DoS 攻击下具备隐私保护的多智能体系统均值趋同控制

胡沁伶 [1] 郑 宁 [1] 徐 明 [1] 伍益明 [1] 何熊熊 [2]

摘 要 均值趋同是一种广泛应用于分布式计算和控制的算法, 旨在系统通过相邻节点间信息交互、更新, 最终促使系统中所有节点以它们初始值的均值达成一致. 研究拒绝服务 (Denial-of-service, DoS) 攻击下的分布式离散时间多智能体系统均值趋同问题. 首先, 给出一种基于状态分解思想的分布式网络节点状态信息处理机制, 可保证系统中所有节点输出值的隐私. 然后, 利用分解后的节点状态值及分析给出的网络通信拓扑条件, 提出一种适用于无向通信拓扑的多智能体系统均值趋同控制方法. 理论分析表明, 该方法能够有效抵御 DoS 攻击的影响, 且实现系统输出值均值趋同. 最后, 通过仿真实例验证了该方法的有效性.

关键词 多智能体系统, 均值趋同, 拒绝服务攻击, 隐私保护, 网络安全

引用格式 胡沁伶, 郑宁, 徐明, 伍益明, 何熊熊. DoS 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, 2022, **48**(8): 1961−1971

DOI 10.16383/j.aas.c201019

Privacy-preserving Average Consensus Control for Multi-agent Systems Under DoS Attacks

HU Qin-Ling[1] ZHENG Ning[1] XU Ming[1] WU Yi-Ming[1] HE Xiong-Xiong[2]

Abstract Average consensus is a widely used algorithm for distributed computing and control, where all the nodes in the network constantly communicate and update their states in order to achieve an agreement. In this paper, we study the average consensus problem for discrete-time multi-agent systems under DoS attacks. First, a distributed network node state value processing mechanism based on state decomposition is given, which can ensure the privacy of the output values of all nodes in the system. Then, through using the decomposed node state values and the network topology conditions given by the analysis, an average output consensus control law for distributed discrete-time multi-agent systems is proposed. Theoretical analysis shows that the proposed method can effectively resist the influence of DoS attacks on the system, and achieve the convergence of the average value of system initial outputs. Finally, numerical examples are presented to show the validity of the proposed method.

Key words Multi-agent systems, average consensus, denial-of-service attack, privacy-preserving, cyber security

Citation Hu Qin-Ling, Zheng Ning, Xu Ming, Wu Yi-Ming, He Xiong-Xiong. Privacy-preserving average consensus control for multi-agent systems under DoS attacks. *Acta Automatica Sinica*, 2022, **48**(8): 1961−1971

多智能体系统是由多个具有一定传感、计算、执行和通信能力的智能个体组成的网络系统, 作为分布式人工智能的重要分支, 已成为解决大型、复杂、分布式及难预测问题的重要手段[1−2]. 趋同问题作为多智能体系统分布式协调控制领域中一个最基本的研究课题, 是指在没有协调中心的情况下, 系统中每个节点仅根据相互间传递的信息, 将智能体动力学与网络通信拓扑耦合成复杂网络, 并设计合适的分布式控制方法, 从而在有限时间内实现所有节点状态值的一致或同步.

然而具备分布式网络特点的多智能体系统由于普遍规模庞大, 单个节点结构简单且节点地理位置分散等原因, 使得系统中易产生脆弱点, 这就使其在推广应用中面临两项基本挑战: 1) 节点状态信息的隐私泄露问题; 2) 节点或节点间的通信链路可能会遭受网络攻击的问题, 如欺骗攻击、拒绝服务 (Denial-of-service, DoS) 攻击等.

针对节点状态信息的隐私泄露问题, 即在考虑多智能体网络趋同的同时, 保证系统中节点的初始状态值不被泄露, 已有较多研究人员开展相关的工作. 其中, 有学者借助于传统的安全多方计算方法,

收稿日期 2020-12-09 录用日期 2021-03-02
Manuscript received December 9, 2020; accepted March 2, 2021
国家自然科学基金 (61803135, 61873239, 62073109) 和浙江省公益技术应用研究项目 (LGF21F020011) 资助
Supported by National Natural Science Foundation of China (61803135, 61873239, 62073109) and Zhejiang Provincial Public Welfare Research Project of China (LGF21F020011)
本文责任编委 鲁仁全
Recommended by Associate Editor LU Ren-Quan
1. 杭州电子科技大学网络空间安全学院 杭州 310018 2. 浙江工业大学信息工程学院 杭州 310023
1. School of Cyberspace, Hangzhou Dianzi University, Hangzhou 310018 2. College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023

例如 Yao 等[3] 提出混淆电路算法, Shamir 等[4] 提出秘钥共享算法等. 然而这类通用的隐私保护方法因计算和通信消耗较大, 不适用于单个智能体节点结构较为简单的分布式系统, 尤其是受到硬实时约束的一类多智能体系统应用. 如上述的混淆电路的计算延迟为秒级[5], 而对于多智能体系统一些典型应用如多无人飞行器编队的实时控制, 其容许的计算延迟仅为毫秒级[6]. 针对多智能体系统均值趋同过程中节点信息泄露问题, 有研究人员提出了一系列专门的隐私保护策略[7−10]. 这些方法大多基于模糊处理的思想, 即通过加入噪声来掩盖真实的状态值. 其中一种常用的手段是差分隐私方法[11], 然而这种差分隐私下的模糊处理方法会影响最终趋同值的精度, 即使系统无法收敛到精确的节点初始状态的平均值. 最近文献 [12] 提出的一种基于相关噪声混淆技术的改进方法, 克服了传统差分隐私方法中精度下降的问题, 但却需要较多的算力. 最近的文献 [13] 采用一种基于状态分解的方法, 将每个节点的初始状态分解为两个随机的子状态, 只令其中一个子状态参与相邻节点间的信息交互, 而另一子状态保留在本节点内部, 不参与邻居间信息传递. 只要两个随机子状态的和满足特定条件, 在作者所设计的趋同算法下, 系统能够达成均值趋同, 且保护每个节点的状态信息不被泄露.

此外, 有学者研究基于可观测性的方法用来保护多智能体系统中节点的隐私[14−16]. 基本思想是设计网络的交互拓扑结构以最小化某个节点的观测性, 本质上相当于最小化该节点推断网络中其他节点初始状态的能力. 然而, 这类基于可观测性的方法仍然存在隐私泄露的风险. 为了提高对隐私攻击的抵御能力, 另一种常见的方法是使用加密技术. 然而, 虽然基于密码学的方法可以很容易地在聚合器或第三方[17] 的帮助下实现隐私保护, 例如基于云的控制或运算[18−20], 但是由于分散密钥管理的困难, 在没有聚合器或第三方的情况下, 将基于密码学的方法应用到完全分散的均值趋同问题是很困难的. 同时, 基于密码学的方法也将显著增加通信和计算开销[21], 往往不适用于资源有限或受硬实时约束的分布式网络控制系统.

以上的工作均是在安全的通信环境下完成的, 然而在实际应用场景中, 由于物理设备和通信拓扑结构都有可能遭受网络攻击, 导致以往有关多智能体系统趋同研究的失效, 这使得针对多智能体系统在网络攻击下的趋同研究发展迅速, 并取得了一些显著成果[22−26]. 目前多智能体系统中常见的网络攻击主要有两种形式: 欺骗攻击[22, 25, 27−28] 和 DoS 攻击[29−33]. 其中 DoS 攻击是多智能体系统中最常见也是最容易实现的攻击形式, 只要攻击者掌握系统元器件之间的通信协议, 即可利用攻击设备开展干扰、阻塞通信信道、用数据淹没网络等方式启动 DoS 攻击. 在 DoS 攻击影响下, 智能体间交互的状态信息因传递受阻而致使系统无法达成一致. 近年来, 研究者们从控制理论的角度对 DoS 攻击下的系统趋同问题进行了研究. 其中, 有研究人员通过构建依赖于参数的通用 Lyapunov 函数设计一种趋同方法[31], 使其能够适用于因通信链路存在随机攻击导致通信拓扑随机切换的情况. 此外, 有研究者通过设计一个独立于全局信息的可靠分布式事件触发器[32], 很好地解决了大规模 DoS 攻击下的一致性问题. 更有研究者开始研究异构多智能体系统在通信链路遭受攻击时的趋同问题[33], 通过设计基于观测器的控制器, 实现在通信链路存在 DoS 攻击时两层节点间的趋同问题. 而在本文中, 考虑多智能体之间通信链路遭受 DoS 攻击的情况, 通过攻击开始时刻与攻击链路矩阵刻画 DoS 攻击模型, 通过增强网络拓扑以满足所谓的 r-鲁棒图来刻画信息流的局部冗余量[34], 从而抵御 DoS 攻击的影响.

然而, 针对趋同问题, 将网络攻击和隐私保护两者结合起来考虑的研究还鲜有见文献报道. 2019 年 Fiore 等[24] 率先开展了同时考虑隐私保护和网络攻击的研究工作, 但所得成果仍存在一定的局限性: 1) 所提方法虽能保护节点隐私且最终达成状态值趋同, 却无法确保系统达成均值趋同; 2) 作者仅考虑了欺骗攻击下的控制器设计问题, 因此所得结论并不适用于网络中存有 DoS 攻击的系统.

基于上述观察与分析, 本文主要致力于研究 DoS 攻击下具备节点信息隐私保护的多智能体系统均值趋同问题, 从而补充现有趋同算法的相关结果. 同时, 考虑实际环境对测量条件等的限制, 不易直接获取节点的真实状态值[35], 为此本文围绕节点的输出值, 即通过观测矩阵获取的系统输出 y, 进行趋同控制器的设计工作. 本文的主要贡献包括:

1) 针对 DoS 攻击在多智能体系统分布式协同控制中的攻击特性和发生范围, 及对网络拓扑连通性的影响, 建立相应数学模型;

2) 针对一类 DoS 攻击下的无向通信网络多智能体系统, 提出一种基于状态分解的节点信息隐私保护策略. 当满足特定条件时, 所提策略可确保系统输出状态不被窃听者准确推断出来;

3) 针对 DoS 攻击的影响, 分析给出了系统中节点通信拓扑的鲁棒性条件, 并据此设计一种基于输出量测值 y 的分布式控制方法, 理论分析并证明

系统可容忍特定数目的链路遭受 DoS 破坏, 并实现输出均值趋同.

本文内容结构为: 第 1 节介绍本文所需要用到的图论知识, 网络拓扑图的相关性质以及均值趋同算法; 第 2 节主要对 DoS 攻击模型和拟解决问题进行描述; 第 3 节提出系统在 DoS 攻击下的隐私保护均值趋同控制方法, 并分别对在攻击下的网络拓扑鲁棒性、系统收敛性以及隐私保护能力进行分析; 第 4 节通过一组仿真实例验证算法的有效性; 第 5 节是总结与展望.

1　预备知识

1.1　图论知识

考虑由 M 个智能体组成的多智能体系统, 节点之间为双向传递信息, 其通信网络可抽象地用一个无向加权图 $\mathcal{G}=(\mathcal{V}, \mathcal{E}, A)$表示. 其中 $\mathcal{V}=\{v_1, v_2, \cdots, v_M\}$ 表示节点集合, $\mathcal{E}\subset\mathcal{V}\times\mathcal{V}$ 表示边集. 两个节点之间的连接关系用邻接矩阵 (权重矩阵) $\boldsymbol{A}=[a_{ij}]\in\mathbf{R}^{M\times M}$ 表示, 如果 $(v_j, v_i)\in\mathcal{E}$, 则 $a_{ij}>0$; 否则 $a_{ij}=0$. 在无向图中, 邻接矩阵是对称的, 即如果 $(v_j, v_i)\in\mathcal{E}$, 则同时有 $(v_i, v_j)\in\mathcal{E}$, 且 $a_{ij}=a_{ji}$. 本文不考虑节点自环情况, 即令 $a_{ii}=0$. 节点 v_i 的邻居集合表示为 $\mathcal{N}_i=\{v_j\in\mathcal{V}|(v_j, v_i)\in\mathcal{E}\}$. 无向图 $\mathcal{G}$ 对应的 Laplacian 矩阵为 $\boldsymbol{L}=\boldsymbol{D}-A$, 其中 $\boldsymbol{D}$ 为度矩阵, 定义为:

$$\boldsymbol{D}=\mathrm{diag}\left\{\sum_{j=1}^{M}a_{1j}, \sum_{j=1}^{M}a_{2j}, \cdots, \sum_{j=1}^{M}a_{Mj}\right\}$$

除了上述无向图的基本知识, 本文的研究工作还用到了 r-可达集合和 r-鲁棒图的概念. 这两个概念最早由文献 [36] 提出, 随后被文献 [22, 27] 等利用并扩展, 主要用于分析节点间拓扑抵御网络攻击的鲁棒性. 经笔者少许修改, 具体定义如下:

定义 1[36]**.** r-可达集合: 对于图 $\mathcal{G}=(\mathcal{V}, \mathcal{E})$ 及其中一非空子集 $\mathcal{S}\subset\mathcal{V}$, 如果 $\mathcal{S}$ 中至少有一个节点 v_i 在集合 $\mathcal{N}_i\setminus\mathcal{S}$ 中有不少于 r 个节点, 则称 $\mathcal{S}$ 为 r-可达集合.

定义 2[36]**.** r-鲁棒图: 对于图 $\mathcal{G}=(\mathcal{V}, \mathcal{E})$, 如果对 $\mathcal{V}$ 中任意一对非空子集 $\mathcal{S}_1, \mathcal{S}_2\subset\mathcal{V}$, $\mathcal{S}_1\cap\mathcal{S}_2=\emptyset$, 保证至少有一个子集为 r-可达集合, 则称 $\mathcal{G}$ 为 r-鲁棒图.

以下是一些关于 r-鲁棒图的基本性质.

引理 1[22]**.** 考虑一个 r-鲁棒图 $\mathcal{G}=(\mathcal{V}, \mathcal{E})$, 令 $\hat{\mathcal{G}}$ 表示 $\mathcal{G}$ 中每个节点至多移除 s $(s<r)$ 条边后的图, 则 $\hat{\mathcal{G}}$ 是一个 $(r-s)$-鲁棒图.

引理 2[22]**.** 对于一个无向图 $\mathcal{G}$, 如果 $\mathcal{G}$ 满足 1-鲁棒图, 则有 $\mathcal{G}$ 为连通图.

1.2　均值趋同算法

考虑有 M 个节点组成的无向加权多智能体系统. 为了让系统实现均值趋同, 也就是所有节点的状态 $x_i[k]$ 最终收敛到它们初始状态的平均值 $\sum_{i=1}^{M}x_i[0]/M$, 根据文献 [13, 37], 其节点动态更新方程可设计为:

$$x_i[k+1]=x_i[k]+\varepsilon\sum_{v_j\in\mathcal{N}_i}a_{ij}(x_j[k]-x_i[k]) \tag{1}$$

式中, $x_i[k]$为节点 v_i 在 k 时刻的状态值, $\varepsilon\in(0, 1/\Delta)$ 为系统增益系数, Δ通常定义为:

$$\Delta:=\max_{i=1, 2, \cdots, M}|\mathcal{N}_i| \tag{2}$$

文献 [38] 表明, 当系统拓扑满足连通图, 且存在 $\eta>0$ 使得 $\eta\le a_{ij}<1$ 时, 系统可在更新规则 (1) 下实现均值趋同, 即:

$$\lim_{k\to+\infty}x_i[k]=\frac{\sum\limits_{i=1}^{M}x_i[0]}{M},\quad\forall v_i\in\mathcal{V} \tag{3}$$

2　问题描述

本文研究对象为如下 M 个智能个体组成的一阶离散时间多智能体系统, 其动力学模型为:

$$\begin{cases}x_i[k+1]=x_i[k]+u_i\\ y_i[k]=nC_ix_i[k],\quad v_i=1, 2, \cdots, M\end{cases} \tag{4}$$

式中, $x_i[k]\in\mathbf{R}^N$ 为系统的状态值, u_i 为控制输入, $y_i[k]\in\mathbf{R}^Q$ 为系统经通信链路传输得到的量测信号, 需要注意的是, 由于通信链路中存在 DoS 攻击, $y_i[k]$ 可能遭受影响而无法被邻居节点接收到. $nC_i\in\mathbf{R}^{Q\times N}$ 为观测矩阵, 其中 n 为从观测矩阵中抽取出的系数, $n\in\mathbf{R}_+$ 为大于 0 的正实数.

2.1　攻击模型

本文所讨论的 DoS 攻击表现为某种传输尝试失败的情况[39], 其存在于多智能体系统中各智能体之间的通信链路中, 即当通信图中两个节点间的链路发生 DoS 攻击时, 其通信链路将会被切断, 此时两个节点无法通过该链路进行信息交互, 进而达到攻击多智能体系统的目的. 在多智能体系统分布式协同控制中, 运载节点输出量测值的通信链路遭遇 DoS 攻击的示意图如图 1 所示.

本文以 Adeversory (P, k_0) 刻画系统遭遇 DoS 攻击的情况. 其中 $P=[p_{ij}[k]]\in\mathbf{R}^{M\times M}$ 表示攻击状态矩阵, 当节点 v_i 和节点 v_j 之间在 k 时刻发生 DoS

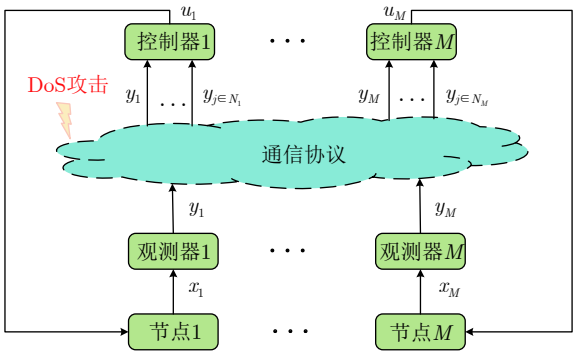


图 1 DoS 攻击下的多智能系统框图

Fig. 1 The diagram of the multi-agent system under DoS attacks

攻击时, $p_{ij}[k]=0$; 否则 $p_{ij}[k]=1$. k_0 为系统遭遇 DoS 攻击的开始时刻.

考虑攻击者资源的有限, 本文假设攻击发生范围满足 f-本地有界[22] 的定义, 该假设在文献 [22–23, 25] 中被广泛采用. 结合 DoS 攻击, 具体定义如下:

定义 3 (f-本地有界 DoS 攻击) **.** 对于系统中的任一节点, 如果与其相邻节点的通信链路中, 任意时刻遭遇 DoS 攻击的链路条数至多不超过 f 条, 则称此类攻击模型为 f-本地有界 DoS 攻击.

2.2 系统假设

结合上述给出的 Adeversory (P, k_0) 和攻击发生范围模型, 本文对所研究的系统作出如下假设:

假设 1. 系统中任意一个节点的通信链路中在任意时刻至多有 f 条链路同时遭受 DoS 攻击, 即满足定义 3 攻击模型. 具体地, 则对于任意 $v_i\in\mathcal{V}$, 在任意时刻 k, 都有下式成立:

$$M-\sum_{j=1}^{M}p_{ij}[k]\leq f$$

虽然本文考虑的是固定无向拓扑, 但在 DoS 攻击影响下, 可以看到系统的通信图却会与之发生变化. 因此, 本文接下去用时变图符号 $\mathcal{G}[k]=(\mathcal{V},\mathcal{E}[k],\mathrm{A}[k])$ 表示系统在 DoS 攻击影响下的真实通信情况.

假设 2. 存在一个标量 η 满足 $0<\eta<1$, 对于所有的 $i,j\in\{1,\cdots,M\}$, 如果 $a_{ij}[k]>0$, 那么 $\eta\leq a_{ij}[k]<1$.

假设 3. 系统任意节点状态值 $x_i\in\mathbf{R}^N$ 受限于一个非空闭凸集, 表示为 $\mathcal{X}_i\in\mathbf{R}^N$, 令 $\mathcal{X}=\cap_{i=1}^{M}\mathcal{X}_i$, 则 $\mathcal{X}\neq\emptyset$.

根据上述假设, 可以得出系统具备如下属性:

引理 3[38]**.** 当系统的网络通信图为有向连通图且邻接矩阵为双随机矩阵时, 并且满足假设 2 和 3 时, 那么对于系统中任意节点 $v_i\in\mathcal{V}$ 在动态更新式 (1) 下, 有:

$$\lim_{k\to\infty}\|x_i[k]-h[k]\|=0 \tag{5}$$

$$\lim_{k\to\infty}\|w_i[k]-h[k]\|=0 \tag{6}$$

式中, $\{h[k]\}$ 为一个定义的辅助序列, 对于每个时刻 k, 有:

$$h[k]=\frac{1}{M}\sum_{i=1}^{M}w_i[k] \tag{7}$$

$$w_i[k]=\sum_{j=1}^{M}a_{ij}[k]x_j[k] \tag{8}$$

根据文献 [38], 因邻接矩阵为双随机矩阵, 由式 (7) ~ (8) 可得:

$$h[k]=\frac{1}{M}\sum_{i=1}^{M}\Big(\sum_{j=1}^{M}a_{ij}[k]x_j[k]\Big)=\frac{1}{M}\sum_{j=1}^{M}x_j[k] \tag{9}$$

将式 (9) 代入式 (5), 即:

$$\lim_{k\to\infty}x_i[k]=\frac{1}{M}\sum_{j=1}^{M}x_j[k] \tag{10}$$

引理 4. 当系统的网络通信图为无向连通图, 并且满足假设 2 和 3 时, 那么由引理 3 可知, 对于系统中任意节点 $v_i\in\mathcal{V}$ 在动态更新式 (1) 下, 式 (10) 仍然成立.

证明. 根据引理 3 可知, 在网络通信图为有向图情况下, 邻接矩阵为双随机矩阵表明在该网络通信图中, 所有节点通信链路满足出度等于入度的条件, 而在无向图中, 该条件同样成立, 因此在无向图中, 式 (10) 仍然成立. □

针对上述建立的网络攻击模型和相关的系统假设, 本文的研究目标是, 设计一种控制策略, 使得: 1) 系统的输出达到趋同并且趋同值是等于所有智能体初始输出状态的平均值; 2) 在整个趋同过程中保护每个节点的信息值隐私.

3 控制器设计

3.1 DoS 攻击下网络拓扑鲁棒性条件

首先对网络通信链路图的鲁棒性条件进行讨论, 以便于开展后续控制器的设计工作.

引理 5. 考虑多智能体系统 (4), 如果其网络拓

扑结构满足 $(f+1)$-鲁棒的无向图, 那么系统在遭受 f-本地有界 DoS 攻击下, 即满足假设 1, 其通信图仍可保持连通性.

证明. 根据假设 1 可知, 网络中每个节点任意时刻至多有 f 条通信链路遭受 DoS 攻击破坏. 再由引理 1 可知, 此时网络拓扑结构至少是 1-鲁棒图. 最后由引理 2 可知, 系统网络拓扑仍然能够保持连通性. □

3.2　DoS 攻击下隐私保护控制

上述小节给出了系统遭受 DoS 攻击下通信网络仍旧保持连通的条件, 接下去本小节给出本文核心的控制器设计方法.

受文献 [13] 启发, 此处引入状态分解方法: 将每个节点的状态值 x_i 分解成两个子状态, 用 x_i^α 和 x_i^β 表示. 值得注意的是, 初始状态的子状态值 $x_i^\alpha[0]$ 和 $x_i^\beta[0]$ 可在所有实数中任取, 但需满足条件: $x_i^\alpha[0]+x_i^\beta[0]=2x_i[0]$.

为便于理解, 本文以 5 个节点的无向连通图为例, 通信拓扑如图 2 所示. 从示例图中可以看出: 子状态 x_i^α 充当原 x_i 的作用, 即与邻居节点进行信息交互, 并且实际上是节点 v_i 的邻居节点唯一可以获知的状态信息. 而另一个子状态 x_i^β 同样存在于该分布式信息交互中, 但是其仅与 x_i^α 进行信息交互. 也就是说子状态 x_i^β 的存在, 对于节点 v_i 的邻居节点是不可见的. 例如, 在图 2(b) 中, 节点 v_1 中的 x_1^α 相当于图 2(a) 中 x_1 的角色和邻居节点进行信息交互, 而 x_1^β 仅对节点 v_1 自身可见, 而对其他节点不可见. 但是它又可以影响 x_1^α 的变化. 两个子状态 x_i^α 和 x_i^β 之间的耦合权重是对称的, 表示为 $a_{i,\,\alpha\beta}[k]$, 并且所有的 $a_{i,\,\alpha\beta}[k]$ 满足 $\eta \le a_{i,\,\alpha\beta}[k] < 1$.

基于上述方法, 本文给出具体的具备隐私保护的输出均值趋同控制协议:

$$\begin{cases} x^\alpha[k+1] = P_\varepsilon[k]y[k] + \varepsilon A_{\alpha\beta}[k](x^\alpha[k]-x^\beta[k]) \\ x^\beta[k+1] = x^\beta[k] + \varepsilon A_{\alpha\beta}[k](x^\beta[k]-x^\alpha[k]) \\ y[k] = nCx^\alpha[k] \end{cases} \tag{11}$$

定义:

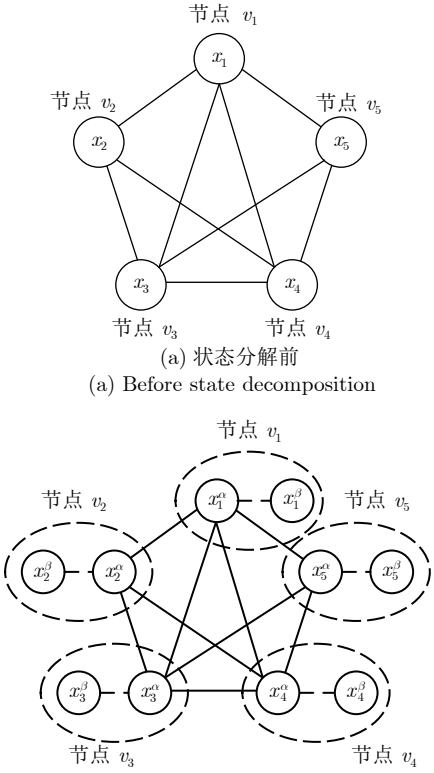
$$x^\alpha[k] := \left[x_1^\alpha[k], \cdots, x_M^\alpha[k]\right]^{\mathrm{T}} \in \mathbf{R}^M \tag{12}$$

$$x^\beta[k] := \left[x_1^\beta[k], \cdots, x_M^\beta[k]\right]^{\mathrm{T}} \in \mathbf{R}^M \tag{13}$$

并且

$$P_\varepsilon[k] := \boldsymbol{I} - \varepsilon L'[k] \tag{14}$$

式中, $\boldsymbol{I}$ 为单位矩阵, $\boldsymbol{L}'[k]$ 为 DoS 攻击下的 Laplacian 矩阵, 其满足:



(a) 状态分解前
(a) Before state decomposition

(b) 状态分解后
(b) After state decomposition

图 2　5 个节点组成的示例图
Fig. 2　Example of network with 5 nodes

$$\boldsymbol{L}'[k] = \boldsymbol{D}'[k] - \boldsymbol{A}'[k] \tag{15}$$

式中, DoS 攻击下系统对应的邻接矩阵为 $\boldsymbol{A}'[k] = [a'_{ij}[k]]$, 满足:

$$a'_{ij}[k] = \begin{cases} a_{ij}[k], & k < k_0 \\ a_{ij}[k]p_{ij}[k], & k \ge k_0 \end{cases} \tag{16}$$

$\boldsymbol{D}'[k]$ 为对应于邻接矩阵 $A'[k]$ 的度矩阵.

另外, 在协议 (11) 中, $y[k] = nCx^\alpha[k]$ 为系统的状态输出方程, $\boldsymbol{C}$ 为输出方程的观测矩阵, 定义为:

$$\boldsymbol{C} := [\boldsymbol{e}_1,\ \cdots,\ \boldsymbol{e}_{M-1},\ \boldsymbol{e}_M]^{\mathrm{T}} \in \mathbf{R}^{M\times M} \tag{17}$$

式中, $\boldsymbol{e}_i$ 表示 $\mathbf{R}^M$ 中第 i 个规范基向量, 该向量中第 i 个位置数为 1, 其他位置数为 0.

注 1. 考虑实际环境中不同情况, 当 $n \in (0,\,1)$ 时, 系统输出方程将会缩小状态值进行信息交互, 适用于节点状态值过大的情况; 当 $n=1$ 时, 系统状态输出方程将会输出原本节点需要进行信息交互的状态值; 当 $n \in (1,\,\infty)$ 时, 系统状态输出方程将会放大状态值进行信息交互, 适用于节点状态值过小的情况.

值得注意的是, 对于系统中的节点, 用于和邻居节点进行信息交互的状态值 $x^\alpha[k]$ 是无法被邻居节点获取的, 需通过系统状态输出方程传递给邻居

节点. 简言之每个节点经过信息交互接收到的邻居节点的值并不是$x^\alpha[k]$, 而是经过输出方程输出的$y[k]$.

令$\boldsymbol{A}_{\alpha\beta}[k]$为每个节点$v_i$, $i=1,2,\cdots,M$的$x_i^\alpha[k]$和$x_i^\beta[k]$两个子状态之间的耦合权重$a_{i,\alpha\beta}[k]$对应的矩阵, 定义为

$$\boldsymbol{A}_{\alpha\beta}[k]:=[a_{1,\alpha\beta}[k],\cdots,a_{M,\alpha\beta}[k]]\in\mathbf{R}^M \tag{18}$$

为便于叙述, 本文考虑节点的状态值及输出值为一维的情况, 即令$N=1$, $Q=1$. 从而, 基于输出状态值的控制协议可表示为:

$$\begin{cases} x_i^\alpha[k+1]=x_i^\alpha[k]+\varepsilon\displaystyle\sum_{j=1}^{M}a'_{ij}[k](y_j[k]-y_i[k])+\\ \qquad\qquad \varepsilon a_{i,\alpha\beta}[k](x_i^\beta[k]-x_i^\alpha[k])\\ x_i^\beta[k+1]=x_i^\beta[k]+\varepsilon a_{i,\alpha\beta}[k](x_i^\alpha[k]-x_i^\beta[k])\\ y_i[k]=nx_i^\alpha[k] \end{cases} \tag{19}$$

事实上, 只要向量状态中的每个标量状态元素都有独立的耦合权重, 本节所提出的控制方法所有分析及结果同样适用于向量状态的情况.

注 2. 与文献 [13, 37] 的更新式 (1) 相比, 本文给出的协议 (19) 中, 由于每个可见子状态的邻居数增加了一个 (不可见子状态), 因此ε的上限从$1/\Delta$降低为$1/(\Delta+1)$.

注 3. 相比于文献 [13, 37] 设计的更新式 (1), 本文在协议 (19) 的设计过程中考虑了系统通信链路中存在 DoS 攻击的情况, 可确保在存在一定能力 DoS 攻击时, 系统在协议 (19) 的约束下实现均值趋同.

3.3 输出均值趋同分析

在给出本文主要结论前, 需要下述引理知识.

引理 6. 考虑多智能体系统 (4), 如果其网络通信图是一个无向连通图, 则对于状态分解后的网络, 所有节点子状态总和是固定不变的.

证明. 由输出方程$y_i[k]=nx_i^\alpha[k]$, 推导可得:

$$y_i[k+1]=nx_i^\alpha[k+1] \tag{20}$$

再将式 (20) 代入式 (19), 可得:

$$\begin{cases} \dfrac{1}{n}y_i[k+1]=x_i^\alpha[k+1]=x_i^\alpha[k]+\\ \qquad\qquad \varepsilon\displaystyle\sum_{j=1}^{M}a'_{ij}[k](y_j[k]-y_i[k])+\\ \qquad\qquad \varepsilon a_{i,\alpha\beta}[k](x_i^\beta[k]-x_i^\alpha[k])\\ x_i^\beta[k+1]=x_i^\beta[k]+\varepsilon a_{i,\alpha\beta}[k](x_i^\alpha[k]-x_i^\beta[k]) \end{cases} \tag{21}$$

进一步, 由式 (21), 可得:

$$\begin{aligned} &x_i^\alpha[k+1]+x_i^\beta[k+1]=x_i^\alpha[k]+\\ &\quad x_i^\beta[k]+\varepsilon\sum_{j=1}^{M}a'_{ij}[k](y_j[k]-y_i[k])+\\ &\quad \varepsilon a_{i,\alpha\beta}[k](x_i^\beta[k]-x_i^\alpha[k])+\\ &\quad \varepsilon a_{i,\alpha\beta}[k](x_i^\alpha[k]-x_i^\beta[k])=\\ &\quad x_i^\alpha[k]+x_i^\beta[k]+\varepsilon\sum_{j=1}^{M}a'_{ij}[k](y_j[k]-y_i[k]) \end{aligned} \tag{22}$$

因此有:

$$\begin{aligned} \sum_{i=1}^{M}(x_i^\alpha[k+1]+x_i^\beta[k+1])=&\sum_{i=1}^{M}(x_i^\alpha[k]+x_i^\beta[k])+\\ \varepsilon\sum_{i=1}^{M}&\left\{\sum_{j=1}^{M}a'_{ij}[k]\big(y_j[k]-y_i[k]\big)\right\} \end{aligned} \tag{23}$$

而在式 (23) 中的$\sum_{i=1}^{M}\left\{\sum_{j=1}^{M}a'_{ij}[k]\big(y_j[k]-y_i[k]\big)\right\}$部分, 可进一步分解为下式:

$$\begin{aligned} &\sum_{i=1}^{M}\left\{\sum_{j=1}^{M}a'_{ij}[k]\big(y_j[k]-y_i[k]\big)\right\}=\\ &\quad\sum_{i=1}^{M}\{a'_{i1}[k](y_1[k]-y_i[k])+a'_{i2}[k](y_2[k]-\\ &\quad y_i[k])+\ldots+a'_{iM}[k](y_M[k]-y_i[k])\}=\\ &\quad a'_{11}[k](y_1[k]-y_1[k])+a'_{12}[k](y_2[k]-y_1[k])+\cdots+\\ &\quad a'_{1M}[k](y_M[k]-y_1[k])+\\ &\quad a'_{21}[k](y_1[k]-y_2[k])+a'_{22}[k](y_2[k]-y_2[k])+\cdots+\\ &\quad a'_{2M}[k](y_M[k]-y_2[k])+\\ &\quad a'_{31}[k](y_1[k]-y_3[k])+a'_{32}[k](y_2[k]-y_3[k])+\cdots+\\ &\quad a'_{3M}[k](y_M[k]-y_3[k])+\\ &\qquad\vdots\\ &\quad a'_{M1}[k](y_1[k]-y_M[k])+a'_{M2}[k](y_2[k]-\\ &\quad y_M[k])+\cdots+a'_{MM}[k](y_M[k]-y_M[k]) \end{aligned} \tag{24}$$

根据无向图属性: $a'_{ij}[k]=a'_{ji}[k]$, 对于任意v_i, $v_j\in\mathcal{V}$, 有:

$$a'_{ij}[k](y_j[k]-y_i[k])=-a'_{ji}[k](y_i[k]-y_j[k]) \tag{25}$$

将式 (25) 代入式 (24), 可得:

$$\sum_{i=1}^{M}\left\{\sum_{j=1}^{M}a'_{ij}[k]\big(y_j[k]-y_i[k]\big)\right\}=0 \tag{26}$$

将式 (26) 代入式 (23), 可得:

$$\sum_{i=1}^{M}(x_i^{\alpha}[k+1]+x_i^{\beta}[k+1])=\sum_{i=1}^{M}(x_i^{\alpha}[k]+x_i^{\beta}[k]) \quad (27)$$

由式 (27) 容易看出, 对于进行状态分解后的网络, 系统节点子状态的和是固定不变的. □

下面给出本文的主要结论.

定理 1. 考虑 DoS 攻击下多智能体系统 (4), 在满足假设 1、2 和 3 条件下, 若其通信拓扑满足 $(f+1)$-鲁棒图, 且系统节点在所给的分布式协议 (19) 下进行状态更新, 则系统可实现输出值均值趋同.

证明. 由于系统的通信图是一个 $(f+1)$-鲁棒图, 根据引理 5 可知, 系统在满足假设 1 的 DoS 攻击下, 其网络图仍能够保持连通. 显然, 经过状态分解之后的系统同样能够保证网络图的连通性. 根据引理 6, 当 $k=0$ 时, 有:

$$\frac{1}{2M}\sum_{i=1}^{M}(x_i^{\alpha}[1]+x_i^{\beta}[1])=\frac{1}{2M}\sum_{i=1}^{M}(x_i^{\alpha}[0]+x_i^{\beta}[0]) \quad (28)$$

随后, 根据引理 4 和式 (28) 可知, 系统可以实现均值趋同, 即任意节点的子状态 $x_i^{\alpha}[k]$ 和 $x_i^{\beta}[k]$ 都将收敛至:

$$\frac{1}{2M}\sum_{i=1}^{M}(x_i^{\alpha}[1]+x_i^{\beta}[1])$$

再根据式 (28) 和状态分解约束条件 $x_i^{\alpha}[0]+x_i^{\beta}[0]=2x_i[0]$, 可得:

$$\lim_{k\to\infty}x_i^{\alpha}[k]=\lim_{k\to\infty}x_i^{\beta}[k]=\frac{1}{M}\sum_{j=1}^{M}x_j[0] \quad (29)$$

最后, 根据式 (29) 和输出方程 $y_i[k]=nx_i^{\alpha}[k]$, 可得:

$$\lim_{k\to\infty}y_i[k]=\lim_{k\to\infty}nx_i^{\alpha}[k]=\frac{n}{M}\sum_{j=1}^{M}x_j[0] \quad (30)$$

□

注 4. 相比于文献 [13] 设计的隐私保护状态更新协议, 本文在协议 (19) 的设计过程中进一步考虑了在实际环境对测量条件等的限制导致难以获得系统中节点的真实状态值的情况, 引入了节点输出值的概念, 通过观测矩阵获取的系统输出 y 进行协议 (19) 的设计, 可确保系统在该协议下实现输出值均值趋同.

3.4　隐私保护分析

本节对趋同控制过程中单个节点信息的隐私保护进行分析. 本文考虑两种隐私窃听者: 好奇窃听者和外部窃听者. 好奇窃听者是指一类能够正确遵循所有控制协议步骤但具有好奇性的节点, 这类节点会收集接收到的数据并试图猜测其他节点的状态信息. 而外部窃听者是指一类了解整个网络拓扑结构的外部节点, 并能够窃听某些内部节点的通信链路从而获得在该通信链路交互的信息.

一般来说, 这里的外部窃听者比好奇窃听者更具有破坏力, 因为外部窃听者会窃听多个节点通信链路上交互的信息, 而好奇窃听者只能窃听该节点通信链路交互的信息, 但好奇窃听者有一个外部窃听者无法得知的信息, 即该好奇窃听者的初始状态值.

定义好奇窃听者 $v_i\in\mathcal{V}$ 在第 k 次迭代时所获得的信息为: $I_i[k]=\{a'_{ip}[k]|_{v_p\in\mathcal{N}_i}, y_p[k]|_{v_p\in\mathcal{N}_i}, x_i[k], x_i^{\alpha}[k], x_i^{\beta}[k], a_{i,\alpha\beta}[k]\}$. 随着状态值迭代更新, 窃听者 v_i 收集获得的信息表示为 $I_i=\bigcup_{k=0}^{\infty}I_i[k]$.

定义 4. 如果窃听者无法以任何精度保证估计节点状态信息 $x_i[0]$ 的值, 则称节点 v_i 得到了隐私保护.

在给出结论前, 需要用到下述引理.

引理 7[13]**.** 在采用状态分解方法的信息交互通信中, 如果正常节点 v_j 具有至少一个不与好奇窃听节点 v_i 直接相连的正常邻居节点 v_m, 则对于节点 v_j 的任意初始状态 $\bar{x}_j[0]\neq x_j[0]$, 窃听节点 v_i 获得的信息始终满足 $\bar{I}_i=I_i$.

引理 8[13]**.** 在采用状态分解方法的信息交互通信中, 如果正常节点 v_j 存在至少一个正常邻居节点 v_m, 其 $a_{jm}[0]$ 的值对于外部窃听者不可见, 则节点 v_j 的任意初始状态的任何变化都可以完全通过对外部窃听者不可见的 $a_{jm}[0]$, $a_{j,\alpha\beta}[0]$ 和 $a_{m,\alpha\beta}[0]$ 的变化来补偿, 因此外部窃听者无法以任何精度保证估计正常节点 v_j 的初始状态值 $x_j[0]$.

定理 2. 考虑 DoS 攻击下多智能体系统 (4), 对于系统中任意正常节点 $v_j\in\mathcal{V}$, 如果 v_j 在所给的分布式协议 (19) 下进行状态更新, 则在整个信息交互过程中, 其状态信息值 $x_j[0]$ 具备隐私保护.

证明. 首先, 分析系统存在好奇窃听者 v_i 的情况. 对于任意正常节点 v_j, 在所给的分布式协议 (19) 下, 其初始状态显然满足 $\bar{x}_j[0]\neq x_j[0]$, $\bar{I}_i=I_i$. 再由引理 6 可知, 该条件下好奇窃听者无法准确估计节点 v_j 的初始值, 因此节点 v_j 的状态值 $x_j[0]$ 得到了隐私保护.

随后, 分析系统存在外部窃听者的情况. 在本文所提的分布式算法 (19) 下, 外部窃听者对于系统中任意正常节点 $v_j\in\mathcal{V}$ 的其中之一子状态不可见. 根据引理 7, v_j 初始状态值的变化则对于外部窃听者不可见, 故外部窃听者无法准确估计正常节点 v_j 的

初始值, 因此节点 v_j 的状态值$x_j[0]$ 得到了隐私保护. □

4 数值仿真

本节通过一些仿真算例来验证提出算法的有效性. 同时与文献 [13] 中的算法进行了比较, 以验证其优势.

考虑由 5 个节点组成的无向图网络, 其通信拓扑如图 3 所示. 每个节点分别赋予初始状态值 $x_1[0]=3,\ x_2[0]=6,\ x_3[0]=9,\ x_4[0]=12,\ x_5[0]=15$.

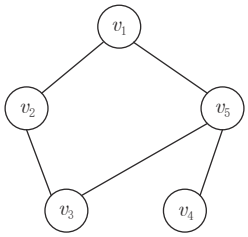


图 3 5 个节点组成的通信图

Fig. 3 Network topology of multi-agent system with 5 nodes

首先考虑节点采用文献 [13] 中的控制律更新状态, 令 $\varepsilon=0.2$, 对应的邻接矩阵设置为:

$$\boldsymbol{A}_1=0.75\begin{bmatrix}0&1&0&0&1\\1&0&1&0&0\\0&1&0&0&1\\0&0&0&0&1\\1&0&1&1&0\end{bmatrix} \tag{31}$$

值得注意的是, 文献 [13] 并未涉及输出方程, 因此不妨用 x_i 表示节点 v_i 的内部状态, x_i^+ 表示和邻居节点进行交互的状态值.

考虑存在外部窃听者试图猜测节点 v_1 的初始状态值 $x_1[0]$. 可以构造以下的外部窃听者对节点 v_1 的初始状态值进行估计:

$$z[k+1]=z[k]+x_1^+[k+1]-\left(x_1^+[k]+\varepsilon\sum_{j=1}^{M}a_{1j}(x_j^+[k]-x_1^+[k])\right) \tag{32}$$

式中, $z[k]$表示窃听者在 k 时刻获知的观测状态. 显然, 外部窃听者初始观测状态 $z[0]=x_1^+[0]$. 假设该外部窃听者除了不可获知节点 v_1 与 v_2 之间的权值 $a_{12}[0]$外, 具备整个网络其他的拓扑信息. 此时给 $a_{12}[0]$ 随机赋值 0.7.

系统中各个节点的状态变化轨迹如图 4 所示. 图中实线表示各节点 α 子状态值变化曲线, 点连线表示外部窃听者对节点 v_1 初始状态值的猜测曲线,

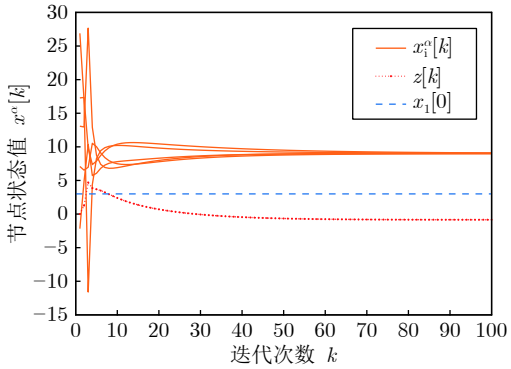


图 4 控制方法下的各节点状态轨迹[13]

Fig. 4 State trajectory of each node with control law[13]

虚线表示节点 v_1 的初始状态值. 可以看出, 在无 DoS 攻击影响下, 文献 [13] 中的控制方法可使系统准确收敛到所有节点初始状态的均值 9, 且外部窃听者不能准确推断节点 v_1 的初始状态 $x_1[0]=3$.

然后, 考虑网络中存在 DoS 攻击的情况. 不妨令 $f=2$, 相应的, Adversory (P, k_0) 中的攻击矩阵为:

$$\boldsymbol{P}=\begin{bmatrix}1&0&1&1&1\\0&1&1&1&1\\1&1&1&1&0\\1&1&1&1&1\\1&1&0&1&1\end{bmatrix} \tag{33}$$

令 $k_0=5$, 即表示系统在第 5 次状态更新时, 开始发生 DoS 攻击, 并且持续发生至更新结束.

图 5 显示的是在上述攻击影响下基于文献 [13] 中算法的系统各个节点的状态变化曲线. 图 5 中实线表示各节点 α 子状态值变化曲线, 点连线表示外部窃听者对节点 v_1 初始状态值的猜测曲线, 虚线表示节点 v_1 的初始状态值. 从中可以明显看到, 在 DoS 攻击下, 尽管系统仍具备保护隐私的能力, 但系统中各个节点的状态则无法达成趋同.

考虑遭受同样满足 Adversory (P, k_0) 的 DoS 攻击, 重新设计网络拓扑结构, 使其满足定理 1 中的通信图要求, 即满足 3-鲁棒图. 设计后的网络通信图如图 6 所示.

同样令 $\varepsilon=0.2$, 邻接矩阵设置为:

$$\boldsymbol{A}_2=0.75\begin{bmatrix}0&1&1&1&1\\1&0&1&1&0\\1&1&0&1&1\\1&1&1&0&1\\1&0&1&1&0\end{bmatrix} \tag{34}$$

同样考虑存在一个外部窃听者试图猜测节点 v_1 的初始状态值. 考虑此时传输的是输出状态, 令 y_i

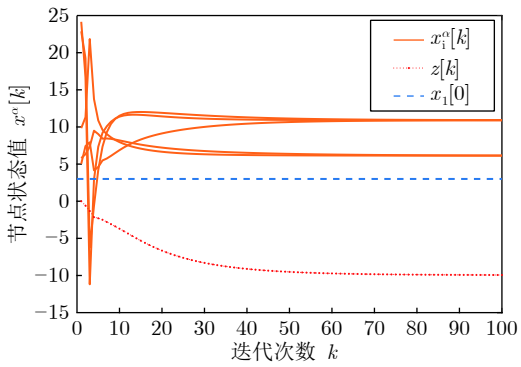


图 5　DoS 攻击影响下各节点状态轨迹[13]

Fig. 5　State trajectory of each node with control law under DoS attacks[13]

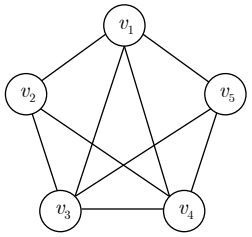


图 6　5 个节点组成的新通信图

Fig. 6　New network topology of 5 nodes

表示节点 v_i 和邻居节点进行交互的信息值, 输出方程中观测矩阵前系数 $n = 1$, 构造以下的外部窃听者对节点 v_1 的初始状态值进行猜测:

$$z[k+1] = z[k] + y_1[k+1] - \left(y_1[k] + \varepsilon \sum_{j=1}^{M} a_{1j}(y_j[k] - y_1[k]) \right) \tag{35}$$

外部窃听者初始观测状态 $z[0] = y_1[0]$. 根据定理 2 中的条件, 实例中假设外部窃听者可以获得除了节点 v_1 和 v_2 之间的权值 $a_{12}[0]$ 外的整个拓扑图信息. 同样为 $a_{12}[0]$ 随机赋值 0.7. 因此, 根据定理 1 和定理 2 可知, 系统节点在满足图 6 所给的通信图下, 在本文控制协议下可以达成均值趋同, 且每个节点的状态信息得到隐私保护.

图 7 显示的是在 DoS 攻击下系统节点在本文所提控制方法下的状态变化轨迹. 其中实线表示各节点输出值变化曲线, 点连线表示外部窃听者对节点 v_1 初始状态值的猜测曲线, 虚线表示节点 v_1 的初始状态值. 由图 7 可以看出, 系统在 DoS 攻击下仍然可以收敛到所有节点初始输出状态的平均值 9, 且外部窃听者无法成功估计节点 v_1 的初始值. 仿真结果验证了本文方法的有效性.

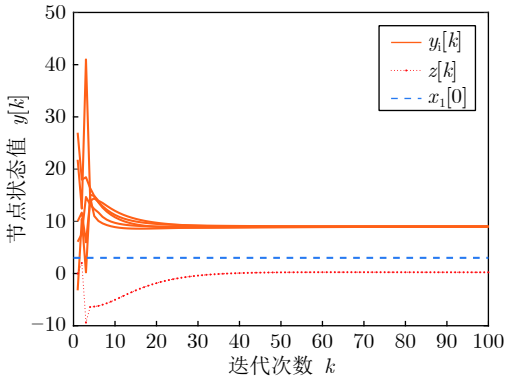


图 7　DoS 攻击下系统中各节点在本文所提算法下的输出状态轨迹

Fig. 7　State trajectory of each node with our proposed control law under DoS attacks

5　结束语

本文针对 DoS 攻击下无向多智能体网络趋同问题, 提出了一种具有节点信息隐私保护能力的趋同方法, 实现了 DoS 攻击下系统输出值的均值趋同. 结合状态分解方法, 设计了节点输出状态的初始值隐私保护策略, 保证所有节点信息在分布式算法邻居间信息交互过程中不被窃听者窃取. 进一步, 借助 r-鲁棒图概念, 分析给出了系统在所建立 DoS 攻击模型下的通信网络拓扑要求, 并设计了相应的输出均值趋同控制协议. 最后, 通过一组数值仿真实验验证了所提方法的有效性.

同时, 本文目前研究设计的输出均值趋同控制协议还存在着一定的不足值得在进一步的研究中进行改进. 1) 该输出均值趋同协议目前仅适用于无向拓扑网络, 而在现实生活中, 有向图更加普遍, 因此接下来的研究将致力于针对有向拓扑情况, 对本文所设计的输出均值趋同控制协议进行进一步的改进. 2) 本文为了抵御系统所遭受的 DoS 攻击, 通过增加通信链路来增强网络拓扑的连通性, 这一定程度上增加了通信的成本, 因此寻找一个能够解决通信成本大幅度增加的问题的方法是接下来值得研究的方向.

References

1　Li Tao, Meng Yang, Zhang Ji-Feng. An overview on quantized consensus and consensus with limited data rate of multi-agent systems. *Acta Automatica Sinica*, 2013, **39**(11): 1805−1811
(李韬, 孟杨, 张纪峰. 多自主体量化趋同与有限数据率趋同综述. 自动化学报, 2013, **39**(11): 1805−1811)

2　Yang Hong-Yong, Guo Lei, Zhang Yu-Ling, Yao Xiu-Ming. Delay consensus of fractional-order multi-agent systems with sampling delays. *Acta Automatica Sinica*, 2014, **40**(9): 2022−2028
(杨洪勇, 郭雷, 张玉玲, 姚秀明. 离散时间分数阶多自主体系统的

时延一致性. 自动化学报, 2014, **40**(9): 2022−2028)

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

4 Shamir A. How to share a secret. *Communication of the ACM*, 1979, **22**: 612−613

5 Kreuter B, Shelat A, Shen C H. Billion-gate secure computation with malicious adversaries. In: Proceedings of the 21st USENIX Conference on Security Symposium. Berkeley, USA: USENIX Association, 2012. 285−300

6 Chen W, Cai S. Ad hoc peer-to-peer network architecture for vehicle safety communications. *IEEE Communications Magazine*, 2005, **43**(4): 100−107

7 Huang Z, Mitra S, Vaidya N. Differentially private distributed optimization. In: Proceedings of the 2015 International Conference on Distributed Computing and Networking. NewYork, USA: ACM, 2015. 1−10

8 Nozari E, Tallapragada P, Cortes J. Differentially private average consensus: Obstructions, trade-offs, and optimal algorithm design. *Automatica*, 2017, **81**(7): 221−231

9 Katewa V, Pasqualetti F, Gupta V. On privacy vs cooperation in multi-agent systems. *International Journal of Control*, 2018, **91**(7): 1693−1707

10 Huang Z, Mitra S, Dullerud G. Differentially private iterative synchronous consensus. In: Proceedings of the 2012 ACM Workshop on Privacy in The Electronic Society. NewYork, USA: ACM, 2012. 81−90

11 Nozari E, Tallapragada P, Cortes J. Differentially private average consensus with optimal noise selection. *IFAC-PapersOnline*, 2015, **48**(22): 203−205

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

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

14 Kia S, Cortes J, Martinez S. Dynamic average consensus under limited control authority and privacy requirements. *International Journal of Robust and Nonlinear Control*, 2015, **25**(13): 1941−1966

15 Pequito S, Kar S, Sundaram S, Aguiar A P. Design of communication networks for distributed computation with privacy guarantees. In: Proceedings of the 2015 54th IEEE Conference on Decision and Control (CDC). Osaka, Japan: IEEE, 2015. 1370−1376

16 Alaeddini A, Morgansen K, Mesbahi M. Adaptive communication networks with privacy guarantees. In: Proceedings of the 2017 American Control Conference (ACC). Seattle, USA: IEEE, 2017. 4460−4465

17 Gupta N, Chopra N. Confidentiality in distributed average information consensus. In: Proceedings of the 2016 IEEE 55th Conference on Decision and Control (CDC). Las Vegas, USA: IEEE, 2016. 6709−6714

18 Kogiso K, Fujita T. Cyber-security enhancement of networked control systems using homomorphic encryption. In: Proceedings of the 2015 54th IEEE Conference on Decision and Control (CDC). Osaka, Japan: IEEE, 2015. 6836−6843

19 Shoukry Y, Gatsis K, Alanwar A, Pappas G J, Seshia S A, Srivastava M, et al. Privacy-aware quadratic optimization using partially homomorphic encryption. In: Proceedings of the 2016 IEEE 55th Conference on Decision and Control (CDC). Las Vegas, USA: IEEE, 2016. 5053−5058

20 Lagendijk R L, Erkin Z, Barni M. Encrypted signal processing for privacy protection: Conveying the utility of homomorphic encryption and multiparty computation. *IEEE Signal Processing Magazine*, 2013, **30**(1): 82−105

21 Zhang C, Ahmad M, Wang Y. Admm based privacy-preserving decentralized optimization. *IEEE Transactions on Information Forensics and Security*, 2019, **14**(3): 565−580

22 LeBlanc H J, Zhang H, Koutsoukos X, Sundaram S. Resilient asymptotic consensus in robust networks. *IEEE Journal on Selected Areas in Communications*, 2013, **31**(4): 766−781

23 Wu Y, He X. Secure consensus control for multiagent systems with attacks and communication delays. *IEEE/CAA Journal of Automatica Sinica*, 2017, **4**(1): 136−142

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

25 Huang Jin-Bo, Wu Yi-Ming, He Xiong-Xiong, Chang Li-Ping. Secure consensus control for heterogeneous multi-agent systems with trusted nodes. *Sci Sin Inform*, 2019, **49**(5): 599−612
(黄锦波, 伍益明, 何熊熊, 常丽萍. 信任节点机制下的异构多智能体系统安全一致性控制. 信息科学, 2019, **49**(5): 599−612)

26 Usevitch J, Panagou D. Resilient leaderfollower consensus to arbitrary reference values in time-varying graphs. *IEEE Transactions on Automatic Control*, 2020, **65**(4): 1755−1762

27 He W, Mo Z, Han Q L, Qian F. Secure impulsive synchronization in lipschitztype multi-agent systems subject to deception attacks. *IEEE/CAA Journal of Automatica Sinica*, 2020, **5**(7): 1326−1334

28 Fu W, Qin J, Shi Y, Zheng W X, Kang Y. Resilient consensus of discrete-time complex cyber-physical networks under deception attacks. *IEEE Transactions on Industrial Informatics*, 2020, **16**(7): 4868−4877

29 Zhang D, Liu L, Feng G. Consensus of heterogeneous linear multiagent systems subject to aperiodic sampled-data and dos attack. *IEEE Transactions on Cybernetics*, 2019, **49**(4): 1501−1511

30 Feng Z, Hu G. Secure cooperative event-triggered control of linear multiagent systems under dos attacks. *IEEE Transactions on Control Systems Technology*, 2020, **28**(3): 741−752

31 Yang Y, Xu H, Yue D. Observer-based distributed secure consensus control of a class of linear multi-agent systems subject to random attacks. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2019, **66**(8): 3089−3099

32 Xu W, Hu G, Ho D W, Feng Z. Distributed secure cooperative control under denial-of-service attacks from multiple adversaries. *IEEE Transactions on Cybernetics*, 2020, **50**(8): 3458−3467

33 Wen G, Wang P, Huang T, Lü J, Zhang F. Distributed consensus of layered multi-agent systems subject to attacks on edges. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2020, **67**(9): 3152−3162

34 Zhang H, Fata E, Sundaram S. A notion of robustness in complex networks. *IEEE Transactions on Control of Network Systems*, 2015, **2**(3): 310−320

35 Yin X, Tan C, Huang J. Output consensus for networks of heterogeneous linear agents. In: Proceedings of the 2017 36th Chinese Control Conference (CCC). Dalian, China: IEEE, 2017. 7679−7683

36 Zhang H, Sundaram S. Robustness of information diffusion al-

gorithms to locally bounded adversaries. In: Proceedings of the 2012 American Control Conference (ACC). Montréal, Canada: IEEE, 2012. 5855−5861

37 Olfati-Saber R, Fax J A, Murray R M. Consensus and cooperation in networked multi-agent systems. *Proceedings of the IEEE*, 2007, **95**(1): 215−233

38 Nedic A, Ozdaglar A, Parrilo P A. Constrained consensus and optimization in multi-agent networks. *IEEE Transactions on Automatic Control*, 2010, **55**(4): 922−938

39 Yang Y, Li Y, Yue D, Tian Y, Ding X. Distributed secure consensus control with event-triggering for multiagent systems under dos attacks. *IEEE Transactions on Cybernetics*, DOI: 10.1109/TCYB.2020.2979342.

胡沁伶 杭州电子科技大学网络空间安全学院硕士研究生. 2015 年获杭州电子科技大学学士学位. 主要研究方向为多智能体系统网络安全和隐私保护. E-mail: Hazelhu0601@126.com

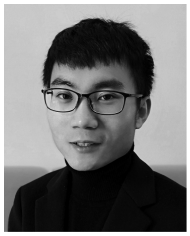
(**HU Qin-Ling** Master student at the School of Cyberspace, Hangzhou Dianzi University. She received her bachelor degree from Hangzhou Dianzi University in 2015. Her research interest covers cyber-security for multi-agent systems and privacy-preserving.)

郑 宁 杭州电子科技大学网络空间安全学院研究员. 1990 年获浙江大学硕士学位. 主要研究方向为信息安全, 信息管理系统和多智能体系统.
E-mail: nzheng@hdu.edu.cn

(**ZHENG Ning** Professor at the School of Cyberspace, Hangzhou Dianzi University. He received his master degree from Zhejiang University in 1990. His research interest covers information security, information management system and multi-agent network.)

徐 明 杭州电子科技大学网络空间安全学院教授. 2004 年获浙江大学博士学位. 主要研究方向为网络信息安全和数字取证.
E-mail: mxu@hdu.edu.cn

(**XU Ming** Professor at the School of Cyberspace, Hangzhou Dianzi University. He received his Ph.D. degree from Zhejiang University in 2004. His research interest covers network security and digital forensics.)



伍益明 杭州电子科技大学网络空间安全学院副教授. 2016 年获浙江工业大学控制科学与工程博士学位. 主要研究方向为分布式系统安全控制, 多智能体系统网络安全和迭代学习控制. 本文通信作者.
E-mail: ymwu@hdu.edu.cn

(**WU Yi-Ming** Associate professor at the School of Cyberspace, Hangzhou Dianzi University. He received his Ph.D. degree from Zhejiang University of Technology in 2016. His research interest covers distributed system secure control, cyber-security for multi-agent systems and iterative learning control. Corresponding author of this paper.)

何熊熊 浙江工业大学信息工程学院教授. 1997 年获浙江大学博士学位. 主要研究方向为迭代学习控制和智能控制及其在多智能体系统和传感器网络中的应用.
E-mail: hxx@zjut.edu.cn

(**HE Xiong-Xiong** Professor at the College of Information Engineering, Zhejiang University of Technology. He received his Ph.D. degree from Zhejiang University in 1997. His research interest covers iterative learning control, intelligent control, applications in multi-agent systems and sensor networks.)