个符号图，其中 $\mathcal{V} = \{v_1, v_2, ..., v_n\}$ 表示由点组成的集合，$\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ 为边组成的集合 $A = [a_{ij}]_{n\times n}$ 代表图 $\mathcal{G}$ 的邻接矩阵。若 $(v_j, v_i) \in \mathcal{E}$，则 $a_{ij} = a_{ji} \neq 0$，否则 $a_{ij} = 0$. 若 $a_{ij} > 0$，则表示节点 v_i 和节点 v_j 之间的权重值为正，表示节点 v_i 和节点 v_j 之间为合作关系。若 $a_{ij} < 0$，表示节点 v_i 和节点 v_j 之间的权重值为负，表示节点 v_i 和节点 v_j 之间为竞争关系。D 表示图 $\mathcal{G}$ 的度矩阵。$N_i = \{v_j \in \mathcal{V} \mid (v_i, v_j) \in \mathcal{E}\}$ 表示节点 v_i 的邻居节点组成的集合。

2.2. 输出掩码

本文的目的是防止在二分一致性计算过程中节点的真实初始状态被其他节点识别。根据对每个节点初始状态隐私保护的要求，本文引入一个函数作为输出掩码来掩饰内部状态 $x_i(t)$. 输出掩码的定义如下：定义 1 [24]：称函数 $h_i(t, x_i, \pi_i)$ 是节点 i 的一个具有隐私保护性质的掩码函数。如果节点 i 满足以下条件：

(1):$h_i(0, x_i, \pi_i) \neq x_i \quad \forall x_i \in R^n, i = 1, 2, \cdots, N$;

(2):$h_i(t, x_i, \pi_i)$ 满足初始条件的不可分辨性；

(3): 对于 $\forall \pi_i \in R^n$，$h_i(t, x_i, \pi_i)$ 不保留邻域；

(4): 对于每个固定的 t 和 $\pi_i, i = 1, 2, \cdots, N$，$h_i(t, x_i, \pi_i)$ 对于 $x_i(t)$ 严格递增；

(5): 对于每个固定的 $x_i(t)$ 和 $\pi_i, i = 1, 2, \cdots, N$，$|h_i(t, x_i, \pi_i) - x_i|$ 对于 t 递增，并且 $\lim_{t\to\infty} h_i(t, x_i, \pi_i) = x_i, i = 1, 2, \cdots, n$。定义 1 是输出掩码的定义，满足网络节点初始状态隐私保护的要求。根据输出掩码的定义，现为每个节点构造如下的输出掩码：

$$y_i = h_i(t, x_i, \pi_i)$$
$$h_i(t, x_i, \pi_i) = (1 + \phi_i e^{-\sigma_i t})(x_i(t) + \gamma_i e^{-\delta_i t})$$

其中 y_i 为节点 i 的输出掩码 $h_i(t, x_i, \pi_i)$ 的输出状态，$\phi_i, \sigma_i, \delta_i$ 为标量，$\gamma_i \in R^n$ 与 $x_i, i = 1, 2, \cdots, n$ 具有相同的维数。掩码系统的通信拓扑应满足以下假设：

$$\{N_i \cup \{i\}\} \not\subset \{N_j \cup \{j\}\}, \quad \forall i, j = 1, 2 \cdots, N, \quad i \neq j \tag{1}$$

即每个节点与其邻居节点的交互有独立性，不会相互影响。给出以下引理 1，阐述选择以上函数作为本文的掩码函数的理由引理 1：函数 $h_i(t, x_i, \pi_i) = (1 + \phi_i e^{-\sigma_i t})(x_i(t) + \gamma_i e^{-\delta_i t})$ 是一个屏蔽内部状态 $x_i(t)$ 的隐私输出掩码。

证明. 由于 $h_i(0, x_i, \pi_i) = (1 + \phi_i)(x_i(t) + \gamma_i) \neq x_i$，则定义 *1* 中条件 *(1)* 满足。在定义 *1* 中条件

(1) 中，$\dot{x}_i(t)$ 和函数 $h_i(t,x_i,\pi_i)$ 的输出状态 $y_i(t)$ 无法确定参数 $\pi_i=\{\phi_i,\sigma_i,\delta_i,\gamma_i\}$，因此不能绘制 $h_i(t,x_i,\pi_i)$ 的曲线，从而保证节点初始值的隐私。则 $h_i(t,x_i,\pi_i)$ 满足定义 *1* 中条件 *(2)*。若 $x^*\in R^n,|x(0)-x^*|<\xi$，则

$$\begin{aligned}&|h_i(0,x,\pi)-x^*|\\&=|(I+\phi)(x+\gamma)-x^*|\\&\leq|(I+\phi)x-x^*|+|(I+\phi)\gamma|\end{aligned}$$

其中 $\phi=\mathrm{diag}(\phi_1,\cdots,\phi_n)$, $\gamma=[\gamma_1,\cdots,\gamma_n]^T$。此不等式不属于 x^* 的 ξ 邻域。则满足定义 *1* 中条件 *(3)*。

由于 $[\partial h_i(t,x_i,\pi_i)/\partial x_i]=(1+\phi_i e^{-\sigma_i t})>0$，则对于每个固定的 t 和 π_i，函数 $h_i(t,x_i,\pi_i)$ 随 x_i 严格递增。则满足定义 *1* 中条件 *(4)*。

由于 $|h_i(t,x_i,\pi_i)-x_i|=\gamma_i e^{-\delta_i t}+\phi_i e^{-\delta_i t}x_i+\phi_i\gamma_i e^{(-\sigma_i+\delta_i)t)}$ 则对于每个固定的 x_i 和 π_i，函数 $h_i(t,x_i,\pi_i)$ 随 t 递增。并且有 $\lim_{t\to\infty}h_i(t,x_i,\pi_i)=x_i$ 则满足定义 *1* 中条件 *(5)*。
综上函数 $h_i(t,x_i,\pi_i)$ 满足定义 *1* 中的 *5* 个条件，则可用于屏蔽节点内部状态。

2.3. 算法描述

考虑一个包含 n 个节点的连续时间合作 -竞争多智能体系统如下：

$$\begin{aligned}&\dot{x}_i(t)=c_iu_i(t)\\&u_i(t)=-\sum_{j\in N_i}|a_{ij}|(x_i(t)-\mathrm{sgn}(a_{ij})x_j(t))\end{aligned}\tag{2}$$

其中 $x_i\in R^n$ 是状态向量，$u_i\in R^n$ 是节点 i 的控制输入，c_i 是节点 i 的控制器增益。(2) 式的增广形式可以写为：

$$\dot{X}(t)=-CLX(t)$$

其中 L 为拉普拉斯矩阵，有 $L=D-A$. $C=\mathrm{diag}[c_1\cdots c_n]$，$D=\mathrm{diag}[d_1\cdots d_n]$，$A=[a_{ij}]_{n\times n}$ 为图 $\mathcal{G}$ 的邻接矩阵，$X(t)=[x_1\cdots x_n]^\top$。在传统的二分一致性算法中，节点的初始值可以由诚实但好奇的节点和窃听者计算，这可能导致节点隐私的泄露。为了保护节点的初始值，提出了基于输出掩码的合作 -竞争多智能体系统隐私保护方案。选取掩码函数

$$\begin{aligned}&y_i=h_i(t,x_i,\pi_i),\\&h_i(t,x_i,\pi_i)=(1+\phi_i e^{-\sigma_i t})(x_i(t)+\gamma_i e^{-\delta_i t}).\end{aligned}\tag{3}$$

则 (3) 式增广形式为：

$$Y(t)=(I+\Phi e^{-\Sigma t})(X(t)+e^{-\Delta t}\gamma)\tag{4}$$

其中 $I\in R^n,\Phi=\mathrm{diag}[\phi_1\cdots\phi_n],\Sigma=\mathrm{diag}[\sigma_1\cdots\sigma_n],X(t)=[x_1\cdots x_n]^\top,\Delta=\mathrm{diag}[\delta_1\cdots\delta_n]$, $\gamma=[\gamma_1\cdots\gamma_n]^\top$. 代入掩码函数，(2) 式可写为：

$$\begin{aligned}\dot{x}_i(t) &= c_i u_i(t),\\ u_i(t) &= -\sum_{j\in N_i} |a_{ij}|(y_i(t) - \mathrm{sgn}(a_{ij})y_j(t))\end{aligned} \tag{5}$$

其增广形式为:

$$\dot{X}(t) = -CL(I + \Phi e^{-\Sigma t})(X + e^{-\Delta t}\gamma)$$

简写为:

$$\dot{X}(t) = -CLNX - CLNM(t) \tag{6}$$

其中：$N(t) = I + \Phi e^{-\Sigma t}, M(t) = e^{-\Delta t}\gamma$

接下来证明此算法使得原系统 (2) 式在实现二分一致性的同时，可以保护节点初始值。

3. 二分一致性以及隐私性分析

为了方便后续定理的提出以及证明，首先给出以下定义及引理。引理 2 [25]：多智能体系统的渐近一致性问题可以转化为跟踪误差系统的渐近稳定性问题。引理 3 [22]：对于以下系统

$$\dot{v} = -\alpha(v) + \beta(v,t) + \zeta(t) \tag{7}$$

如果 $\alpha(v) \in \mathcal{K}_\infty^2, \beta \in \mathcal{KL}_\infty^{1,e}, \zeta \in \mathcal{L}^2$，那么这些解都可以推广到 ∞，并且对于 $\forall v_0 \geq 0$ 和 $\forall t_0 \geq 0$ 有界。此外 $\mathrm{limt} \to \infty \mathrm{v(t)} = 0, \forall \mathrm{v_0} \geq 0$ 和 $\forall t_0 \geq 0$ 引理 4 [26]：对于以下系统 $(R_+ \times R^n \to R^n)$:

$$\dot{x} = g(t,x) \tag{8}$$

如果存在一个连续可微函数 $V(t,x), \alpha_1, \alpha_2, \alpha_3 \in \mathcal{K}_\infty^2, \beta \in \mathcal{KL}_\infty^{1,e}, \zeta \in \mathcal{L}^2$，满足:

$$\alpha_1(\|X\|) \leq V(t,x) \leq \alpha_2(\|X\|),$$

$$\frac{\partial V}{\partial t} + \frac{\partial V}{\partial x} g(t,x) \leq \alpha_3(\|X\|) + \beta(\|X\|, t - t_0) + \zeta(t - t_0).$$

其中 $\forall t \geq t_0, t_0 \geq 0$ 且 $x_0 \in R^n$。那么当 $t \to \infty$ 时，上式的任意解在 t_0 处收敛到 0. 定义 2 [8]：系统2所对应的无向符号图 $\mathcal{G}$ 是结构平衡的，当且仅当所有节点可以分为正、负两部分，分别定义为 $\mathcal{V}_1$ 和 $\mathcal{V}_2$, 并且满足 $\mathcal{V}_1 \cap \mathcal{V}_2 = \emptyset, \mathcal{V}_1 \cup \mathcal{V}_2 = \mathcal{V}$。引理 5：存在给定对角矩阵 $Q = \mathrm{diag}_n\{\delta_i\}$. 常数 $\delta_i \in \{\pm 1\}$. δ_i，L 为图 $\mathcal{G}$ 拉普拉斯矩阵. A 为图 $\mathcal{G}$ 邻接矩阵，D 为度矩阵。当图 $\mathcal{G}$ 为结构平衡图时，存在对角矩阵 Q ，使得 QLQ 非负。

证明.

$$\begin{aligned} Z(t) &= QX(t), Q^{-1} = Q \\ Q^{-1}z(t) &= Q^{-1}QX(t) = X(t) \\ QZ(t) &= X(t) \\ \dot{X}(t) &= -CLx(t) = -CLQZ(t) \\ Q\dot{X}(t) &= -QCLQZ(t) \\ \dot{Z}(t) &= -QCLQZ(t). \end{aligned}$$

由于控制器增益 C 和矩阵 Q 都为对角矩阵，则 $QC = CQ$，所以

$$\begin{aligned} \dot{Z}(t) &= -CQLQZ(t) \\ &= -CL_QZ(t). \end{aligned}$$

此处 $L_Q = QLQ, L_Q$ 为非负，因此可以把二分一致性问题转换为平均一致性问题。

定理 1：对于连续时间合作 -竞争多智能体系统系统 (2)，考虑每个节点插入掩码函数 (3)，且存在一个上述对角矩阵 Q，使得 $L_Q = QLQ$, 则系统 (6) 可实现二分一致性。

证明. 给定如下李亚普洛夫函数：

$$V(t) = \frac{1}{2}(X - X_0)^\top e^{-It}(X - X_0). \tag{9}$$

那么 *(9)* 式沿着 *(6)* 式轨迹的时间导数可以按以下方式获得：

$$\begin{aligned} \dot{V}(t) =& \frac{\partial V}{\partial X}\dot{X} + \frac{\partial V}{\partial t} \\ =& (X - X_0)^\top e^{-It}\dot{X} - \frac{1}{2}(X - X_0)^\top I_{4n}e^{-It}(X - X_0) \\ =& (X - X_0)^\top e^{-It}(-CLNX - CLNM(t)) - \frac{1}{2}(X - X_0)^\top 1_{4n}e^{-I_{4n}t}(X - X_0) \\ =& X^\top e^{-It}(-CLNX - CLNM(t)) - X_0^\top e^{-It}(-CLNX - CLNM(t)) \\ & - \frac{1}{2}X^\top Ie^{-It}(X - X_0) + \frac{1}{2}x_0^\top I_{4n}e^{-It}(X - X_0). \\ =& - X^\top e^{-I_{4n}t}CLNX - X^\top e^{-I_{4n}t}CLNM(t) + X_0^\top e^{-I_{4n}t}CLNX + X_0^\top e^{-It}CLNM(t) \\ & - \frac{1}{2}X^\top Ie^{-I_{4n}t}X + \frac{1}{2}X^\top Ie^{-It}X_0 + \frac{1}{2}x_0^\top Ie^{-I_{4n}t}X - \frac{1}{2}x_0^\top Ie^{-It}X_0. \\ =& - X^\top e^{-It}CLNX - X^\top e^{-It}CLNM(t) + X_0^\top e^{-It}CLNX + X_0^\top e^{-It}CLNM(t) \\ & - \frac{1}{2}X^\top Ie^{-I_{4n}t}X + X_0^\top Ie^{-It}X - \frac{1}{2}x_0^\top Ie^{-It}X_0. \\ =& - X^\top e^{-It}CLNX - \frac{1}{2}X^\top Ie^{-It}X \qquad (10a) \\ & - X^\top e^{-It}CLNM(t) + X_0^\top e^{-It}CLNX + X_0^\top I_{4n}e^{-I_{4n}t}X \qquad (10b) \\ & + X_0^\top e^{-I_{4n}t}CLNM(t) - \frac{1}{2}X_0^\top Ie^{-It}X_0. \qquad (10c) \end{aligned}$$

对于 *(10a)* 式

$$\begin{aligned} a =& - X^{\top} e^{-I_{4n}t} CLNX - \frac{1}{2} X^{\top} I_{4n} e^{-I_{4n}t} X \\ =& X^{\top} e^{-I_{4n}t} (-CLN - \frac{1}{2} I) X. \end{aligned}$$

由于 $N = I + \Phi e^{-\Sigma t}, M(t) = e^{-\Delta t}$, 可以得到

$$a \leq X^{\top} e^{-I_{4n}t} (-CL - \frac{1}{2} I) X.$$

$as\ -CL - \frac{1}{2} I \leq -\frac{1}{2} I \leq 0$，那么 $a \leq 0$
对于 *(10b)* 式

$$\begin{aligned} b =& - X^{\top} e^{-I_{4n}t} CLNM(t)\gamma + X_0^{\top} e^{-I_{4n}t} CLNX + X_0^{\top} I_{4n} e^{-I_{4n}t} X \\ \leq& \|X\| \|CL\| \|N(0)\gamma\| \max e^{-\sigma_i t} e^{-t} + \|X_0^{\top}\| \|CL\| N(0) \|X\| e^{-t} + \|X_0^{\top}\| \|X\| e^{-t} \\ =& \beta_1(\|z\|, t) + \beta_2(\|z\|, t) + \beta_3(\|z\|, t). \end{aligned}$$

其中 $\beta_i(\|X\|, t) \in \mathcal{KL}_{\infty}^{1,e}, \quad i = 1, 2, 3$
对于 *(10c)* 式

$$\begin{aligned} c =& X_0^{\top} e^{-I_{4n}t} CLNM(t)\gamma - \frac{1}{2} x_0^{\top} I_{4n} e^{-I_{4n}t} X_0 \\ \leq& \|X_0\| \|CL\| \|N(0)\gamma\| \max e^{-\sigma_i t} e^{-t} + \frac{1}{2} \|X_0^{\top} X_0\| e^{-t} \\ =& \zeta_1(t) + \zeta_2(t). \end{aligned}$$

其中 $\{\zeta_1(t), \zeta_2(t)\} \in \mathcal{L}^e$ 综上可得，

$$\dot{V} \leq -\alpha(\|X\|) + \beta(\|X\|, t) + \zeta(t). \tag{11}$$

其中 $\beta(\|X\|, t), \zeta(t)$ 分别代指 $\beta_i(\|X\|, t)$ 和 $\zeta_j(t), \quad i = 1, 2, 3; \quad j = 1, 2.$ 根据引理 *2* 和引理 *3* 结论，$V(0) = V(0, X(0))$ 作为初始值，则可以得出定理 *1* 的结论。

定理 2：对于连续时间合作 -竞争多智能体系统系统 (2)，每个节点插入掩码函数 (3)，在所提出的状态分解隐私保护方法下，节点 v_i 的初始值的隐私可以在与邻居的信息交换过程中得到保护。

证明. 如上所述，输出掩码方法存在一个严格的假设，即系统应满足条件 *(1)*. 根据本文所提出输出掩码的隐私保护方法，对于每个节点 v_i 不会直接与 v_i 的其他任何邻居连接。因此可以满足掩码函数需满足的条件 *(1)*，并且节点之间未被掩盖的界限并不涉及隐私泄露问题，因此使用本文所提出的隐私保护算法，能保护系统隐私。

注意：在隐私性分析中，不需要文献 [22] 中提出的输出掩码方法对通信拓扑结构的约束条件。并且现有文献（如 [27]）对通信拓扑结构也有类似的严格要求。因此与现有许多方法相比，本文所提出的方法操作更为简便适用

4. 数值仿真

本节给出了一个数值模拟示例，以证明隐私保护机制可以防止隐私泄露。为了方便且不失一般性，考虑 6 个节点的拓扑结构图，以 (图 1) 为例。令 6 个节点的初始值 $X(0)$ 为 $[2,-4,1,0,-2,5]$，邻接矩阵如下：

$$A=\begin{pmatrix} 0 & 1 & 0 & -1 & 0 & 0 \\ 1 & 0 & 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ -1 & 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & -1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}. \tag{12}$$

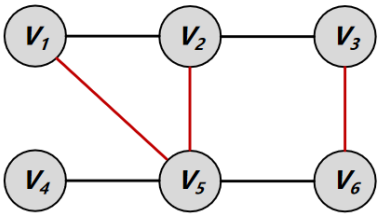

Figure 1. Network topology diagram with 6 nodes

图 1. 6 个节点的网络拓扑图

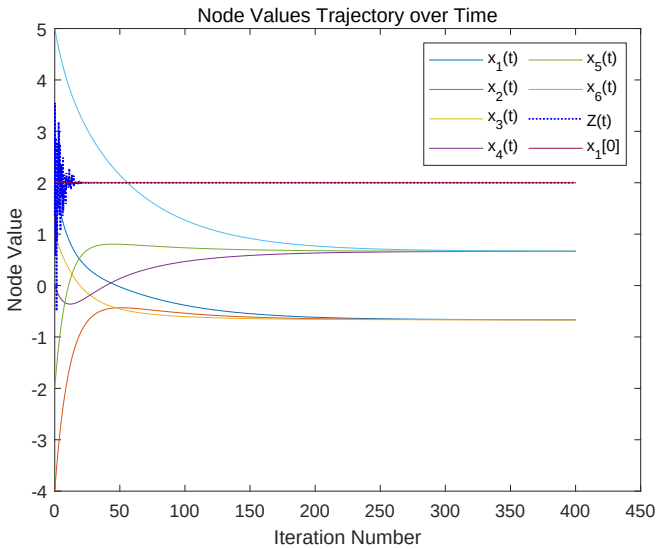

Figure 2. The red horizontal solid line represents the initial value of v_1. The blue dashed line represents the state of external eavesdroppers. Eavesdropper z can infer the initial value $x_1[0]$ of node v_1

图 2. 红色的实线表示节点 v_1 的初始值，蓝色虚线表示外部窃听者随时间变化的状态。窃听者可以观测到节点 v_1 的初始值 $x_1[0]$

受文献 [19] 启发，考虑类似的攻击者如下：

$$\dot{z}(t) = \dot{x}_i^+(t) - (-c_i \sum_{j \in N_i} |a_{ij}(t)|(x_i^+(t) - \mathrm{sgn}(a_{ij}(t))x_j^+(t))). \tag{13}$$

其中 $x_i^+(t) = x_i(t) + w_i(t)$. $w_i(t)$ 为每个节点在其状态中添加的随机噪声：

$$w_i(t) = \begin{cases} v_i(t) \quad \text{if} \quad t = 0 \\ s^t v_i(t) - s^{t-dt} v_i(t - dt) \quad \text{if} \quad t \neq 0. \end{cases} \tag{14}$$

其中 s 为 $(0, 1)$ 之间的常数。每个节点在 t 时刻生成一个随机变量 $v_i(t)$，$v_i(t)$ 是均值为 0，方差为 1 的标准正态分布。本文假设所有随机变量 $v_i(t) i = 1, \cdots, n$ 是相互独立的。以图 (1) 为例，在应用隐私保护机制之前，外部窃听者希望获得节点 v_1 的初始值。外部窃听着设置为

$$\begin{aligned} \dot{z}(t) =& \dot{x}_1^+(t) - (-c_1|a_{12}(t)|(x_1^+(t) - \mathrm{sgn}(a_{12}(t))x_2^+(t)) \\ & - c_1|a_{14}(t)|(x_1^+(t) - \mathrm{sgn}(a_{14}(t))x_4^+(t))). \end{aligned} \tag{15}$$

其中 $s = 0.8$.

使用隐私保护机制以后，加入如下掩码函数：

$$y_i = (1 + \phi_i e^{-\sigma_i t})(x_i(t) + \gamma_i e^{-\delta_i t}).$$

其中取 $c = 0.02, \phi = 0.5, \sigma = 0.4, \gamma = 0.3, \delta = 0.7$.

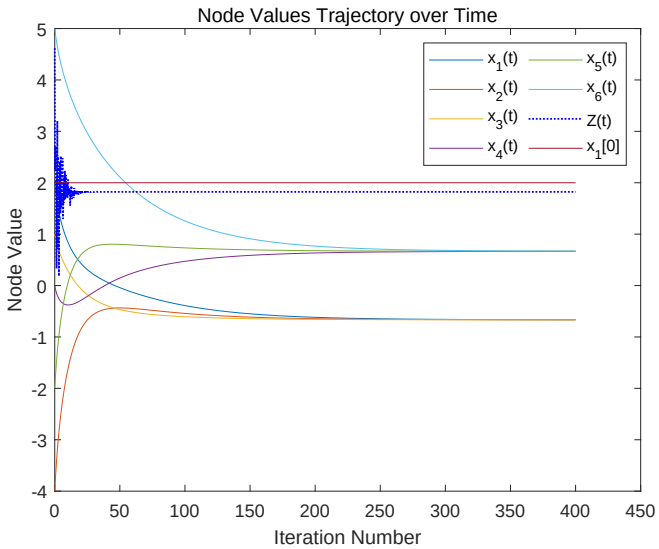

Figure 3. The red horizontal solid line represents the initial value of v_1. The blue dashed line represents the state of external eavesdroppers. Eavesdropper z cann ot infer the initial value of node v_1.

图 3. 红色的实线表示节点 v_1 的初始值，蓝色虚线表示外部窃听者随时间变化的状态，窃听者无法观测到节点 v_1 的初始值

使得 $x_i^+(t) = y_i(t) + w_i(t)$. 相应的仿真结果如图 2~3所示。从图 2 可以看出，虽然这 6 个节点在隐私保护机制之前可以实现二分一致性，但节点 v_1 的初始值可以被窃听者检测到, 不能实现隐私保护。从图 3可以看出，通过本文设计的隐私保护机制，加入掩码函数，在实现二分一致性的同时，相同的窃听者无法计算节点 v_1 的初始值，可以实现隐私保护。本文使用的攻击者参考文献 [19]，文献 [19] 仿真实验中对比可得此攻击者在文献 [28] 中无法保护系统隐私，而在本文输出掩码隐私保护机制下实现隐私保护。

5. 总结

本文研究了合作 -竞争多智能体系统中基于输出掩码的隐私保护二分一致性问题。在无向符号图中，考虑连续时间系统，设计了一种基于输出掩码的隐私保护机制来保护节点隐私，同时确保系统达到二分一致性。首先本文通过对合作 -竞争多智能体系统进行正规变换，把原系统求解二分一致性解问题转变为求解平均一致性解问题。接着对于每个节点输出输出掩码，再根据李亚普洛夫函数，把平均一致性问题转换为全局吸引行问题。再对李亚普洛夫函数求导得到一致性的证明。再经过分析得到隐私性的证明。最后，数值模拟验证了所提出的隐私保护算法的有效性。

参考文献

[1] Mi, W., Luo, L. and Zhong, S. (2023) Fixed-Time Consensus Tracking for Multi-Agent Systems with a Nonholomonic Dynamics. *IEEE Transactions on Automatic Control*, **68**, 1161-1168. https://doi.org/10.1109/tac.2022.3148312

[2] Dai, J., Yi, J. and Chai, L. (2024) Accelerating the Convergence Rate of Consensus for Second-Order Multi-Agent Systems by Memory Information. *Automatica*, **166**, Article 111727. https://doi.org/10.1016/j.automatica.2024.111727

[3] Zhang, W., Mao, S., Huang, J., Kocarev, L. and Tang, Y. (2021) Data-driven Resilient Control for Linear Discrete-Time Multi-Agent Networks under Unconfined Cyber-Attacks. *IEEE Transactions on Circuits and Systems I: Regular Papers*, **68**, 776-785. https://doi.org/10.1109/tcsi.2020.3037242

[4] Tang, Z., Hill, D.J. and Liu, T. (2021) Distributed Coordinated Reactive Power Control for Voltage Regulation in Distribution Networks. *IEEE Transactions on Smart Grid*, **12**, 312-323. https://doi.org/10.1109/tsg.2020.3018633

[5] Jiang, Y., Liu, L. and Feng, G. (2024) Fully Distributed Adaptive Control for Output Consensus of Uncertain Discrete-Time Linear Multi-Agent Systems. *Automatica*, **162**, Article 111531. https://doi.org/10.1016/j.automatica.2024.111531

[6] Olfati-Saber, R. and Murray, R.M. (2004) Consensus Problems in Networks of Agents with Switching Topology and Time-Delays. *IEEE Transactions on Automatic Control*, **49**, 1520-1533. https://doi.org/10.1109/tac.2004.834113

[7] Zhai, S. and Zheng, W.X. (2019) On Survival of All Agents in a Network with Cooperative and Competitive Interactions. *IEEE Transactions on Automatic Control*, **64**, 3853-3860. https://doi.org/10.1109/tac.2019.2892521

[8] Altafini, C. (2013) Consensus Problems on Networks with Antagonistic Interactions. *IEEE Transactions on Automatic Control*, **58**, 935-946. https://doi.org/10.1109/tac.2012.2224251

[9] Li, K., Ji, L., Yang, S., Li, H. and Liao, X. (2022) Couple-Group Consensus of Cooperative-Competitive Heterogeneous Multiagent Systems: A Fully Distributed Event-Triggered and Pinning Control Method. *IEEE Transactions on Cybernetics*, **52**, 4907-4915. https://doi.org/10.1109/tcyb.2020.3024551

[10] Ruan, M., Gao, H. and Wang, Y. (2019) Secure and Privacy-Preserving Consensus. *IEEE Transactions on Automatic Control*, **64**, 4035-4049. https://doi.org/10.1109/tac.2019.2890887

[11] Shi, S., Wang, Z., Xiao, M., Jiang, G. and Cao, J. (2024) Consensus Analysis for Cooperative-Competitive Multiagent Systems under False Data Injection Attacks via Dynamic Event-Triggered Observers. *IEEE Transactions on Signal and Information Processing over Networks*, **10**, 195-204. https://doi.org/10.1109/tsipn.2024.3375611

[12] Zhan, J., Hsieh, C.-L., Wang, I-C., Hsu, T.-S., Liau, C.-J. and Wang, D.-W. (2010) Privacy-Preserving Collaborative Recommender Systems. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, **40**, 472-476. https://doi.org/10.1109/tsmcc.2010.2040275

[13] Gao, C., Wang, Z., He, X., Liu, Y. and Yue, D. (2024) Differentially Private Consensus Control for Discrete-Time Multiagent Systems: Encoding-Decoding Schemes. *IEEE Transactions on Automatic Control*, **69**, 5554-5561. https://doi.org/10.1109/tac.2024.3367803

[14] Wang, J., Ke, J. and Zhang, J. (2024) Differentially Private Bipartite Consensus over Signed Networks with Time-Varying Noises. *IEEE Transactions on Automatic Control*, **69**, 5788-5803. https://doi.org/10.1109/tac.2024.3351869

[15] Wang, J. and Zhang, J. (2024) Differentially Private Distributed Stochastic Optimization with Time-Varying Sample Sizes. *IEEE Transactions on Automatic Control*, **69**, 6341-6348. https://doi.org/10.1109/tac.2024.3379387

[16] Yazdani, K., Jones, A., Leahy, K. and Hale, M. (2023) Differentially Private LQ Control. *IEEE Transactions on Automatic Control*, **68**, 1061-1068. https://doi.org/10.1109/tac.2022.3148710

[17] Chen, W., Liu, L. and Liu, G. (2023) Privacy-Preserving Distributed Economic Dispatch of Microgrids: A Dynamic Quantization-Based Consensus Scheme with Homomorphic Encryption. *IEEE Transactions on Smart Grid*, **14**, 701-713. https://doi.org/10.1109/tsg.2022.3189665

[18] Gao, H., Zhang, C., Ahmad, M. and Wang, Y. (2018) Privacy-Preserving Average Consensus on Directed Graphs Using Push-Sum. *2018 IEEE Conference on Communications and Network Security (CNS)*, Beijing, 30 May-1 June 2018, 1-9. https://doi.org/10.1109/cns.2018.8433217

[19] Wang, Y. (2019) Privacy-Preserving Average Consensus via State Decomposition. *IEEE Transactions on Automatic Control*, **64**, 4711-4716. https://doi.org/10.1109/tac.2019.2902731

[20] Hu, J., Sun, Q., Wang, R. and Wang, Y. (2024) An Improved Privacy-Preserving Consensus Strategy for AC Microgrids Based on Output Mask Approach and Node Decomposition Mechanism. *IEEE Transactions on Automation Science and Engineering*, **21**, 642-651. https://doi.org/10.1109/tase.2022.3217677

[21] Liu, Y., Xie, X., Sun, J. and Yang, D. (2024) Event-Triggered Privacy Preservation Consensus Control and Containment Control for Nonlinear Mass: An Output Mask Approach. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, **54**, 4437-4447. https://doi.org/10.1109/tsmc.2024.3379375

[22] Altafini, C. (2020) A System-Theoretic Framework for Privacy Preservation in Continuous-Time Multiagent Dynamics. *Automatica*, **122**, Article 109253. https://doi.org/10.1016/j.automatica.2020.109253

[23] Artstein, Z. (1977) The Limiting Equations of Nonautonomous Ordinary Differential Equations. *Journal of Differential Equations*, **25**, 184-202. https://doi.org/10.1016/0022-0396(77)90199-1

[24] Wang, A., He, H. and Liao, X. (2021) Event-Triggered Privacy-Preserving Average Consensus for Multiagent Networks with Time Delay: An Output Mask Approach. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, **51**, 4520-4531. https://doi.org/10.1109/tsmc.2019.2939680

[25] Qin, J., Zhang, G., Zheng, W.X. and Kang, Y. (2019) Adaptive Sliding Mode Consensus Tracking for Second-Order Nonlinear Multiagent Systems with Actuator Faults. *IEEE Transactions on Cybernetics*, **49**, 1605-1615. https://doi.org/10.1109/tcyb.2018.2805167

[26] Hu, J., Sun, Q., Wang, R., Wang, B., Zhai, M. and Zhang, H. (2022) Privacy-Preserving Sliding Mode Control for Voltage Restoration of AC Microgrids Based on Output Mask Approach. *IEEE Transactions on Industrial Informatics*, **18**, 6818-6827. https://doi.org/10.1109/tii.2022.3141428

[27] Wang, Y., Lu, J., Zheng, W.X. and Shi, K. (2021) Privacy-Preserving Consensus for Multi-Agent Systems via Node Decomposition Strategy. *IEEE Transactions on Circuits and Systems I: Regular Papers*, **68**, 3474-3484. https://doi.org/10.1109/tcsi.2021.3081372

[28] Manitara, N.E. and Hadjicostis, C.N. (2013) Privacy-Preserving Asymptotic Average Consensus. *2013 European Control Conference (ECC)*, Zurich, 17-19 July 2013, 760-765. https://doi.org/10.23919/ecc.2013.6669251