A Distributed Online Reconstruction Algorithm for Spatio-temporal Signals Postprint

Authors: Chi Yuan, Jiang Junzheng, Jiang Junzheng

Date: 2023-02-14T00:00:00+00:00

Abstract

The online reconstruction problem of spatio-temporal signals can be reduced to the recovery of difference-smoothed time-varying graph signals. For this convex optimization problem, existing distributed reconstruction algorithms based on gradient descent exhibit extremely slow convergence when the condition number of the Hessian matrix is large, resulting in significant reconstruction error when the maximum number of iterations is limited within a single observation interval. To address this issue, we propose a distributed online reconstruction algorithm based on the approximate Newton method. First, the original optimization problem is decomposed into a series of local optimization problems on subgraphs through subgraph partitioning, and the solutions to these local problems are obtained. Then, fusion averaging is performed on the local solutions across subgraphs to obtain an approximate global optimal solution. Subsequently, based on the gap between the approximate solution and the actual optimal solution, we prove that the subgraph partitioning and fusion matrix obtained in this manner exhibit sparsity and can serve as an approximate Hessian inverse matrix for the original optimization problem. Finally, this approximate matrix is substituted into the classical Newton method iteration formula, and the structured sparsity of this approximate matrix is utilized to achieve distributed computation. Simulation results demonstrate that, compared with existing algorithms, the proposed algorithm achieves faster convergence, smaller reconstruction error, and requires less communication overhead.

Full Text

A Distributed Online Reconstruction Algorithm for Spatio-Temporal Signals

Guilin University of Electronic Technology Journal, Vol. 43, No. 1, Feb. 2023

School of Information and Communication, Guilin University of Electronic Technology, Guangxi

Abstract

The online reconstruction problem for spatio-temporal signals can be formulated as the recovery of differentially smooth time-varying graph signals. For this convex optimization problem, existing distributed reconstruction algorithms based on gradient descent methods exhibit extremely slow convergence when the Hessian matrix of the optimization problem has a large condition number. When the maximum number of iterations is limited within a single observation interval, the reconstruction error becomes substantial. To address this issue, we propose a distributed online reconstruction algorithm based on an approximate Newton' s method. First, the original optimization problem is decomposed into a series of local optimization problems on subgraphs through subgraph partitioning, and solutions to these local problems are obtained. Then, fusion averaging is performed on the local solutions between subgraphs to obtain an approximate global optimum. By analyzing the gap between this approximate solution and the actual optimal solution, we prove that the subgraph partitioning and fusion matrices obtained in this manner are sparse and can serve as an approximate inverse Hessian matrix for the original optimization problem. Finally, this approximate matrix is substituted into the classical Newton' s method iteration formula. Simulation results demonstrate that the proposed algorithm achieves faster convergence speed, smaller reconstruction error, and requires less communication compared with existing algorithms.

Keywords: spatio-temporal signals; online reconstruction; distributed algorithm; approximate Newton' s method; subgraph decomposition

1. Introduction

In today' s information age, human society is inundated with massive spatio-temporal signals, such as traffic flow data in transportation networks, bioelectrical signals in neuronal networks, temperature or humidity data in wireless sensor networks, and personal preferences in social media [1-2]. As science and technology advance, these signals are gradually evolving toward larger scales and more complex structures. Real-world spatio-temporal signals are often missing or unreliable due to factors like energy constraints, noise interference, and environmental changes [3-4]. For instance, in wireless sensor network monitoring tasks, sensor nodes have limited memory and processing capabilities and cannot continuously observe data. Consequently, the observation data across the entire network may be incomplete [3-4]. Additionally, sensor nodes may produce anomalous monitoring data due to noise interference or drastic environmental changes [5], resulting in missing reliable signal values at these nodes.

When spatio-temporal signals suffer from data missing, it is necessary to exploit the correlation characteristics of these signals for reconstruction to ensure subsequent signal processing tasks can proceed normally [6-7]. Graph signal processing theory extends traditional signal processing concepts such as Fourier analysis and convolution to irregular graph domains, providing a powerful tool for studying signals on non-regular domains [8-9]. The correlation of spatio-temporal signals in both spatial and temporal domains can generally be described by the differential smoothness of time-varying graph signals [10-11]. Differential smoothness refers to the property that a signal and its variations are similar between adjacent vertices on a graph. For example, in a temperature sensor network, geographically closer nodes typically exhibit more similar temperature values and variation trends. The online reconstruction model for differentially smooth time-varying graph signals can effectively reconstruct missing spatio-temporal signals. However, existing distributed algorithms for solving this model are based on gradient descent methods [10], which suffer from relatively slow convergence and are sensitive to the condition number of the optimization problem' s Hessian matrix [12-13]. Slow convergence leads to significantly increased communication within distributed systems [14-15]. Therefore, there is a need to design a new distributed algorithm with faster and more stable convergence.

To address the slow convergence issue of existing algorithms [10], this paper proposes a distributed online reconstruction algorithm for spatio-temporal signals based on approximate Newton' s method. The principle involves first decomposing the optimization problem into a series of local problems on subgraphs through subgraph partitioning to obtain local solutions, then performing fusion averaging of local solutions between subgraphs to obtain an approximate global optimum. By analyzing the gap between the approximate solution and the actual optimal solution, we prove that the decomposition and fusion matrices obtained in this manner are sparse and can serve as an approximate inverse Hessian matrix for the original optimization problem. Finally, this approximate matrix is substituted into the classical Newton' s method iteration formula. Simulation results demonstrate that the proposed algorithm achieves faster convergence, smaller reconstruction error, and lower communication requirements compared with existing algorithms.

---

2. Spatio-Temporal Signal Online Reconstruction Model

2.1 Graph Signal Processing Fundamentals

Any spatially distributed N-node network can be modeled as a graph. A graph G is a non-linear data structure consisting of a set of vertices and edges connecting them, denoted as G = (V, E, W), where V is the vertex set, E is the edge set, and W ^(N×N) is the weighted adjacency matrix. The weighted adjacency matrix completely describes the topology of a graph, with its weights satisfying the condition $W_{ij} \geq 0$. The stronger the correlation or similarity between vertices i and j on the graph, the larger the edge weight W_{ij} if vertices i and j are connected.

A graph signal can be defined as a mapping from the vertex set V of graph G to the real number set . If the vertices of the graph are ordered in some fashion, the graph signal can be denoted as an N-dimensional vector x, where x_i represents the signal value at vertex i. A graph signal that varies over time during a period T is called a time-varying graph signal and can be represented as a matrix composed of a set of graph signal vectors at each time instant.

The smoothness of a graph signal refers to the property that the signal values differ only slightly between adjacent vertices on the graph. The smoothness of a graph signal can be measured by the Laplacian quadratic form $x^T L x$, where L is the graph Laplacian matrix. Based on the definition of the Laplacian matrix, the Laplacian quadratic form can be rewritten as $x^T L x = \Sigma_{(i,j)} E W_{ij} (x_i - x_j)^2$. The smaller the value of the Laplacian quadratic form, the smaller the difference in signal values between vertices connected by edges with larger weights, indicating a smoother graph signal. Conversely, a larger value indicates a less smooth graph signal.

2.2 Differentially Smooth Graph Signal Online Reconstruction Model

The incomplete spatio-temporal signal can be abstracted as a time-varying graph signal sampled over a period T. The sampling observation model for time-varying graph signals is given by:

$y_t = S_t x_t + n_t$, (1)

where y_t is the actual observed value of the graph signal at time t, x_t is the true value of the signal, n_t is independent and identically distributed observation noise for $t = 1, 2, \cdots, T$, and S_t is a diagonal sampling matrix. The diagonal elements of S_t are $s_{t,i}$ $\{0,1\}$, where $s_{t,i} = 1$ if vertex i is sampled at time t, and 0 otherwise. When the diagonal elements of S_t contain zeros, it indicates that the observed data has missing values.

For smooth graph signals with missing data, a commonly used smooth graph signal online reconstruction model can be employed to recover the incomplete signal:

$\min_{(x_t)} ||S_t x_t - y_t||_2^2 + \quad x_t^T L x_t$, (4)

where the first term is a data fidelity term that forces the reconstructed signal to be close to the true values at the sampling points, and the second term is a smoothness regularization term that enhances the smoothness of the reconstructed signal. However, this model only considers the smoothness of graph signals in the vertex domain of the graph and does not fully exploit the correlation of signals in the temporal domain, resulting in larger reconstruction errors [16].

In reality, spatio-temporal signals exhibit smoothness not only in the spatial domain but also in the temporal domain. Therefore, researchers have proposed a differentially smooth graph signal online reconstruction model that achieves better reconstruction performance for spatio-temporal signals:

min_(x_t) ||S_t x_t - y_t||*2^2 + (x_t - x̃*(t-1))^T L (x_t - x̃_(t-1)), (5)

where x̃*(t-1) is the reconstructed graph signal from the previous time instant. According to this model, the current time signal can be reconstructed using the differential smoothness property of time-varying graph signals. When the graph signal is smooth and we let x̃*(t-1) = x_(t-1), the reconstruction model degenerates to a form that considers the smoothness across both time and vertex domains. Therefore, this online reconstruction model fully accounts for the smoothness of graph signals in the joint time-vertex domain.

Although this online reconstruction model has the advantages of low computational complexity and can be solved in a distributed manner [10], its solution algorithm is a first-order method that is susceptible to the condition number of the Hessian matrix of the optimization problem. To address this limitation, this paper designs an approximate Newton' s method-based algorithm for solving the optimization model.

---

3. Distributed Online Reconstruction Algorithm

3.1 Distributed Algorithm Based on Gradient Descent

The distributed algorithm based on gradient descent for solving the optimization problem (5) has a relatively slow convergence speed. The gradient g_t(x_t) at iteration k is given by:

g_t(x_t^(k)) = 2(S_t + L) x_t^(k) - 2 y_t - 2 L x̃_(t-1), (6)

where k is the iteration number. Since the graph Laplacian matrix L is a local matrix with the same sparsity pattern as the graph, the computation of the gradient can be implemented in a distributed fashion. Specifically, the i-th element of the gradient vector [g_t(x_t)]_i is:

[g_t(x_t)]*i = 2 s_t,i x_t,i - 2 y_t,i + 2 Σ*(j N_i) W_{ij} (x_t,i - x_t,j), (9)

where N_i is the set of first-order neighbors of vertex i. To compute [g_t(x_t)]_i, each vertex only needs to exchange information once with its first-order neighbors.

3.2 Approximate Newton' s Method

Newton's method is a second-order algorithm with faster convergence that is less sensitive to the condition number of the Hessian matrix [17], which meets the convergence speed requirements for online reconstruction algorithms. However, Newton' s method requires matrix inversion. For the optimization problem (5), the Hessian matrix is:

$$H_t = 2(S_t + L). \quad (7)$$

Matrix inversion is a centralized operation. For the entire graph, inverting a matrix of graph-scale size cannot be computed in a distributed manner. Moreover, the computational complexity of matrix inversion is high. To maintain the convergence speed advantage of second-order algorithms while enabling distributed computation, we seek an approximate matrix P_t with structured sparsity to replace the inverse Hessian matrix H_t^{-1}, yielding the approximate Newton iteration:

$$\hat{x}_t = x_t - P_t g_t(x_t). \quad (10)$$

3.3 Hessian Inverse Approximation via Subgraph Decomposition

To obtain the approximate inverse Hessian matrix P_t for the optimization problem, we partition graph G into a series of overlapping subgraphs centered at each vertex i. Let V_i be the set of r-hop neighborhood vertices of vertex i, and E_i be the set of edges connecting these vertices, forming subgraph $G_i = (V_i, E_i)$. Through this graph partitioning approach, the global optimization problem on graph G can be decomposed into a series of sub-optimization problems on subgraphs G_i:

$$\min_{x_t} \|R_i (S_t + L) x_t - R_i y_t - R_i L \tilde{x}_{t-1}\|_2^2, \quad (13)$$

where R_i is a diagonal matrix representing a local operation. The j-th diagonal element of R_i equals 1 if vertex j of graph G is within subgraph G_i, and 0 otherwise. The objective function of the sub-optimization problem depends only on the optimization variables within the local operation $R_i x_t$, without requiring information outside the subgraph. This sub-optimization problem is a local optimization problem whose local optimal solution can be obtained from data within the subgraph:

$$\hat{x}_{t,i} = (R_i (S_t + L) R_i)^{\dagger} R_i (y_t + L \tilde{x}_{t-1}), \quad (14)$$

where † denotes the pseudo-inverse. Since matrix R_i has non-zero rows and columns only for vertices within subgraph G_i, the computation of the pseudo-inverse of this N-order matrix can be performed locally within the subgraph, with computational complexity equivalent to inverting a small matrix of order $|V_i|$.

The local solution obtained directly through (14) suffers from boundary effects due to information loss within each subgraph G_i. To compensate for this missing information across subgraphs, we perform information fusion on the local solutions:

$$\hat{x}_t = \Sigma_i w_i R_i \hat{x}_{t,i}, \quad (15)$$

where $w_i = 1/|V_i|$ is the fusion weight, and $\hat{x}_t$ is the approximate solution to the original optimization problem after subgraph fusion. The parameter r, which controls the size of the subgraph fusion, is typically chosen as a small integer less than N to effectively reduce the scale of subgraph fusion and avoid boundary effects [18].

Combining the subgraph partitioning and fusion processes, we can define matrix P_t as:

$$P_t = \Sigma_j w_j R_j (R_j (S_t + L) R_j)^{\dagger} R_j. \quad (16)$$

Then the approximate solution $\hat{x}_t$ can be compactly written as:

$$\hat{x}_t = P_t (y_t + L \tilde{x}_{(t-1)}). \quad (17)$$

Comparing with the optimal solution of the original problem:

$$x_t^* = (S_t + L)^{(-1)} (y_t + L \tilde{x}_{(t-1)}), \quad (18)$$

we find that the difference between the approximate solution $\hat{x}_t$ and the optimal solution x_t^* originates from the difference between the approximate matrix P_t and the inverse Hessian matrix $H_t^{(-1)}$. The error bound is given by:

$$||\hat{x}_t - x_t^*||_2 \leq ||I - P_t H_t||_2 \, ||x_t^*||_2. \quad (20)$$

It can be proven that there always exists a sufficiently large parameter r such that the relative error between the approximate solution $\hat{x}_t$ and the optimal solution x_t^* is less than . As the parameter r increases, the value of $||I - P_t H_t||_2$ gradually decreases, meaning that the approximate solution $\hat{x}_t$ becomes increasingly accurate [19-20].

3.4 Distributed Implementation

By selecting appropriate parameters (r, w_i) and substituting matrix P_t into the Newton iteration formula (10), we obtain the approximate Newton iteration:

$$x_t^{(k+1)} = x_t^{(k)} - P_t g_t(x_t^{(k)}). \quad (21)$$

Due to the sparsity of matrix P_t, the algorithm can be implemented in a distributed manner. The distributed computational flow is as follows:

1. Initialize: $x_t^{(0)} = \tilde{x}_{(t-1)}$
2. For each vertex i, exchange information with its r-hop neighbors to compute the local gradient $[g_t(x_t^{(k)})]_i$
3. Exchange information with r-hop neighbors to compute the approximate Newton step: $[d_t^{(k)}]_i = -[P_t g_t(x_t^{(k)})]_i$
4. Update local estimate: $[x_t^{(k+1)}]_i = [x_t^{(k)}]_i + [d_t^{(k)}]_i$
5. Repeat until convergence criterion is met: $||x_t^{(k+1)} - x_t^{(k)}||_2 / ||x_t^{(k)}||_2 <$

In each iteration of the distributed approximate Newton algorithm, each vertex only needs to communicate twice with its r-hop neighbors (once for gradient computation and once for Newton step computation), making the entire process distributed.

---

4. Simulation Results and Analysis

We compare the performance of the proposed distributed approximate Newton algorithm (referred to as "this algorithm") with existing algorithms in spatio-temporal signal online reconstruction tasks. The evaluation metrics are set as follows:

- **Relative error** for single-time reconstruction: $_t = ||\hat{x}_t - x_t||_2 / ||x_t||_2$ (22)
- **Cumulative reconstruction error** over the entire time period: $_{total} = (1/(T N_o)) \Sigma_t ||S_t (\hat{x}_t - x_t)||_2$ (23)

where N_o is the number of observed vertices. The iteration termination condition for all algorithms is set as: $||x_t^{(k)} - x_t^{(k-1)}||_2 / ||x_t^{(k-1)}||_2 < 10^{-3}$. (24)

4.1 Synthetic Dataset Experiment

Using the Graph Signal Processing Toolbox (GSPbox) [21], we construct a random sensor graph with $N = 100$ vertices to simulate a real-world wireless sensor network. The topology is shown in [FIGURE 1]. The time-varying graph signal is generated as follows: first, a low-frequency graph signal with energy of 100 is randomly generated as the initial signal at time $t = 1$. Then, subsequent graph signals are generated through $x_{(t+1)} = x_t + L^{(-1/2)} \delta_t$, where δ_t is a zero-mean Gaussian white noise vector with energy of 0.01, and the perturbation energy is controlled to ensure the synthetic signals satisfy differential smoothness.

Observation noise n_t is added to the synthetic graph signals, with each vertex's noise being independent and identically distributed Gaussian white noise with variance 0.01. The sampling method is set as independent and identically distributed random sampling across time with a sampling probability of 0.5, simulating a scenario where actual spatio-temporal observation data is partially missing.

Performance with sufficient maximum iterations: When the maximum number of iterations is sufficient ($K = 200$), the relationship between relative error and iteration number for a single time instant is shown in [FIGURE 2]. Due to the relatively small condition number of the Hessian matrix in this experiment (72.3), the gradient descent algorithm maintains a stable convergence speed, but it is still slower than the proposed algorithm. The proposed algorithm demonstrates faster convergence, reaching the same error precision with significantly less communication cost.

Performance with limited maximum iterations: When the maximum number of iterations is restricted, the relationship between cumulative reconstruction error over the entire time period and maximum communication per observation interval is shown in [FIGURE 3]. When K is small, the gradient descent algorithm cannot obtain accurate reconstructed signals within each time instant due to its slow convergence, resulting in large cumulative error. In contrast, the distributed approximate Newton algorithm has minimal error even when K = 5, demonstrating that it has less stringent requirements on the maximum number of iterations per observation interval and lower communication demands.

4.2 Sea Surface Temperature Dataset Experiment

To further validate the distributed approximate Newton algorithm, we compare its performance with existing algorithms on Pacific sea surface temperature data reconstruction. The dataset covers the Pacific region from 180°W to 120°W and 15°S to 15°N, with a topology shown in [FIGURE 4]. In this experiment, the condition number of the Hessian matrix is larger (523.6), causing the convergence speed of the gradient descent algorithm to become extremely slow and significantly increasing communication requirements to achieve target reconstruction precision.

Performance with sufficient maximum iterations: The relationship between relative error and iteration number is shown in [FIGURE 5]. The distributed approximate Newton algorithm maintains a fast convergence speed even with the large condition number.

Performance with limited maximum iterations: When the maximum iteration count is restricted, the comparison is shown in [FIGURE 6]. The gradient descent algorithm fails to achieve effective reconstruction within limited iterations, while the distributed approximate Newton algorithm still demonstrates good reconstruction performance.

---

5. Conclusion

To address the issue of slow convergence in existing distributed online reconstruction algorithms for spatio-temporal signals, this paper proposes a distributed online reconstruction algorithm based on approximate Newton' s method. The algorithm obtains local solutions to the graph signal reconstruction optimization problem through subgraph partitioning, then performs fusion calculations to obtain an approximate global optimal solution, yielding a sparse approximate inverse Hessian matrix. This approximate matrix is then substituted into the

Newton' s method iteration formula to derive the distributed approximate Newton method. Simulation results show that the distributed approximate Newton algorithm achieves faster convergence, lower communication requirements, and is less sensitive to the condition number of the optimization problem, making it more suitable for distributed online reconstruction tasks of spatio-temporal signals. This study focuses on algorithmic improvements based on existing optimization models. Future work will consider designing new online reconstruction models for spatio-temporal signals to further reduce reconstruction error.

References

[1] ORTEGA A, FROSSARD P, KOVACEVIC J. Graph signal processing: overview, challenges, and applications[J]. Proceedings of the IEEE, 2018, 106(5): 808-828.

[2] SANDRYHAILA A, MOURA J M F. Big data analysis with signal processing on graphs: representation and processing of massive data sets with irregular structure[J]. IEEE Signal Processing Magazine, 2014, 31(5): 80-90.

[3] YICK J, MUKHERJEE B, GHOSAL D. Wireless sensor network survey[J]. Computer Networks, 2008, 52(12): 2292-2330.

[4] LORENZO P D, BANELLI P, BARBAROSSA S. Distributed adaptive learning of graph signals[J]. IEEE Transactions on Signal Processing, 2017, 65(16): 4193-4208.

[5] 图信号处理理论与算法研究 [D]. 桂林电子科技大学, 2019: 1-4.

[6] 基于图信号处理的数据异常检测与恢复 [D]. 北京邮电大学, 2018: 9-12.

[7] 图信号采样与重建研究 [D]. 哈尔滨工业大学, 2017: 1-4.

[8] SHUMAN D I, NARANG S K, FROSSARD P, et al. Signal processing on graphs: extending high-dimensional data analysis to networks and other irregular domains[J]. IEEE Signal Processing Magazine, 2013, 30(3): 83-98.

[9] SANDRYHAILA A, MOURA J M F. Discrete signal processing on graphs[J]. IEEE Transactions on Signal Processing, 2013, 61(7): 1644-1656.

[10] QIU K, MAO X, SHEN J. Time-varying graph signal reconstruction[J]. IEEE Journal of Selected Topics in Signal Processing, 2017, 11(6): 870-883.

[11] MAO X, QIU K, SHEN J. Spatio-temporal signal reconstruction based on differential smoothness[J]. IEEE Transactions on Signal Processing, 2018, 66(23): 6281-6296.

[12] BOYD S, VANDENBERGHE L. Convex optimization[M]. Cambridge: Cambridge University Press, 2004: 41-47.

[13] JAKOVETIC D, XAVIER J, MOURA J M F. Fast distributed gradient methods[J]. IEEE Transactions on Automatic Control, 2014, 59(5): 1131-1146.

[14] MOKHTARI A, LING Q, RIBEIRO A. Network Newton distributed optimization methods[J]. IEEE Transactions on Signal Processing, 2017, 65(1): 146-161.

[15] EISEN M, MOKHTARI A, RIBEIRO A. Decentralized Quasi-Newton methods[J]. IEEE Transactions on Signal Processing, 2017, 65(10): 2613-2628.

[16] NARANG S K, SANOU E, GADDE A, et al. Localized iterative methods for interpolation in graph structured data[C]//IEEE Global Conference on Signal and Information Processing. Piscataway, NJ: IEEE Press, 2013: 491-494.

[17] NOCEDAL J, WRIGHT S J. Numerical optimization[M]. New York: Springer Science & Business Media, 2006: 41-47.

[18] JIANG Z, CHENG C, SUN Y. Nonsubsampled graph filter banks: theory and distributed algorithms[J]. IEEE Transactions on Signal Processing, 2019, 67(15): 3938-3953.

[19] JIANG Z, TAY W P. Decentralised signal processing on graphs via matrix inverse approximation[J]. Signal Processing, 2019, 165(22): 292-302.

[20] CHENG C, JIANG Z, SUN Y. Spatially distributed sampling and reconstruction[J]. Applied and Computational Harmonic Analysis, 2019, 47(1): 109-148.

[21] PERRAUDIN N, SHUMAN D I, PARATTE G. GSPBOX: a toolbox for graph signal processing[EB/OL]. (2016-03-15)[2021-03-20]. https://arxiv.org/pdf/1408.5781.pdf.

Note: Figure translations are in progress. See original paper for figures.