



计算机科学

COMPUTER SCIENCE

有向网络下隐私在线约束优化问题的状态分解分布式对偶平均算法

代祥光, 贺成龙, 管明宇, 张伟, 周炆, 刘建峰, 吕庆国

引用本文

代祥光, 贺成龙, 管明宇, 张伟, 周炆, 刘建峰, 吕庆国. 有向网络下隐私在线约束优化问题的状态分解分布式对偶平均算法[J]. 计算机科学, 2025, 52(8): 411-420.

DAI Xiangguang, HE Chenglong, GUAN Mingyu, ZHANG Wei, ZHOU Yang, LIU Jianfeng, LYU Qingguo. State-decomposition Distributed Dual Averaging Algorithm for Privacy Online Constrained Optimization over Directed Networks [J]. Computer Science, 2025, 52(8): 411-420.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于代理人的区块链双向混币协议](#)

Proxy-based Bidirectional Coin Mixing Mechanism of Blockchain

计算机科学, 2025, 52(8): 385-392. <https://doi.org/10.11896/jsjcx.240600079>

[基于像素区间划分及预测恢复的完全加密图像可逆信息隐藏](#)

Reversible Data Hiding in Fully Encrypted Images Based on Pixel Interval Partitioning and Prediction Recovery

计算机科学, 2025, 52(6A): 240900030-8. <https://doi.org/10.11896/jsjcx.240900030>

[基于对抗生成网络的众包内容隐私保护](#)

Privacy Preservation of Crowdsourcing Content Based on Adversarial Generative Networks

计算机科学, 2025, 52(6A): 250200123-7. <https://doi.org/10.11896/jsjcx.250200123>

[一种结合数据集蒸馏的联邦学习隐私保护方法](#)

Federated Learning Privacy Protection Method Combining Dataset Distillation

计算机科学, 2025, 52(6A): 240500132-7. <https://doi.org/10.11896/jsjcx.240500132>

[YOLO-BFEPS:一种高效注意力增强的跨尺度YOLOv10火灾检测模型](#)

YOLO-BFEPS:Efficient Attention-enhanced Cross-scale YOLOv10 Fire Detection Model

计算机科学, 2025, 52(6A): 240800134-9. <https://doi.org/10.11896/jsjcx.240800134>

有向网络下隐私在线约束优化问题的状态分解分布式对偶平均算法

代祥光¹ 贺成龙² 管明宇¹ 张伟¹ 周炆³ 刘建峰¹ 吕庆国⁴

1 重庆三峡学院智能信息处理与控制重庆高校市级重点实验室 重庆 404100

2 重庆三峡学院数学与统计学院 重庆 404100

3 腾讯科技有限公司 广州 510000

4 重庆大学计算机学院 重庆 400715

(daixiangguang@163.com)

摘要 文中研究了一类具有公共约束集的分布式在线约束优化问题。网络中的各个节点通过局部计算和通信协同求解该问题,每个节点仅能访问自身的局部损失函数,该函数的取值依赖于其每次迭代中的决策变量。然而,由于节点在通信过程中不断广播与隐私相关的信息,大多数现有算法可能面临隐私泄露的风险。为此,提出了一种高效的状态分解分布式对偶平均算法。该算法结合状态分解和梯度调整策略,以增强隐私保护能力,同时消除有向网络中的不平衡性。值得注意的是,该算法无需额外的隐藏信号,也不会显著增加计算量。理论分析表明,该方法不仅能实现次线性遗憾,还能够有效保护各节点损失函数的隐私性。此外,通过仿真实验验证了该算法的收敛性和可行性。

关键词: 在线约束优化;分布式对偶平均算法;状态分解;有向网络;隐私保护

中图分类号 TP301

State-decomposition Distributed Dual Averaging Algorithm for Privacy Online Constrained Optimization over Directed Networks

DAI Xiangguang¹, HE Chenglong², GUAN Mingyu¹, ZHANG Wei¹, ZHOU Yang³, LIU Jianfeng¹ and LYU Qingguo⁴

1 Key Laboratory of Intelligent Information Processing and Control, Chongqing Three Gorges University, Chongqing 404100, China

2 School of Mathematics and Statistics, Chongqing Three Gorges University, Chongqing 404100, China

3 Tencent Technology Co., Ltd., Guangzhou 510000, China

4 Computer Science, Chongqing University, Chongqing 400715, China

Abstract This paper investigates a class of distributed online constrained optimization problems with a common constraint set. In this setting, nodes in the network collaborate to solve the problem through local computation and communication. Each node can only access its own local loss function, whose value depends on its decision variables at each iteration. However, since nodes continuously broadcast information related to their private data during communication, most existing algorithms face the risk of privacy leakage. To address this issue, this paper proposes an efficient state decomposition-based distributed dual averaging algorithm. This algorithm integrates state decomposition and gradient adjustment strategies to enhance privacy protection while mitigating imbalances in directed networks. Notably, it does not require additional hidden signals or significantly increase computational complexity. Theoretical analysis shows that the proposed method achieves the desired sublinear regret while effectively preserving the privacy of each node's loss function. Furthermore, simulation experiments confirm the convergence and feasibility of the algorithm.

Keywords Online constrained optimization, Distributed dual averaging, State decomposition, Directed networks, Privacy protection

到稿日期:2025-03-17 返修日期:2025-05-07

基金项目:重庆市自然科学基金(CSTB2023NSCQ-LZX0135, CSTB2022NSCQ-MSX1627, CSTB2022NSCQ-MSX1285);重庆市教委科学技术研究项目(KJZD-M202201204, KJZD-K202201205, KJQN202301230)

This work was supported by the Chongqing Natural Science Foundation (CSTB2023NSCQ-LZX0135, CSTB2022NSCQ-MSX1627, CSTB2022NSCQ-MSX1285) and Chongqing Municipal Education Commission Science and Technology Research Project (KJZD-M202201204, KJZD-K202201205, KJQN202301230).

通信作者:张伟(weizhang@sanxiao.edu.cn)

1 引言

近年来,随着交通控制、计算机网络和传感器网络等实际应用对优化问题的需求不断增长,优化技术在数学、控制和人工智能等多个学科领域^[1-2]也得到了广泛应用。传统方法在求解大规模优化问题时面临诸多挑战,主要原因在于中央节点计算能力的限制。而多节点网络优化技术的发展,使得分布式优化逐渐受到广泛关注,因为它对通信、计算和存储资源的需求较低。多节点优化问题可以简单描述为:在由多个节点组成的通信网络中,各节点通过信息交换协同求解全局优化问题。

1.1 文献综述

截至目前,针对不同的分布式优化场景,研究人员已设计出多种算法。在这些场景中,最简单的情况是优化无约束目标函数,而更复杂的情况涉及全局约束集的分布式优化,即决策变量被限制在一个非空闭凸集内。现有的分布式优化算法大致可分为两类:1)基于共识的算法,这类方法主要依赖于节点间信息的一致性,典型算法包括异步分布式梯度算法^[3]、随机镜像下降算法^[4]以及分布式拟单调次梯度算法^[5]等;2)基于对偶分解的算法,通过对偶变量来分解优化问题,代表性算法包括分布式对偶近端梯度算法^[6]和分布式前向-后向布雷格曼分裂算法^[7]。此外,一些研究进一步扩展了分布式优化框架。例如,文献^[8]研究了分布式聚合优化,其中局部目标函数不仅依赖于自身决策变量,还依赖于其他节点的变量(称为聚合变量)。基于此,研究者提出了分布式聚合梯度跟踪算法并证明了其收敛性。文献^[9]提出了一种新颖的分布式优化算法,在其中引入盈余变量以克服不平衡通信网络,并放宽了双随机矩阵条件,通过构造行随机矩阵控制决策变量、列随机矩阵控制盈余变量。在应用方面,文献^[10]提出了一种并行分布式优化方法,用于微电网能源管理。该方法允许节点并行计算、交换信息和更新决策,在不牺牲最优解精度的前提下,提高了收敛速度。此外,文献^[11]研究了一般的平滑、非凸、有限和优化问题,并提出了一种基于方差缩减的梯度下降算法,该算法在较温和的假设条件下具有较低的计算复杂度。在实际优化问题中,由于信息传输通常是单向的,因此通信网络往往是不平衡的^[12-14]。为解决这一问题,研究者开发了多种算法,例如推和策略和基于盈余^[15]的策略。文献^[16-17]在对偶平均算法中采用次梯度缩放策略,其关键思想是利用权重矩阵的归一化 Perron 特征向量估计值,对各节点的次梯度进行调整,使其能够根据自身及邻居节点的状态加权平均地更新自身变量。这一思想正是本文的研究重点。

除此之外,在优化算法的迭代过程中,每个节点需要与相邻节点交换信息,并根据目标函数的梯度进行迭代计算,这往往涉及隐私泄露的问题。在实际优化任务中,节点通常不希望其私有信息被外界或其他节点获取。例如,在医疗保健系统中,每位患者都希望自己的个人信息保持私密性,而不被泄露。因此,研究具备隐私保护机制的分布式优化算法具有重要的现实意义。在分布式优化中,已有多种隐私保护策略被提出。其中,差分隐私策略是最常见的方法之一,其基本原理是在数据中添加服从特定分布的噪声,以防止外部观察者准

确获取原始数据。例如,文献^[17]采用了这种方法进行隐私保护。然而,由于噪声的引入,该方法在提升隐私安全性的同时,可能会降低优化精度。因此,研究者需要在求解精度和隐私保护水平之间进行权衡。另一种常见的隐私保护方法是数据加密,该技术在每次信息交互时都进行加密解密的操作,从而保障数据安全。然而,在资源受限的网络环境下,数据加密会显著增加通信和计算成本,限制其实用性。而对于某些特定网络结构的分布式优化问题,研究者还提出了一些针对性的隐私保护方法。例如,文献^[18]提出了一种基于函数分解的隐私保护方法,该方法能够保留每个节点的私有次梯度信息,从而同时保障隐私安全性和优化算法的收敛精度。而文献^[19]在一般共识算法的基础上提出了状态分解机制来保护隐私,其核心在于阐述这种状态分解机制对于隐私保护的有效性,能够确保准确的共识。这一方法正是本文研究的核心内容。

1.2 主要贡献

在上述研究的基础上,本文针对非平衡有向网络上的在线分布式约束优化问题,结合隐私保护进行了深入探讨,提出了一种基于对偶平均策略的在线优化算法,通过状态分解来实现隐私保护。本文的主要贡献体现在以下几个方面:

1)针对非平衡有向网络上的分布式约束优化问题,提出了一种状态分解分布式对偶平均算法,并严格证明了该算法的遗憾相对时间呈次线性,保证了算法的有效性。

2)现有基于差分隐私技术的优化算法通常只能收敛到近似最优解,而本文提出的算法能够收敛至精确最优解。与基于同态加密技术的优化方法相比,本算法更加简洁,且不会显著增加计算负担。

3)提出了一种高效的隐私保护机制,适用于带约束的分布式优化问题。与文献^[19]中的算法相比,本文使用的权重矩阵为行随机矩阵,该矩阵能够更灵活地适应不同的网络拓扑结构。此外,通过梯度缩放策略,本算法能够有效克服有向网络中的不平衡问题,提升优化效果。

符号说明:对于具有 N 个节点的多节点网络系统,它们之间的通信结构为有向网络,用 $G=(\mathcal{V},\mathcal{E},\mathbf{A})$ 来表示,其中 $\mathcal{V}=\{1,2,\dots,N\}$ 是节点集, $\mathcal{E}\subseteq\mathcal{V}\times\mathcal{V}$ 是边集, $\mathbf{A}=[a_{ij}]\in\mathbb{R}^{N\times N}$ 是权重矩阵。当 $(i,j)\in\mathcal{E}$ 或 $i=j$ 时, $a_{ij}>0$,否则 $a_{ij}=0$ 。在该网络中,节点 j 从其入邻居集 $\mathcal{N}_j^{\text{in}}=\{i\in\mathcal{V}\mid(i,j)\in\mathcal{E}\}$ 中的节点接收信息,出邻居集表示为 $\mathcal{N}_j^{\text{out}}=\{i\in\mathcal{V}\mid(j,i)\in\mathcal{E}\}$ 。假设限制在该有向网络中的节点只能访问其入邻居集中节点的信息。

2 问题重述

对于每个节点 $i\in\mathcal{V}$,都有一组局部损失函数 $\{f_{i,k}\}_{k=1}^K$,其具体形式是未知的。在每次迭代中,节点根据当前的决策变量产生相应的损失。因此,分布式优化问题可以表示为:在每次迭代 $k\in\{1,2,\dots,K\}$ 中,网络中的每个节点以合作的方式最小化全局目标函数:

$$\begin{aligned} \min f_k(\mathbf{x}) &= \sum_{i=1}^N f_{i,k}(\mathbf{x}) \\ \text{s. t. } \mathbf{x} &\in \Omega \end{aligned} \quad (1)$$

其中, $\Omega \in \mathbb{R}^m$ 是非空, 闭凸的; $\mathbf{x} \in \Omega$ 是全局决策变量; 每个节点 i 只能访问关于 f_i 的信息, 对任何节点来说全局函数 f 是未知的。

在本文中, 将遗憾作为评估所提算法有效性的度量, 以下是节点 $j \in \mathcal{V}$ 在时间范围 K 内的遗憾定义^[20]:

$$R_{j,k} = \sum_{k=1}^K \sum_{i=1}^N f_{i,k}(\mathbf{x}_{j,k}) - \sum_{k=1}^K \sum_{i=1}^N f_{i,k}(\mathbf{x}^*) \quad (2)$$

其中, $\mathbf{x}_{j,k}, j \in \{1, 2, \dots, N\}$ 是 \mathbf{x} 在时间 k 时的局部变量; $\mathbf{x}^* = \arg \min_{\mathbf{x} \in \Omega} \sum_{k=1}^K \sum_{i=1}^N f_{i,k}(\mathbf{x})$ 是在时间范围 K 上的最佳决策变量。本文的目的是设计一种算法, 使得式(1)中的遗憾相对于时间 K 呈次线性, 即 $\lim_{K \rightarrow \infty} R_{j,k}/K = 0, j \in \mathcal{V}$ 。根据遗憾的定义, 若动态遗憾相对于时间增长呈次线性, 则说明算法所生成的决策变量逐渐趋近于最优决策变量, 从而实现渐近最优性, 说明了算法的有效性。然而, 当节点之间存在通信延迟或噪声时, 信息传递的及时性和准确性将受到影响, 从而影响算法对动态最优解的跟踪能力, 导致动态遗憾的增加。因此, 分析不受影响情况下的动态遗憾有助于深入理解算法的性能边界, 并为算法的改进和优化提供理论依据。正因如此, 我们得出信道环境是良好的。以下给出相关假设。

假设 1 通信拓扑 \mathcal{G} 是一个固定且强连通的有向网络, 其权重矩阵 \mathbf{A} 是行随机的。

假设 2 非空闭凸集 Ω 直径的上界为 $M > 0$; 局部损失函数 $f_{i,k}(\mathbf{x})$ 是凸函数, 其次梯度 $\mathbf{g}_{i,k}(\mathbf{x})$ 是 L -利普西茨连续 ($L > 0$), 即对于 $\forall k \geq 1, i \in \mathcal{V}, \mathbf{x} \in \Omega$, 有 $\|\mathbf{g}_{i,k}(\mathbf{x})\| \leq L$ 。

3 算法构建

3.1 分布式次梯度下降算法

在提出该算法之前, 本文根据文献[21]中提出的基于梯度调整的分布式次梯度算法进行研究。它可以描述为: 在时间 k , 对于 $\forall i \in \mathcal{V}$, 有

$$\mathbf{x}_{i,k+1} = \sum_{j=1}^N a_{ij} \mathbf{x}_{j,k} - \eta_k \frac{\mathbf{g}_{i,k}}{\mathbf{z}_{i,k}}, \mathbf{z}_{i,k} = \sum_{j=1}^N a_{ij} \mathbf{z}_{j,k} \quad (3)$$

其中, 决策变量 $\mathbf{x}_{i,k} \in \mathbb{R}^m$, 函数 $f_i(\mathbf{x})$ 在 $\sum_{j=1}^N a_{ij} \mathbf{x}_{j,k}$ 处的次梯度为 $\mathbf{g}_{i,k}$, 递减步长为 $\eta_k, \mathbf{z}_{i,k} \in \mathbb{R}^N, \mathbf{z}_{i,k}$ 是向量 $\mathbf{z}_{i,k}$ 的第 i 个元素, $z_{i,0} = 1, z_{ij,0} = 0 (i \neq j)$ 。

引理 1 在假设 1 下, 对于任意 $i, j \in \mathcal{V}$ 和 k , 存在 $C_1 = C_1(\lambda, \mathbf{A}), \lambda \in (|\lambda_2(\mathbf{A})|, 1)$, 使得:

$$\|[\mathbf{A}^k]_{ij} - \pi_j\| \leq C_1 \lambda^k, \|y_{i,k} - \pi_i\| \leq C_1 \lambda^k$$

其中, $\boldsymbol{\pi} = [\pi_1, \pi_2, \dots, \pi_N]^T$ 是 \mathbf{A} 的特征值 1 对应的归一化左 Perron 特征向量, $\mu^{-1} \leq y_{i,k} \leq 1, \boldsymbol{\pi}^T \mathbf{A} = \boldsymbol{\pi}^T, \boldsymbol{\pi}^T \mathbf{1}_N = 1$ 。引理 1 是 Perron 定理的推论, 在此不做具体证明。

该分布式次梯度算法为分布式在线优化问题提供了一种有效的解决方案。对于具有简单约束形式和凸目标函数的优化问题, 该算法能够有效求解。然而, 在面对具有复杂约束的优化问题时, 算法在某些方面仍需进行改进。

3.2 分布式对偶平均算法

为了求解有约束条件的优化问题, 本文引入分布式对偶平均算法^[21], 定义投影运算符 $\prod_{\Omega}(\mathbf{v}, \eta)$ 为:

$$\prod_{\Omega}(\mathbf{v}, \eta) = \arg \min_{\mathbf{u} \in \Omega} \left\{ \langle \mathbf{v}, \mathbf{u} \rangle + \frac{h(\mathbf{u})}{\eta} \right\}$$

其中, $\eta > 0, \mathbf{v} \in \mathbb{R}^m$, 近端函数 $h(\mathbf{u})$ 具有 1-强凸性质, 且满足对于所有的 $\mathbf{u} \in \Omega, h(\mathbf{0}) = 0, h(\mathbf{u}) > 0$ 。则对于问题(1)可以通过对偶平均算法求解^[21], 对 $\forall i \in \mathcal{V}$, 有:

$$\mathbf{z}_{i,k+1} = \mathbf{z}_{i,k} + \mathbf{g}_{i,k}, \mathbf{x}_{i,k+1} = \prod_{\Omega}(\mathbf{v}, \eta_k) \quad (4)$$

其中, $\mathbf{g}_{i,k} \in \partial f_{i,k}(\mathbf{x}_{i,k}), \eta_k > 0$ 是递减步长, $\mathbf{x}_{i,k+1}$ 通过将 $\mathbf{z}_{i,k+1}$ 映射到 Ω 上进行更新。对偶平均算法可以通过选择不同的近端函数 $h(\mathbf{u})$, 利用对偶变量 $\mathbf{z}_{i,k+1}$, 为解决非光滑带约束优化问题提供有效方案。然而该算法无法有效解决本文所讨论的有向网络上的隐私保护分布式在线优化问题。

3.3 状态分解分布式对偶平均算法

本文基于如式(4)所示的分布式对偶平均算法和文献[22]中的状态分解方法, 旨在解决由式(1)所代表的问题模型及隐私保护问题。在每次迭代中, 节点 i 的更新状态如下:

$$\begin{cases} \mathbf{z}_{i,k+1}^{\phi} = \sum_{j=1}^N a_{ij} \mathbf{z}_{j,k}^{\phi} + \alpha_{i\phi\varphi} (\mathbf{z}_{i,k}^{\varphi} - \mathbf{z}_{i,k}^{\phi}) \\ \mathbf{z}_{i,k+1}^{\varphi} = \mathbf{z}_{i,k}^{\varphi} + \alpha_{i\phi\varphi} (\mathbf{z}_{i,k}^{\phi} - \mathbf{z}_{i,k}^{\varphi}) + \frac{\mathbf{g}_{i,k}}{y_{ii,k}} \\ \mathbf{y}_{i,k+1} = \sum_{j=1}^N a_{ij} \mathbf{y}_{j,k} \\ \mathbf{x}_{i,k+1}^{\phi} = \prod_{\Omega}(\mathbf{z}_{i,k+1}^{\phi}, \eta_k) \\ \mathbf{x}_{i,k+1}^{\varphi} = \prod_{\Omega}(\mathbf{z}_{i,k+1}^{\varphi}, \eta_k) \end{cases} \quad (5)$$

其中, $\mathbf{g}_{i,k}$ 是 $f_{i,k}(\mathbf{x})$ 在 $\alpha_{i\phi\varphi} \mathbf{x}_{i,k}^{\phi} + (1 - \alpha_{i\phi\varphi}) \mathbf{x}_{i,k}^{\varphi}$ 的次梯度; $\alpha_{i\phi\varphi}$ 是节点 i 的状态 $\mathbf{z}_{i,k}^{\varphi}$ 和 $\mathbf{z}_{i,k}^{\phi}$ 之间的权重; $y_{ii,k}$ 是 $\mathbf{y}_{i,k}$ 的第 i 个元素, 是处理有向网络的不平衡性的估计变量, $\mathbf{y}_{i,k}$ 的引入是为了补偿有向图中节点权重分布不均造成的信息失衡, 通过动态追踪 Perron 向量, 并对梯度进行归一化, 使整个系统能正确收敛至全局最优点; $\eta_k > 0$ 为递减步长。状态分解分布式对偶平均算法的更新流程如算法 1 所示。

算法 1 状态分解分布式对偶平均算法

初始化输入: $\mathbf{z}_{i,1} \in \mathbb{R}^m, \mathbf{x}_{i,1}^{\phi} = \prod_{\Omega}(\mathbf{z}_{i,1}^{\phi}, \eta_0), \eta_0 > 0, \mathbf{x}_{i,1}^{\varphi} = \prod_{\Omega}(\mathbf{z}_{i,1}^{\varphi}, \eta_0), \mathbf{y}_{i,1} = \mathbf{e}_i \in \mathbb{R}^N$

设置: $k=1$

1. 当 $k \leq K$.
2. 对节点 $i \in \mathcal{V}$, 发送 $a_{ij} \mathbf{z}_{i,k}^{\phi}$ 和 $a_{ij} \mathbf{y}_{i,k}$ 到邻居集 $j \in \mathcal{N}^{\text{out}}$ 中的节点, 从邻居集 $j \in \mathcal{N}^{\text{in}}$ 中的节点接收 $a_{ij} \mathbf{z}_{j,k}^{\phi}$ 和 $a_{ij} \mathbf{y}_{j,k}$.
3. 更新 $\mathbf{z}_{i,k+1}^{\phi}, \mathbf{z}_{i,k+1}^{\varphi}, \mathbf{y}_{i,k+1}, \mathbf{x}_{i,k+1}^{\phi}$ 和 $\mathbf{x}_{i,k+1}^{\varphi}$,

$$\mathbf{z}_{i,k+1}^{\phi} = \sum_{j=1}^N a_{ij} \mathbf{z}_{j,k}^{\phi} + \alpha_{i\phi\varphi} (\mathbf{z}_{i,k}^{\varphi} - \mathbf{z}_{i,k}^{\phi}),$$

$$\mathbf{z}_{i,k+1}^{\varphi} = \mathbf{z}_{i,k}^{\varphi} + \alpha_{i\phi\varphi} (\mathbf{z}_{i,k}^{\phi} - \mathbf{z}_{i,k}^{\varphi}) + \frac{\mathbf{g}_{i,k}}{y_{ii,k}},$$

$$\mathbf{y}_{i,k+1} = \sum_{j=1}^N a_{ij} \mathbf{y}_{j,k},$$

$$\mathbf{x}_{i,k+1}^{\phi} = \prod_{\Omega}(\mathbf{z}_{i,k+1}^{\phi}, \eta_k),$$

$$\mathbf{x}_{i,k+1}^{\varphi} = \prod_{\Omega}(\mathbf{z}_{i,k+1}^{\varphi}, \eta_k).$$
4. 返回 $\mathbf{x}_{i,k+1}^{\phi}$ 和 $\mathbf{x}_{i,k+1}^{\varphi}$.

本文算法是一种基于多节点共识方法^[19]的优化算法。其核心思想是将每个节点的状态 $\mathbf{x}_{i,k}$ 分解为两个子状态: $\mathbf{x}_{i,k+1}^{\phi}$ 和 $\mathbf{x}_{i,k+1}^{\varphi}$ 。其中, 子状态 $\mathbf{x}_{i,k+1}^{\phi}$ 在节点间的交互中继承了原始状态 $\mathbf{x}_{i,k}$ 的作用, 实际上, 它是节点 i 唯一可观测的状态

值。而子状态 $\mathbf{x}_{i,k+1}^e$ 仅通过与 $\mathbf{x}_{i,k+1}^f$ 的交互参与算法的迭代。换句话说,子状态 $\mathbf{x}_{i,k+1}^e$ 的更新是在节点内部进行的,因此,对于节点 i 的相邻节点来说, $\mathbf{x}_{i,k+1}^e$ 是不可见的。例如,图 1(a) 展示了状态分解前的网络结构。以图 1(b) 中的节点 1 为例, $\mathbf{x}_{1,k+1}^f$ 在节点间交互中充当原始状态 \mathbf{x}_1 的角色,而 $\mathbf{x}_{1,k+1}^e$ 对节点 1 以外的其他节点是不可见的,尽管它会影响到 $\mathbf{x}_{1,k+1}^f$ 的更新。在文献[15]中已严格证明,在满足约束条件 $\mathbf{x}_{i,k+1}^e + \mathbf{x}_{i,k+1}^f = 2\mathbf{x}_{i,k+1}$ 以及 $0 < \zeta \leq \alpha_{i\phi} < 1$ 的前提下,分解后的两个状态的初始值可以从实数集中任意选取。因此,该算法的初始值无需额外约束。

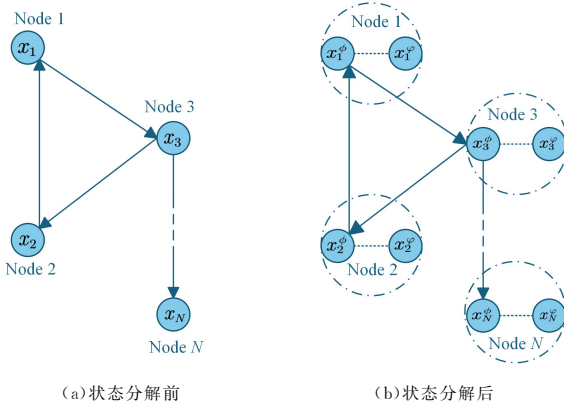


图 1 状态分解的示意图

Fig. 1 Schematic diagram of state decomposition

4 理论分析

在本章中,对算法 1 进行理论分析。

4.1 主要结果

令:

$$\mathbf{z}_k = \begin{pmatrix} \mathbf{z}_{1,k} \\ \mathbf{z}_{2,k} \\ \vdots \\ \mathbf{z}_{N,k} \\ \mathbf{z}_{N+1,k} \\ \vdots \\ \mathbf{z}_{2N,k} \end{pmatrix} = \begin{pmatrix} \mathbf{z}_{1,k}^f \\ \mathbf{z}_{2,k}^f \\ \vdots \\ \mathbf{z}_{N,k}^f \\ \mathbf{z}_{1,k}^e \\ \vdots \\ \mathbf{z}_{N,k}^e \end{pmatrix}, \mathbf{x}_k = \begin{pmatrix} \mathbf{x}_{1,k} \\ \mathbf{x}_{2,k} \\ \vdots \\ \mathbf{x}_{N,k} \\ \mathbf{x}_{N+1,k} \\ \vdots \\ \mathbf{x}_{2N,k} \end{pmatrix} = \begin{pmatrix} \mathbf{x}_{1,k}^f \\ \mathbf{x}_{2,k}^f \\ \vdots \\ \mathbf{x}_{N,k}^f \\ \mathbf{x}_{1,k}^e \\ \vdots \\ \mathbf{x}_{N,k}^e \end{pmatrix}$$

算法 1 可以改写为:

$$\begin{cases} \mathbf{z}_{i,k+1} = \sum_{j=1}^{2N} h_{ij} \mathbf{z}_{j,k} + \mathbf{s}_{i,k}, i=1, 2, \dots, 2N \\ \mathbf{y}_{i,k+1} = \sum_{j=1}^N a_{ij} \mathbf{y}_{j,k} \\ \mathbf{x}_{i,k+1} = \prod_{\Omega} (\mathbf{z}_{i,k+1}, \eta_k) \end{cases} \quad (6)$$

当 $i=1, 2, \dots, N, \mathbf{s}_{i,k}=0$; 当 $i=N+1, N+2, \dots, 2N,$

$\mathbf{s}_{i,k} = \mathbf{g}_{i-N,k} / \mathbf{y}_{i-Ni-N,k}$ 。新的权重矩阵 \mathbf{H} 如下:

$$\mathbf{H} = [h_{ij}]_{2N \times 2N} = \begin{pmatrix} \mathbf{A} - \mathbf{Q} & \mathbf{Q} \\ \mathbf{Q} & \mathbf{I} - \mathbf{Q} \end{pmatrix}$$

其中, $\mathbf{Q} = \text{diag}\{\alpha_{1\phi}, \alpha_{2\phi}, \dots, \alpha_{N\phi}\}$ 。由于网络结构 G 是强连通的并且权重矩阵为行随机矩阵,对于新的权重矩阵 \mathbf{H} 所对应的有向网络 \bar{G} ,可以通过选择合适的状态权重 $0 < \alpha_{i\phi} < a_{ii}, \forall i \in \mathcal{V}$,使得网络结构 \bar{G} 也是强连通的,并且其

权重矩阵 \mathbf{H} 也是行随机矩阵。

引理 2 在假设 1 下,令 $0 < \alpha_{i\phi} < a_{ii}$,对于任意 $i, j \in \bar{\mathcal{V}}, \bar{\mathcal{V}}$ 是 \bar{G} 的节点集,存在 $C_2 = C_2(\gamma, \mathbf{H})$ 和 $\gamma \in (|\lambda_2(\mathbf{H})|, 1)$,使得 $\|[\mathbf{H}^k]_{ij} - q_j\| \leq C_2 \gamma^k$ 。其中, $q = 1/2[\pi_1, \pi_2, \dots, \pi_N, \pi_1, \pi_2, \dots, \pi_N]^T$ 是 \mathbf{H} 的特征值 1 对应的归一化的左 Perron 特征向量。引理 2 是引理 1 的直接应用。

引理 3 由式(3)可得,对任意 $\mathbf{v}, \mathbf{u} \in \mathbb{R}^m$ 和递减步长 $\eta > 0$,有 $\|\prod_{\Omega}^h(\mathbf{v}, \eta) - \prod_{\Omega}^h(\mathbf{u}, \eta)\| \leq \eta \|\mathbf{v} - \mathbf{u}\|$ 。设 \mathcal{H}_{Ω}^h 是对任意元素 $\mathbf{v}, \mathbf{u} \in \mathbb{R}^m$ 和递减步长 $\eta > 0$ 的共轭函数,定义 $\mathcal{H}_{\Omega}^h(\mathbf{v}, \eta) = \sup_{\mathbf{u} \in \Omega} \{\langle \mathbf{v}, \mathbf{u} \rangle - h(\mathbf{u}) / \eta\}$,基于共轭函数 $\mathcal{H}_{\Omega}^h(\cdot)$, $\prod_{\Omega}^h(\cdot)$,下式成立:

$$\mathcal{H}_{\Omega}^h(-\mathbf{v} - \mathbf{u}, \eta) \leq \frac{\alpha}{2} \|\mathbf{u}\|^2 + \mathcal{H}_{\Omega}^h(-\mathbf{v}, \eta) - \langle \nabla \mathcal{H}_{\Omega}^h(-\mathbf{v}, \eta), \mathbf{u} \rangle, \nabla \mathcal{H}_{\Omega}^h(-\mathbf{v}, \eta) = \prod_{\Omega}^h(\mathbf{v}, \eta)$$

引理 3 的证明可参考文献[5]附录中的引理 1。

引理 4 在假设 1 和假设 2 下,定义加权平均变量 $\bar{\mathbf{z}}_{k+1} = \sum_{i=1}^{2N} q_i \mathbf{z}_{i,k+1}$,对式(6)重写后的算法中的 $\mathbf{z}_{i,k+1}$ 和 $\bar{\mathbf{z}}_{k+1}$,有:

$$\|\mathbf{z}_{i,k+1} - \bar{\mathbf{z}}_{k+1}\| \leq C_2 \sum_{j=1}^{2N} \|\mathbf{z}_{j,1}\| + \frac{\mu N L C_2}{1 - \gamma} + 2\mu L = \Gamma$$

证明:把式(6)中的 $\mathbf{z}_{i,k+1}$ 代入 $\bar{\mathbf{z}}_{k+1} = \sum_{i=1}^{2N} q_i \mathbf{z}_{i,k+1}$,可以得到:

$$\begin{aligned} \bar{\mathbf{z}}_{k+1} &= \sum_{i=1}^{2N} q_i \left[\sum_{j=1}^{2N} h_{ij} \mathbf{z}_{j,k} + \mathbf{s}_{i,k} \right] \\ &= \bar{\mathbf{z}}_k + \sum_{i=N+1}^{2N} q_i \frac{\mathbf{g}_{i-N,k}}{\mathbf{y}_{i-Ni-N,k}} \\ &= \bar{\mathbf{z}}_k + \frac{1}{2} \sum_{i=1}^N \pi_i \frac{\mathbf{g}_{i,k}}{\mathbf{y}_{ii,k}} \end{aligned} \quad (7)$$

式(7)左右两边对 k 加和,化简得:

$$\bar{\mathbf{z}}_{k+1} = \bar{\mathbf{z}}_1 + \sum_{k=1}^k \sum_{i=N+1}^{2N} q_i \frac{\mathbf{g}_{i-N,k}}{\mathbf{y}_{i-Ni-N,k}} \quad (8)$$

由式(6)可得:

$$\mathbf{z}_{i,k+1} = \sum_{j=1}^{2N} [\mathbf{H}^k]_{ij} \mathbf{z}_{j,1} + \frac{\mathbf{g}_{i,k}}{\mathbf{y}_{ii,k}} + \sum_{l=1}^{k-1} \sum_{j=N+1}^{2N} [\mathbf{H}^{k-l}]_{ij} \frac{\mathbf{g}_{i-N,l}}{\mathbf{y}_{j-Nj-N,l}} \quad (9)$$

由式(8)和式(9)可得:

$$\mathbf{z}_{i,k+1} - \bar{\mathbf{z}}_{k+1} = \bar{\omega}_1 + \bar{\omega}_2 \quad (10)$$

其中:

$$\begin{aligned} \bar{\omega}_1 &= \sum_{j=1}^{2N} ([\mathbf{H}^k]_{ij} - q_j) \mathbf{z}_{j,1} \\ \bar{\omega}_2 &= \sum_{l=1}^{k-1} \sum_{j=N+1}^{2N} ([\mathbf{H}^{k-l}]_{ij} - q_j) \frac{\mathbf{g}_{i-N,l}}{\mathbf{y}_{j-Nj-N,l}} - \\ &\quad \sum_{i=N+1}^{2N} q_i \frac{\mathbf{g}_{i-N,k}}{\mathbf{y}_{i-Ni-N,k}} + \frac{\mathbf{g}_{i,k}}{\mathbf{y}_{ii,k}} \end{aligned}$$

根据引理 2,可以得出:

$$\|\bar{\omega}_1\| = \left\| \sum_{j=1}^{2N} ([\mathbf{H}^k]_{ij} - q_j) \mathbf{z}_{j,1} \right\| \leq C_2 \gamma^k \sum_{j=1}^{2N} \|\mathbf{z}_{j,1}\|$$

此外,根据假设 2 和引理 1 可以得到:

$$\begin{aligned} \|\bar{\omega}_2\| &\leq \left\| \sum_{l=1}^{k-1} \sum_{j=N+1}^{2N} ([\mathbf{H}^{k-l}]_{ij} - q_j) \frac{\mathbf{g}_{i-N,l}}{\mathbf{y}_{j-Nj-N,l}} \right\| + \\ &\quad \left\| \sum_{i=N+1}^{2N} q_i \frac{\mathbf{g}_{i-N,k}}{\mathbf{y}_{i-Ni-N,k}} \right\| + \left\| \frac{\mathbf{g}_{i,k}}{\mathbf{y}_{ii,k}} \right\| \end{aligned}$$

$$\leq \mu NLC_2 \sum_{l=1}^{k-1} \gamma^{k-l} + 2\mu L$$

最后可得:

$$\begin{aligned} \|\mathbf{z}_{i,k+1} - \bar{\mathbf{z}}_{k+1}\| &\leq \|\bar{\omega}_1\| + \|\bar{\omega}_2\| \\ &\leq C_2 \gamma^k \sum_{j=1}^{2N} \|\mathbf{z}_{j,1}\| + 2\mu L + \\ &\quad \mu NLC_2 \sum_{l=1}^{k-1} \gamma^{k-l} \\ &\leq C_2 \sum_{j=1}^{2N} \|\mathbf{z}_{j,1}\| + \frac{\mu NLC_2}{1-\gamma} + 2\mu L = \Gamma \end{aligned} \quad (11)$$

定理 1 在假设 1 和假设 2 下,由算法更新的 $x_{i,k}$,递减

步长 $\eta_k = \eta_0 / \sqrt{k+1}$,变量 $\bar{\mathbf{x}}_{k+1} = \prod_{\Omega}^h(\bar{\mathbf{z}}_{k+1}, \eta_k)$,最优解 $\mathbf{x}^* =$

$\arg \min_{\mathbf{x} \in \Omega} \sum_{k=1}^K \sum_{i=1}^N f_{i,k}(\mathbf{x})$,存在 σ_1 和 σ_2 ,使得:

$$R_{j,k} \leq \sigma_1 \sqrt{K} + \sigma_2$$

其中:

$$\sigma_1 = \frac{2h(\mathbf{x}^*)}{\eta_0} + \frac{\mu^2 L^2 \eta_0}{4} + 6\eta_0 NLF$$

$$\sigma_2 = \mathcal{H}_{\Omega}^h(-\bar{\mathbf{z}}_1, \eta_0) + \langle \bar{\mathbf{z}}_1, \mathbf{x}^* \rangle + \frac{\mu NMLC_2}{1-\gamma}$$

证明:由式(2)动态遗憾定义可得:

$$\begin{aligned} &\sum_{k=1}^K \sum_{i=1}^N f_{i,k}(\mathbf{x}_{j,k}) - \sum_{k=1}^K \sum_{i=1}^N f_{i,k}(\mathbf{x}^*) \\ &\leq \sum_{k=1}^K \sum_{i=1}^N (f_{i,k}(\mathbf{x}_{j,k}) - f_{i,k}(\mathbf{x}_{i,k}) + \langle \mathbf{g}_{i,k}, \mathbf{x}_{i,k} - \bar{\mathbf{x}}_k \rangle) + \\ &\quad \sum_{k=1}^K \sum_{i=1}^N \langle \mathbf{g}_{i,k}, \bar{\mathbf{x}}_k - \mathbf{x}^* \rangle \end{aligned} \quad (12)$$

对 $f: \mathbb{R}^m \rightarrow \mathbb{R}$, $\forall \mathbf{v}, \mathbf{u} \in \mathbb{R}^m$,次梯度 $g(\mathbf{v}) \in \partial f(\mathbf{v}) \subseteq \mathbb{R}^m$,

则满足 $f(\mathbf{v}) - f(\mathbf{u}) \leq \langle g(\mathbf{v}), \mathbf{v} - \mathbf{u} \rangle$. 不等式放缩由此可得.

则从式(12)可得:

$$\sum_{k=1}^K \sum_{i=1}^N f_{i,k}(\mathbf{x}_{j,k}) - \sum_{k=1}^K \sum_{i=1}^N f_{i,k}(\mathbf{x}^*) \leq \theta_1 + \theta_2 \quad (13)$$

其中:

$$\begin{aligned} \theta_1 &= \sum_{k=1}^K \sum_{n=1}^N \langle \mathbf{g}_{i,k}, \bar{\mathbf{x}}_k - \mathbf{x}^* \rangle \\ \theta_2 &= \sum_{k=1}^K \sum_{i=1}^N (f_{i,k}(\mathbf{x}_{j,k}) - f_{i,k}(\mathbf{x}_{i,k})) + \sum_{k=1}^K \sum_{i=1}^N \langle \mathbf{g}_{i,k}, \mathbf{x}_{i,k} - \bar{\mathbf{x}}_k \rangle \end{aligned}$$

对于 θ_1 ,利用柯西施瓦茨不等式可得:

$$\begin{aligned} \theta_1 &= \sum_{k=1}^K \sum_{i=1}^N \langle \mathbf{g}_{i,k}, \bar{\mathbf{x}}_k - \mathbf{x}^* \rangle \\ &= \sum_{k=1}^K \sum_{i=N+1}^{2N} \left\langle \frac{\mathbf{g}_{i-N,k}}{\mathbf{y}_{i-Ni-N,k}} (\mathbf{y}_{i-Ni-N,k} - q_i), \bar{\mathbf{x}}_k - \mathbf{x}^* \right\rangle + \\ &\quad \sum_{k=1}^K \sum_{i=N+1}^{2N} \left\langle \frac{\mathbf{g}_{i-N,k}}{\mathbf{y}_{i-Ni-N,k}} q_i, \bar{\mathbf{x}}_k - \mathbf{x}^* \right\rangle \\ &\leq \sum_{k=1}^K \sum_{i=N+1}^{2N} \left\| \frac{(\mathbf{y}_{i-Ni-N,k} - q_i) \mathbf{g}_{i-N,k}}{\mathbf{y}_{i-Ni-N,k}} (\bar{\mathbf{x}}_k - \mathbf{x}^*) \right\| + \\ &\quad \sum_{k=1}^K \sum_{i=N+1}^{2N} \left\langle \frac{\mathbf{g}_{i-N,k}}{\mathbf{y}_{i-Ni-N,k}} q_i, \bar{\mathbf{x}}_k - \mathbf{x}^* \right\rangle \end{aligned} \quad (14)$$

由假设 2 可得 $\|\bar{\mathbf{x}}_k - \mathbf{x}^*\| \leq M$,则由式(14)的第一部分可得:

$$\begin{aligned} &\sum_{k=1}^K \sum_{i=1}^{2N} \left\| \frac{(\mathbf{y}_{i-Ni-N,k} - q_i) (\bar{\mathbf{x}}_k - \mathbf{x}^*) \mathbf{g}_{i-N,k}}{\mathbf{y}_{i-Ni-N,k}} \right\| \\ &\leq \sum_{k=1}^K \sum_{i=N+1}^{2N} \mu LC_2 \gamma^k M \leq \frac{\mu NMLC_2}{1-\gamma} \end{aligned} \quad (15)$$

从式(8)可得到:

$$\sum_{k=1}^K \sum_{i=N+1}^{2N} \left\langle \frac{q_i \mathbf{g}_{i-N,k}}{\mathbf{y}_{i-Ni-N,k}}, \mathbf{x}^* \right\rangle = \langle \bar{\mathbf{z}}_{K+1}, \mathbf{x}^* \rangle - \langle \bar{\mathbf{z}}_1, \mathbf{x}^* \rangle \quad (16)$$

因为 $\mathcal{H}_{\Omega}^h(-\bar{\mathbf{z}}_{k+1}, \eta_k) \leq \mathcal{H}_{\Omega}^h(-\bar{\mathbf{z}}_{k+1}, \eta_{k-1})$,结合引理 4,得:

$$\begin{aligned} \mathcal{H}_{\Omega}^h(-\bar{\mathbf{z}}_{k+1}, \eta_k) &\leq \frac{\eta_{k-1}}{2} \left\| \sum_{i=N+1}^{2N} \frac{\mathbf{g}_{i-N,k}}{\mathbf{y}_{i-Ni-N,k}} q_i \right\|^2 - \\ &\quad \left\langle \bar{\mathbf{x}}_k, \sum_{i=N+1}^{2N} \frac{\mathbf{g}_{i-N,k}}{\mathbf{y}_{i-Ni-N,k}} q_i \right\rangle + \\ &\quad \mathcal{H}_{\Omega}^h(-\bar{\mathbf{z}}_k, \eta_{k-1}) \end{aligned} \quad (17)$$

对式(17)关于 k 求加和:

$$\begin{aligned} &\sum_{k=1}^K \left\langle \bar{\mathbf{x}}_k, \sum_{i=N+1}^{2N} \frac{\mathbf{g}_{i-N,k}}{\mathbf{y}_{i-Ni-N,k}} q_i \right\rangle \\ &\leq \mathcal{H}_{\Omega}^h(-\bar{\mathbf{z}}_1, \eta_0) - \mathcal{H}_{\Omega}^h(-\bar{\mathbf{z}}_{K+1}, \eta_K) + \\ &\quad \sum_{k=1}^K \frac{\eta_{k-1}}{2} \left\| \sum_{i=N+1}^{2N} \frac{\mathbf{g}_{i-N,k}}{\mathbf{y}_{i-Ni-N,k}} q_i \right\|^2 \end{aligned} \quad (18)$$

联立式(16)和式(18)可得到:

$$\begin{aligned} &\sum_{k=1}^K \sum_{i=N+1}^{2N} \left\langle \frac{\mathbf{g}_{i-N,k}}{\mathbf{y}_{i-Ni-N,k}} q_i, \bar{\mathbf{x}}_k - \mathbf{x}^* \right\rangle \\ &\leq \mathcal{H}_{\Omega}^h(-\bar{\mathbf{z}}_1, \eta_0) - \mathcal{H}_{\Omega}^h(-\bar{\mathbf{z}}_{K+1}, \eta_K) + \langle \bar{\mathbf{z}}_1, \mathbf{x}^* \rangle + \\ &\quad \sum_{k=1}^K \frac{\eta_{k-1}}{2} \left\| \sum_{i=N+1}^{2N} \frac{\mathbf{g}_{i-N,k}}{\mathbf{y}_{i-Ni-N,k}} q_i \right\|^2 - \langle \bar{\mathbf{z}}_{K+1}, \mathbf{x}^* \rangle \\ &\leq \mathcal{H}_{\Omega}^h(-\bar{\mathbf{z}}_1, \eta_0) + \langle \bar{\mathbf{z}}_1, \mathbf{x}^* \rangle + \langle \bar{\mathbf{z}}_{K+1}, \bar{\mathbf{x}}_{K+1} - \mathbf{x}^* \rangle + \\ &\quad \frac{h(\bar{\mathbf{x}}_{K+1})}{\eta_K} + \frac{\mu^2 L^2}{4} \sum_{k=1}^K \frac{\eta_{k-1}}{2} \end{aligned} \quad (19)$$

令 $p(\mathbf{x}) = \langle \bar{\mathbf{z}}_{K+1}, \mathbf{x} \rangle + h(\mathbf{x})/\eta_K$,由 $h(\mathbf{x})$ 是 1-强凸函数,得

$p(\mathbf{x})$ 是 $1/\eta_K$ 强凸函数;联合 $\bar{\mathbf{x}}_{K+1} = \prod_{\Omega}^h(\bar{\mathbf{z}}_{K+1}, \eta_K)$ 可得 $p(\bar{\mathbf{x}}_{K+1}) - p(\mathbf{x}^*) \leq 0$,则:

$$\langle \bar{\mathbf{z}}_{K+1}, \bar{\mathbf{x}}_{K+1} - \mathbf{x}^* \rangle + \frac{h(\bar{\mathbf{x}}_{K+1})}{\eta_K} \leq \frac{h(\mathbf{x}^*)}{\eta_K} \quad (20)$$

将式(19)和式(20)代入式(14),即可推出:

$$\begin{aligned} \theta_1 &\leq \frac{h(\mathbf{x}^*)}{\eta_K} + \frac{\mu^2 L^2}{4} \sum_{k=1}^K \frac{\eta_{k-1}}{2} + \langle \bar{\mathbf{z}}_1, \mathbf{x}^* \rangle + \\ &\quad \mathcal{H}_{\Omega}^h(-\bar{\mathbf{z}}_1, \eta_0) + \frac{\mu NMLC_2}{1-\gamma} \\ &\leq \left(\frac{2h(\mathbf{x}^*)}{\eta_0} + \frac{\mu^2 L^2 \eta_0}{4} \right) \sqrt{K} + \langle \bar{\mathbf{z}}_1, \mathbf{x}^* \rangle + \\ &\quad \mathcal{H}_{\Omega}^h(-\bar{\mathbf{z}}_1, \eta_0) + \frac{\mu NMLC_2}{1-\gamma} \end{aligned} \quad (21)$$

其中, $\eta_k = \eta_0 / \sqrt{k+1}$,可得 $1/\eta_K \leq 2\sqrt{K}/\eta_0$,则 $\sum_{k=1}^K \eta_{k-1} = \sum_{k=1}^K \eta_0 / \sqrt{k} \leq \eta_0 (2\sqrt{K} - 2) \leq 2\eta_0 \sqrt{K}$. 另一方面,对 θ_2 在假设 2 可推出:

$$\begin{aligned} \theta_2 &\leq \sum_{k=1}^K \sum_{i=1}^N [L \|\mathbf{x}_{j,k} - \mathbf{x}_{i,k}\| + L \|\mathbf{x}_{i,k} - \bar{\mathbf{x}}_k\|] \\ &\leq \sum_{k=1}^K 3NL \max \|\mathbf{x}_{i,k} - \bar{\mathbf{x}}_k\| \end{aligned} \quad (22)$$

对于式(22),结合式(6)和引理 4 可推出:

$$\begin{aligned} &\sum_{k=1}^K 3NL \max \|\mathbf{x}_{i,k} - \bar{\mathbf{x}}_k\| \\ &= 3NL \sum_{k=1}^K \max_{i \in V} \|\Pi_{\Omega}^h(\mathbf{z}_{i,k}, \eta_{k-1}) - \Pi_{\Omega}^h(\bar{\mathbf{z}}_k, \eta_{k-1})\| \\ &\leq 3NLF \sum_{k=1}^K \eta_{k-1} \leq 6\eta_0 NLF \sqrt{K} \end{aligned} \quad (23)$$

将式(21)一式(23)代入式(12),得到遗憾:

$$R_{j,k} \leq \sigma_1 \sqrt{K} + \sigma_2$$

其中:

$$\sigma_1 = \frac{2h(\mathbf{x}^*)}{\eta_0} + \frac{\alpha^2 L^2 \eta_0}{4} + 6\eta_0 NL\Gamma$$

$$\sigma_2 = \mathcal{H}_\Omega^h(-\bar{\mathbf{z}}_1, \eta_0) + \langle \bar{\mathbf{z}}_1, \mathbf{x}^* \rangle + \frac{\mu NMLC_2}{1-\gamma}$$

4.2 隐私证明

本节主要证明算法 1 能够实现隐私保护。在分布式网络的隐私保护研究中,通常考虑 3 种攻击模式:诚实但好奇的节点是指严格遵循算法规则的节点,但会出于好奇收集接收到的数据,并试图推测其他节点的私有信息;外部窃听者是指能够获知网络拓扑结构,并通过窃听通信链路获取交换信息的攻击者;恶意攻击者是指完全不遵循算法规则,可能恶意篡改自身状态,试图获取其他节点的私有信息。在信息泄露的风险方面,恶意攻击者构成的威胁最大,其次是窃听者,而诚实但好奇的节点的威胁相对最小。本文主要针对外部窃听者进行分析。

在实际应用中,节点通常不希望其他节点获知其成本函数相关的信息,因此,节点的私有信息主要体现在目标函数中。基于这一点,本文将目标函数的梯度信息作为私有信息进行保护。对于窃听者而言,无论其无法获取梯度信息,还是只能获得梯度的近似值,都表明该算法能够在保护梯度信息的意义上实现隐私保护。假设窃听者已知节点间的通信拓扑结构以及算法 1 的更新规则,并且能够在某个时间范围 K 内获取相应信息,则可得以下结果:

$$\begin{cases} \mathbf{z}_{i,2}^\phi + \mathbf{z}_{i,2}^\psi = \sum_{j=1}^N a_{ij} \mathbf{z}_{j,1}^\phi + \mathbf{z}_{i,1}^\psi + \frac{\mathbf{g}_{i,1}}{y_{ii,1}} \\ \mathbf{z}_{i,3}^\phi + \mathbf{z}_{i,3}^\psi = \sum_{j=1}^N a_{ij} \mathbf{z}_{j,2}^\phi + \mathbf{z}_{i,2}^\psi + \frac{\mathbf{g}_{i,2}}{y_{ii,2}} \\ \vdots \\ \mathbf{z}_{i,K}^\phi + \mathbf{z}_{i,K}^\psi = \sum_{j=1}^N a_{ij} \mathbf{z}_{j,K-1}^\phi + \mathbf{z}_{i,K-1}^\psi + \frac{\mathbf{g}_{i,K-1}}{y_{ii,K-1}} \end{cases} \quad (24)$$

对窃听者而言,根据算法更新规则,可知变量 $\mathbf{z}_{i,k}^\phi, \mathbf{x}_{i,k+1}^\phi =$

$\prod_{\Omega} (\mathbf{z}_{i,k+1}^\phi, \eta_k)$ 和 $y_{ii,k} (k=1, 2, \dots, K-1)$ 是可以被窃听和计算的;然而变量 $\mathbf{z}_{i,k}^\psi, \mathbf{x}_{i,k+1}^\psi = \prod_{\Omega} (\mathbf{z}_{i,k+1}^\psi, \eta_k)$ 和 $\mathbf{g}_{i,k} (k=1, 2, \dots, K-1)$ 对外部窃听者而言是未知的。这是因为 $\mathbf{z}_{i,k}^\psi$ 的更新过程仅在节点 i 内部进行,并不与其他节点发生交互。在方程组 (24) 中,共有 $2K-1$ 个未知数,而仅有 $K-1$ 个方程,显然无法唯一确定每个未知变量的值。即使窃听者可以利用方程组 (24) 得到近似解,但由于目标函数 f_i 依赖于精确的梯度信息,而不是近似值,该算法仍能够有效保护节点 i 的梯度信息。

在文献[23-25]中,差分隐私技术被用于实现隐私保护,其基本原理是在数据中加入服从特定分布的噪声,从而使邻居节点或外部窃听者无法准确恢复原始数据。然而,这种方法在提升隐私保护能力的同时,也不可避免地降低了优化精度,因而需在隐私保护与计算准确性之间做出权衡。在文献[26-28]中,研究者通过数据加密技术设计了隐私保护型分布式优化算法,其核心思想是节点在每次信息传输和接收时都需要进行加密和解密操作。然而,这种方法显著增加了计算开销。相比之下,本文提出的算法仅利用节点之间的拓扑关

系,并通过额外的简单计算实现隐私保护。因此该方法在保证隐私性的同时,也避免了复杂的加密计算和噪声干扰。然而,本文仅针对外部窃听者攻击模式下的隐私保护优化问题进行研究。在实际应用中,差分隐私技术和数据加密技术更为成熟,并已广泛应用于多个领域,如联邦学习^[29-31]和边缘计算^[31]等。

5 模拟实验

本章通过多个仿真实验,验证了所提算法的有效性。对于后续实验中网络结构的选取,本文遵循相同的操作流程。首先,生成了一个高度稀疏的随机有向网络 G ,该网络由 n 个传感器节点组成,并确保其具备较强的连通性。随后,采用随机均匀抽样的方法,在每个时刻 k 以概率 40% 为 G 中的每条边分配权重,从而构造一个新的有向网络 G_k 。这一过程确保了在整个时间区间内,有向网络 G_k 始终保持强连通性,并且由于节点的入度和出度不相等,该方法生成的权重矩阵满足非平衡性的要求。

实验环境方面,第一个实验及第二个实验的第二部分在 Python 3.12 环境下运行,而其余实验均在 MATLAB R2023b 中完成。

5.1 目标跟踪

动态目标跟踪问题是指多个传感器相互协作,对多个移动目标进行跟踪。在实际应用中,由传感器组成的通信网络往往是不均衡的,这与本文所提出的算法研究背景相一致。在本实验中,假设通信网络由 5 个节点组成,这些节点之间的连接关系是有向且随时间变化的,同时保证整个网络始终保持连通性。所有节点通过协同工作,共同跟踪 3 个移动目标。在时间步 k 时,移动目标的位置可以表示为 $\mathbf{c}_{i,k}^* = [\mathbf{c}_{i,k}^*(1), \mathbf{c}_{i,k}^*(2)] \in \mathbb{R}^2, l = \{1, 2, 3\}$ 。具体来说:

$$\mathbf{c}_{i,k}^* = \begin{bmatrix} \mathbf{c}_{i,k}^*(1) \\ \mathbf{c}_{i,k}^*(2) \end{bmatrix} = \begin{bmatrix} \hat{\Lambda}_{i,k}^* \\ \hat{\Lambda}_{i,k}^* \end{bmatrix} = \begin{bmatrix} v_l \sin(\gamma_l k + \rho_l) \\ \gamma_l v_l \cos(\gamma_l k + \rho_l) \end{bmatrix}$$

其中, $\mathbf{c}_{i,k}^*(1)$ 和 $\mathbf{c}_{i,k}^*(2)$ 分别表示目标 l 在时间 k 时的位置和速度, ρ_l 表示目标 l 的相位, γ_l 表示角频率, v_l 表示振幅。数字 1 和数字 2 分别代表指数,对应该向量在两个不同维度上的取值。设 $\mathbf{c}_k^* = [(\mathbf{c}_{1,k}^*)^T, (\mathbf{c}_{2,k}^*)^T, (\mathbf{c}_{3,k}^*)^T]^T$, 并随机生成测量矩阵 $\omega^i \in \mathbb{R}^{1 \times 6}$ 。随后,每个节点使用测量值 $\phi_k^i = \omega^i \mathbf{c}_k^*$ 来跟踪移动目标。当 $\mathbf{x} \in [-2, 2]^6$ 时,目标跟踪问题可以表示为优化以下目标函数: $f_k(\mathbf{x}) = \sum_{i=1}^5 \|\omega^i \mathbf{x} - \phi_k^i\| / 2$ 。

在模拟实验中,随机变量 ρ_l 和 v_l 分别在区间 $[0, \pi]$ 和 $[0, 2]$ 内服从均匀分布。采样频率设为 100 Hz, 递减步长设定为 $\eta_k = 4 / \sqrt{k+1}$ 。图 2 展示了移动目标在不同时间步 K 的运动轨迹,其中 $\mathbf{c}_{1,k}^*(1), \mathbf{c}_{2,k}^*(1)$, 和 $\mathbf{c}_{3,k}^*(1)$ 表示时间步 K 中移动目标的位置。图中还给出了节点 1 在时间步 k 时,第一、第三和第五维度的状态值 $\mathbf{x}_k^1(1), \mathbf{x}_k^1(3), \mathbf{x}_k^1(5)$ 。从图 2 可以看出,节点 1 的跟踪轨迹与 3 个移动目标的实际运动轨迹基本吻合,说明算法 1 能够有效实现目标跟踪。此外,从图 3 可观察到,算法 1 和文献[17]的算法在动态遗憾方面均呈现次线性趋势,这进一步验证了算法 1 的有效性。

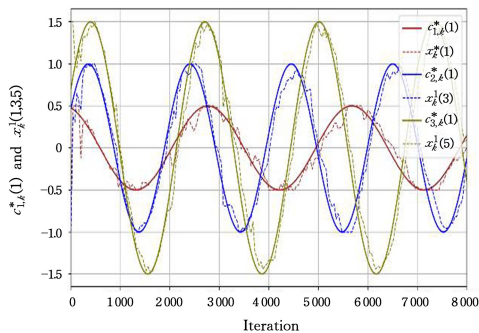


图2 算法1下的目标跟踪

Fig. 2 Target tracking under algorithm 1

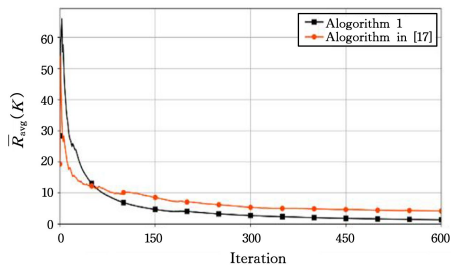


图3 算法1和文献[17]中的算法在目标跟踪下的遗憾随K的变化

Fig. 3 Regret of algorithm 1 and the algorithm in reference [17]

under target tracking varies with K

5.2 动态稀疏信号恢复

在本例中,我们的目标是利用算法1解决动态稀疏信号恢复问题。所有传感器的主要任务是在时间区间 K 内协同搜索向量,以最小化全局损失函数:

$$f(\mathbf{x}) = \sum_{k=1}^K \sum_{i=1}^N \frac{1}{2} \|q_{i,k} - \mathbf{P}_{i,k} \mathbf{x}\|^2 + \rho \|\mathbf{x}\|_1, \mathbf{x} \in \Omega$$

其中, $\mathbf{P}_{i,k} \in \mathbb{R}^{d \times m}$ 为观测矩阵, ρ 为常数正则化参数, $q_{i,k} = \mathbf{P}_{i,k} \bar{\omega}_k + \boldsymbol{\beta}_{i,k}$ 表示测量值, 递减步长为 $\eta_k = 1/4 \sqrt{k+1}$, $\boldsymbol{\beta}_{i,k} \in [-1, 1] \times [-1, 1]$ 为噪声, $\bar{\omega}_k \in [0, 1] \times [0, 1]$ 为动态稀疏信号, 矩阵 $\mathbf{P}_{i,k} = \mathbf{I}_2, i = 1, \dots, 20, \Omega \subseteq [-2.5, 2.5] \times [-2.5, 2.5]$ 。

1) 比较: 本文对文献[17]的算法与算法1进行了性能比较。在解决动态稀疏信号恢复问题时, 算法1具有较优的表现。从图4可以看出, 随着时间 K 的增加, 算法1的动态遗憾逐渐收敛。与文献[17]中的对偶平均算法类似, 算法1同样实现了次线性收敛速度。实验结果不仅验证了算法1的有效性, 也为其理论分析提供了有力支持。

2) 不同的隐私保护水平: 本节探讨了文献[17]的算法在不同隐私保护水平下的遗憾随时间 K 的变化情况。首先简单描述差分隐私的概念: 对算法 \mathcal{B} , $\epsilon > 0$ 是隐私预算, 隐私预算越小, 隐私保护能力越强。对于任意一对相邻数据集 U 和 U' (即它们最多相差一个样本), 对 $\forall S \subseteq \text{Range}(\mathcal{B})$, 有 $P[\mathcal{B}(U) \in S] \leq e^\epsilon \cdot P[\mathcal{B}(U') \in S]$, 其中 S 表示事件, P 表示概率。则称算法 \mathcal{B} 满足 ϵ -差分隐私。这里, 通过添加高斯噪声 $e_{i,k} \sim \mathcal{L}(0, b_k)$ 以达到差分隐私。具体的差分隐私相关概念详见文献[17]。其中, 用尺度参数 b_k 表示隐私保护水平, b_k 值越大, 隐私预算越小, 算法的隐私保护能力越强。从图5可以观察到, 随着隐私级别的提高, 遗憾值增大, 相应的

收敛精度降低。这表明, 在增强隐私保护的同时, 算法的收敛性能可能会受到一定影响。因此, 在采用差分隐私技术进行隐私保护时, 需要在算法收敛性与隐私保护水平之间权衡和优化。

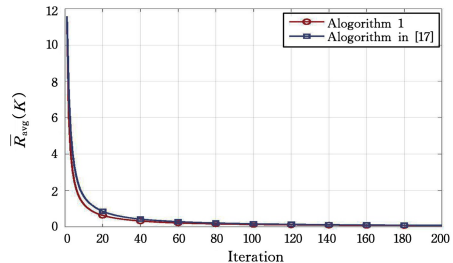


图4 算法1和文献[17]中的算法在动态稀疏信号恢复任务下的遗憾随K的变化

Fig. 4 Regret of algorithm 1 and the algorithm in reference [17] in the dynamic sparse signal recovery varies with K

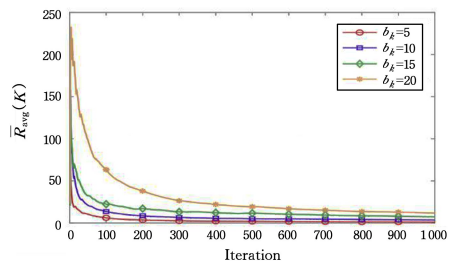


图5 在文献[17]的算法下不同级别隐私保护

Fig. 5 Different levels of privacy protection under the algorithm proposed in reference [17]

5.3 逻辑回归模型进行线性分类

本节探讨了一个受文献[23]启发的线性分类问题, 设计了3个实验以系统地评估所提出算法的可行性与隐私保护能力。具体而言, 针对特定特征空间中移动的一组数据点, 本文考虑由 $n=50$ 个节点组成的分布式系统, 这些节点协同训练一个线性分类器。时间步 $K > 1$, 每个节点 i 在样本 $\{a_{i,t}, \mathbf{c}_k^{i,t}\}_{t \in \{1, 2, \dots, v_i\}}$ 上进行训练, 其中 $a_{i,t} \in \{-1, 1\}$ 为样本标签, $\mathbf{c}_k^{i,t} \in \mathbb{R}^2$ 为节点位置。我们的目标是寻找一个分离超平面 $\{c \in \mathbb{R}^2 \mid \mathbf{y}^T c + h = 0\}$, 由参数 $(\mathbf{y}, h) \in \Omega_1 \times \Omega_2, \Omega_1 \in \mathbb{R}^2, \Omega_2 \in \mathbb{R}$ 描述, 以最优方式区分标签为 $a^{i,t} = 1$ 和 $a^{i,t} = -1$ 的数据点。分类规则如下: 对于所有样本 (i, t) , 若 $(\mathbf{y}_k^*)^T \mathbf{c}_k^{i,t} + h_k^* \geq 0$, 则样本标签为 $a^{i,t} = 1$; 否则, 其标签为 $a^{i,t} = -1$ 。由此, 该线性分类问题可表示为如下优化问题:

$$\min_{\mathbf{y}, h} \sum_{i=1}^n \sum_{t=1}^{v_i} \log(1 + e^{-a^{i,t}(\mathbf{y}^T \mathbf{c}_k^{i,t} + h)}) + \frac{S(\|\mathbf{y}\|^2 + h^2)}{2}$$

其中, $S=1$ 为正则化参数。在实验中, 每个数据点 $\mathbf{c}_k^{i,t}$ 沿着圆周轨迹移动, 其轨迹由 $\mathbf{c}_k^{i,t} = \mathbf{c}_p^{i,t} + r \begin{bmatrix} \cos(k/100) \\ \sin(k/100) \end{bmatrix}$ 描述, 其中 $r=1$ 为圆的半径, $\mathbf{c}_p^{i,t}$ 为随机生成的圆心位置。实验参数设定为 $\eta_k = 20 \sqrt{1/(k+1)}, b_k = 5$ 。本节的重点在于分析算法1与文献[17]中算法在有限时间跨度 $K=200$ 内的收敛性能。实验结果如图6所示, 结论与图4一致, 即算法1适用时变有向网络, 并具有渐近的次线性收敛速度。

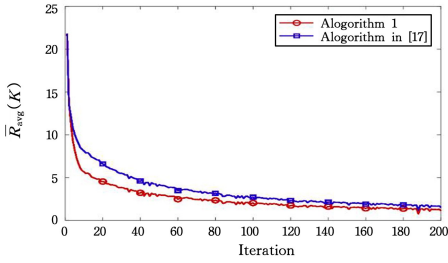


图6 算法1和文献[17]中的算法在线性分类任务下的遗憾随K的变化

Fig. 6 Regret of algorithm 1 and the algorithm in reference [17] in the line classification varies with K

1)机器学习:在前述实验中,已经验证了算法的收敛性。为了进一步验证算法1在大规模机器学习任务中的有效性,采用文献[33]中MNIST数据集训练分布式逻辑回归分类模型。通过与现有分布式算法的性能进行对比,进一步验证算法1在分类任务中具有良好的准确性。如图7所示,该数据集包含13007个样本,其中随机选取12000个样本用于训练,其余1007个用于测试。每个样本被转换为一个维度 $d=784$ 的向量,以供模型训练。图8展示了算法1及文献[34-36]的方法在训练过程中残差(即期望平方误差)的变化趋势。残差随迭代次数的变化可以直观反映算法的收敛速度。虽然4种算法的收敛速率有所不同,但随着迭代次数的增加,残差最终均趋于稳定,表明它们均能在较少的迭代次数内逼近最优解,从而验证了算法1的收敛性。此外,图9展示了4种算法的分类准确率。从实验结果可以直观地评估各算法的分类能力。可以观察到,随着迭代进行,所有算法的分类准确率最终均趋于稳定,并接近1,这表明4种算法均能在该任务中实现较高的分类精度。

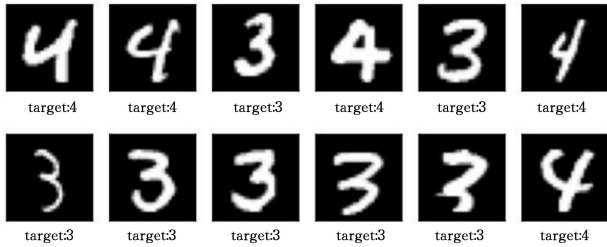


图7 数据集的样本图

Fig. 7 Samples of the dataset

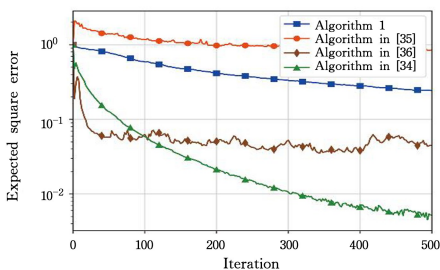


图8 算法1和文献[34-36]中3个相关的算法使用MNIST数据集的训练结果

Fig. 8 Training results of algorithm 1 and three other related algorithms in reference [34-36] using MNIST dataset

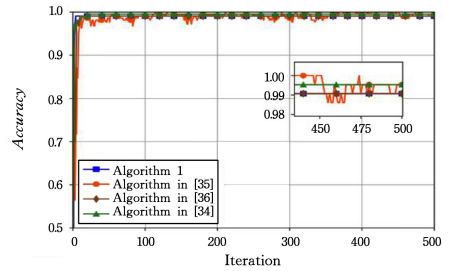


图9 算法1和文献[34-36]3个相关的算法使用MNIST数据集的测试结果

Fig. 9 Testing results of algorithm 1 and three other related algorithms in reference [34-36] using MNIST dataset

2)隐私保护:本节进一步评估了算法1在隐私保护方面的能力。具体地,构建了一个包含窃听者的攻击场景,并且考虑最恶劣的隐私风险,即外部窃听者不仅了解网络结构和通用更新规则,还能够获取网络内共享的信息。本文假设代价函数保持不变,且算法1采用确定性梯度进行更新。将算法1与式(4)所示的分布式对偶平均算法进行对比分析。通过比较两种算法在该威胁模型下的性能表现,验证算法1在抵御信息泄露方面的有效性,从而说明其具备较强的隐私保护能力。图10展示了窃听者iDLG^[37]在式(4)所示的分布式对偶平均算法更新规则下的推理结果。实验表明,在未采取任何隐私保护措施的情况下,外部窃听者能够轻松推导出私有数据,从而威胁系统安全性。相比之下,图11展示了窃听者在算法1约束下的推理结果,表明其无法准确恢复原始数据。这一结果证明了算法1在隐私保护方面的有效性。

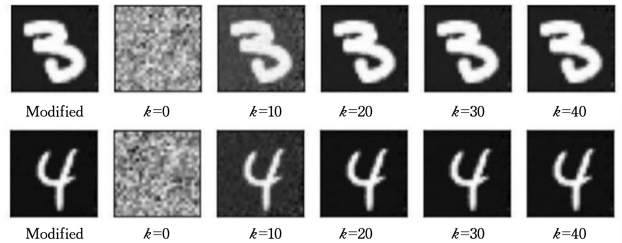


图10 式(4)在文献[37]窃听下的推理结果

Fig. 10 Inference results of equation(4) under the method in reference [37]

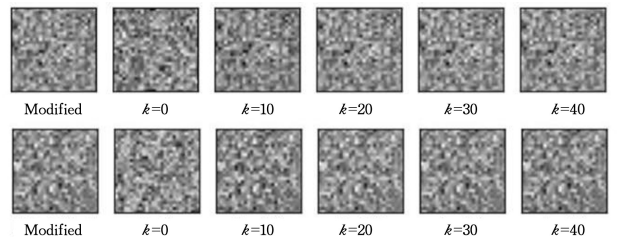


图11 算法1下的推理结果

Fig. 11 Inference results under algorithm 1

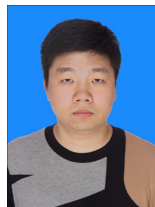
结束语 本文提出了一种高效的状态分解分布式对偶平均算法,适用于时变有向网络环境,并可用于解决涉及隐私保护的在线约束优化问题。该算法基于对偶平均方法求解约束优化问题,并结合状态分解技术以增强隐私保护能力。具体

而言,所提算法将每个节点的状态变量进行分解,其中一个状态变量仅在节点内部进行更新,不与外部通信,而另一个状态变量则用于与相邻节点交换信息。此外,这些变量之间的权重是时变的,并且仅节点本身知晓其具体数值,从而进一步增强了隐私保护机制。理论分析证明,该算法可以实现次线性遗憾,同时有效保护各个节点的损失函数隐私。仿真实验进一步验证了该算法的有效性和实际可行性。

参 考 文 献

- [1] LIU Q S, LE X Y, LI K X, et al. A distributed optimization algorithm based on multiagent network for economic dispatch with region partitioning [J]. *IEEE Transactions on Cybernetics*, 2021, 51(5): 2466-2475.
- [2] YAN J J, CAO J, CAO Y. Distributed continuous-time algorithm for economic dispatch problem over switching communication topology[J]. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2021, 68(6): 2002-2006.
- [3] XU J M, ZHU S Y, SOH Y C, et al. Convergence of asynchronous distributed gradient methods over stochastic networks [J]. *IEEE Transactions on Automatic Control*, 2018, 63(2): 434-448.
- [4] YUAN D M, HONG Y G, HO D W C, et al. Optimal distributed stochastic mirror descent for strongly convex optimization[J]. *Automatica*, 2018, 90: 196-203.
- [5] LIANG S, WANG L Y, YIN G, et al. Distributed quasimonotone subgradient algorithm for nonsmooth convex optimization over directed graphs[J]. *Automatica*, 2019, 101: 175-181.
- [6] NOTARNICOLA I, NOTARSTEFANO G. Asynchronous distributed optimization via randomized dual proximal gradient[J]. *IEEE Transactions on Automatic Control*, 2017, 62(5): 2095-2106.
- [7] XU J M, ZHU S Y, SOH Y C, et al. A bregman splitting scheme for distributed optimization over networks[J]. *IEEE Transactions on Automatic Control*, 2018, 63(11): 3809-3824.
- [8] LI X X, XIE L H, HONG Y G, et al. Distributed aggregative optimization over multi-agent networks[J]. *IEEE Transactions on Automatic Control*, 2022, 67(6): 3165-3171.
- [9] CHENG S S, LIANG S, HONG Y G, et al. Distributed stochastic algorithm for convex optimization over directed graphs[C] // 2019 Chinese Control And Decision Conference(CCDC). IEEE, 2019: 101-106.
- [10] LI Q, LIAO Y X, WU K M, et al. Parallel and distributed optimization method with constraint decomposition for energy management of microgrids[J]. *IEEE Transactions on Smart Grid*, 2021, 12(6): 4627-4640.
- [11] JIANG X, ZENG X L, SUN J, et al. Distributed stochastic gradient tracking algorithm with variance reduction for non-convex optimization[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2023, 34: 5310-5321.
- [12] DAI P C, YU W W, WANG H, et al. Distributed Actor-Critic Algorithms for Multiagent Reinforcement Learning Over Directed Graphs[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2022, 34(10): 7210-7221.
- [13] WU J Y, TIAN Y P. Online Distributed Newton Step Algorithm for Multi-agent Optimization Over General Unbalanced Networks[C] // 2024 IEEE 18th International Conference on Control & Automation(ICCA). IEEE, 2024: 174-179.
- [14] ZHANG J H, MA D. A Novel State Decomposition-Based Privacy-Preserving Algorithm for Distributed Optimization Over Directed Networks[C] // 2024 14th Asian Control Conference(ASCC). IEEE, 2024: 1145-1150.
- [15] MAKRIDIS E, CHARALAMBOUS T. A Linear Push-Pull Average Consensus Algorithm for Delay-Prone Networks[C] // 2024 European Control Conference (ECC). IEEE, 2024: 743-749.
- [16] MAI V S, ABED E H. Distributed optimization over directed graphs with row stochasticity and constraint regularity[J]. *Automatica*, 2019, 102: 94-104.
- [17] LYU Q G, ZHANG K K, DENG S J, et al. Privacy-preserving decentralized dual averaging for online optimization over directed networks[J]. *IEEE Transactions on Industrial Cyber-Physical Systems*, 2023, 1: 79-91.
- [18] WANG H J, LIU K, HAN D Y, et al. Privacy-preserving distributed online stochastic optimization with time-varying distributions[J]. *IEEE Transactions on Control of Network Systems*, 2023, 10(2): 1069-1082.
- [19] WANG Y Q. Privacy-preserving average consensus via state decomposition [J]. *IEEE Transactions on Automatic Control*, 2019, 64(11): 4711-4716.
- [20] LI J, ZHU X, WU Z, et al. Online distributed dual averaging algorithm for multi-agent bandit optimization over time-varying general directed networks[J]. *Information Sciences*, 2021, 581: 678-693.
- [21] DUCHI J C, AGARWAL A, WAINWRIGHT M J. Dual averaging for distributed optimization: convergence analysis and network scaling [J]. *IEEE Transactions on Automatic Control*, 2012, 57(3): 592-606.
- [22] YAN J J, CAO J D. Privacy preservation of optimization algorithm over unbalanced directed graph[J]. *IEEE Transactions on Network Science and Engineering*, 2022, 9(4): 2164-2173.
- [23] CARNEVALE G, FARINA F, NOTARNICOLA L, et al. GT-Adam: gradient tracking with adaptive momentum for distributed online optimization [J]. *IEEE Transactions on Control of Network Systems*, 2023, 10(3): 1436-1448.
- [24] HUANG L Y, WU J F, SHI D W, et al. Differential privacy in distributed optimization with gradient tracking[J]. *IEEE Transactions on Automatic Control*, 2024, 69(9): 5727-5742.
- [25] DING T, ZHU S Y, HE J P, et al. Consensus-based distributed optimization in multi-agent systems: convergence and differential privacy[C] // 2018 IEEE Conference on Decision and Control. IEEE, 2018: 3409-3414.
- [26] BU X Y, DONG T. Differential privacy optimal consensus for multiagent system by using functional perturbation[C] // 2019 6th International Conference on Information, Cybernetics, and Computational Social Systems. IEEE, 2019: 157-162.

- [27] ZHANG C, WANG Y. Enabling privacy-preservation indecentralized optimization[J]. IEEE Transactions on Control of Network Systems, 2019, 6(2): 679-689.
- [28] LU Y, ZHU M. Privacy preserving distributed optimization using homomorphic encryption[J]. Automatica, 2018, 96: 314-325.
- [29] ZANG H R, YANG T T, LIU H B, et al. Study on cryptographic verification of distributed federated learning for internet of things[J]. Computer Science, 2024, 51(S1): 1062-1066.
- [30] SUN M, DING X N, CHENG Q. Federated learning scheme based on differential privacy [J]. Computer Science, 2024, 51(S1): 912-917.
- [31] TANG L T, WANG D, ZHANG L F, et al. Federated learning scheme based on secure multi-party computation and differential privacy[J]. Computer Science, 2022, 49(9): 297-305.
- [32] SUN J M, ZHAO M X. Survey of application of differential privacy in edge computing[J]. Computer Science, 2024, 51(S1): 896-904.
- [33] HAZAN E. Introduction to online convex optimization [J]. Foundations and Trends in Optimization, 2016, 2(3/4): 157-325.
- [34] XIONG Y Y, LI X, YOU K Y, et al. Distributed online optimization in time-varying unbalanced networks without explicit subgradients[J]. IEEE Transactions on Signal Processing, 2022, 70: 4047-4060.
- [35] HAN D Y, LIU K, LIN Y M, et al. Differentially private distributed online learning over time-varying digraphs via dual averaging[J]. International Journal of Robust and Nonlinear Control, 2022, 32(5): 2485-2499.
- [36] LIANG S, WANG L Y, YIN G. Dual averaging push for distributed convex optimization over time-varying directed graph[J]. IEEE Transactions on Automatic Control, 2020, 65(4): 1785-1791.
- [37] HE X Y, ZHENG Z B, CHEN Z F, et al. Adaptive evolution strategies for stochastic zeroth-order optimization [J]. IEEE Transactions on Emerging Topics in Computational Intelligence, 2022, 6(5): 1271-1285.



DAI Xiangguang, born in 1986, Ph.D., professor. His main research interests include optimization algorithms, neural networks, clustering and pattern recognition.



ZHANG Wei, born in 1970, Ph.D., professor. His main research interests include information security and computational intelligence.

(责任编辑:喻藜)