Recovery of Missing Sensor Data by Reconstructing Time-varying Graph Signals

Anindya Mondal, Mayukhmali Das, Aditi Chaterjee, P. Venkateswaran*

Department of Electronics and Telecommunication Engineering

Jadavpur University, India

anindyam.jan@gmail.com*

যাদবপুর বিশ্ববিদ্যালয়
JADAVPUR UNIVERSITY

Project Page

Paper

30th EUROPEAN SIGNAL PROCESSING CONFERENCE EUSIPCO BELGRADE/SERBIA 29 AUG - 2 SEP 2022

Objective

Graph Signal Processing (GSP) is a widely used graph-based method, with multiple applications in biological & social networks, computer vision, IoT and in other domains.

Wireless Sensor Networks (WSN) are a network of sensors deployed for observing physical phenomenons like temperature, humidity, air quality and others.

GSP finds its most promising application in WSNs, where a graph node represents the relative positions of the sensors & a graph signal represents the reading of the sensors.

Sensors are often located in remote locations. If any of them get damaged due to some reason, replacing them becomes difficult. Instead, reconstructing the missing data is easier.

In our work, we explore a recently developed method, based on the minimization of Sobolev norm for GSP to reconstruct missing values in WSNs.

Proposed Approach

G = (ν,ε,**W**) is an weighted undirected graph with weight **W** and x:ε→R is a vertex-indexed graph signal.

Let $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_M]^T$ be a time-varying graph signal, where $\mathbf{x}_t$ denotes the signal at time t (1 < t < M) in G . Here each row of **X** represents a time-series on the corresponding vertex.

Qiu et al. defined the smoothness function as: $S_2(\mathbf{X}) = \sum_{t=1}^{M} S_2(\mathbf{x}_t) = \mathrm{tr}\,(\mathbf{X}^{\mathrm{T}}\mathbf{L}\mathbf{X})$

To include the temporal information, we also define the temporal difference operator as:

$$\mathbf{D}_h = \begin{bmatrix} -1 & & & \\ 1 & -1 & & \\ & 1 & \ddots & \\ & & \ddots & -1 \\ & & & 1 \end{bmatrix}$$

and the temporal difference signal as: $\mathbf{X}\mathbf{D}_h = [\mathbf{x}_2 - \mathbf{x}_1,\ \mathbf{x}_3 - \mathbf{x}_2,\ \ldots,\ \mathbf{x}_M - \mathbf{x}_{M-1}]$

Qiu et al. had formulated the reconstruction problem as:

$$\min_{\overline{\mathbf{X}}} \frac{1}{2}\|\mathbf{J} \circ \overline{\mathbf{X}} - \mathbf{Y}\|_F^2 + \frac{\gamma}{2}\mathbf{tr}((\overline{\mathbf{X}}\mathbf{D}_h)^T\mathbf{L}\overline{\mathbf{X}}\mathbf{D}_h)$$

where **Y** is the sampled matrix, **J** is the sampling matrix. The above expression reconstructs a time-varying graph signal $\overline{\mathbf{X}}$with a small error $\|\mathbf{J} \circ \overline{\mathbf{X}} - \mathbf{Y}\|_F^2$while minimizing the temporal difference graph signal smoothness $\mathbf{tr}((\overline{\mathbf{X}}\mathbf{D}_h)^T\mathbf{L}\overline{\mathbf{X}}\mathbf{D}_h)$

In 2020, Giraldo and Bouwmans proposed an alternative approach for time-varying graph signals reconstruction inspired by the minimization of the Sobolev norm. They formulated the problem as: $\min_{\overline{\mathbf{X}}} \frac{1}{2}\|\mathbf{J} \circ \overline{\mathbf{X}} - \mathbf{Y}\|_F^2 + \frac{\gamma}{2}\mathbf{tr}((\overline{\mathbf{X}}\mathbf{D}_h)^{\mathbf{T}}(\mathbf{L}+\epsilon\mathbf{I})^{\beta}\overline{\mathbf{X}}\mathbf{D}_h)$

Experiments

We test the proposed approach on two widely used datasets: the Intel Lab Dataset and the Mol’ene Dataset.

We represent the sensors as the nodes of a graph and the Euclidean distance among the sensors as edges. The sensor readings are represented by time-varying graph signals.

We randomly remove some readings from the sensor data and reconstruct those values using the proposed method.

We compare the proposed method with four recovery approaches: kNN, EM, LRMC, and, PMF.

The proposed method outperforms the approaches mentioned earlier by as much as 54%.

Motivation

Available methods for missing value recosntruction (like kNN, Expectation-Maximization, Low-rank Matrix Completion and Probablistic Matrix Factorization and others) fail under situations of massive data unavailability.

Primary reason behind their failure is their inability to properly exploit the embedded spatio-temporal dependencies in the sensor data.

In this regard, GSP-based methods outperform the SoTA methods, primarily because of their ability to extract meaningful information from an irregular data.

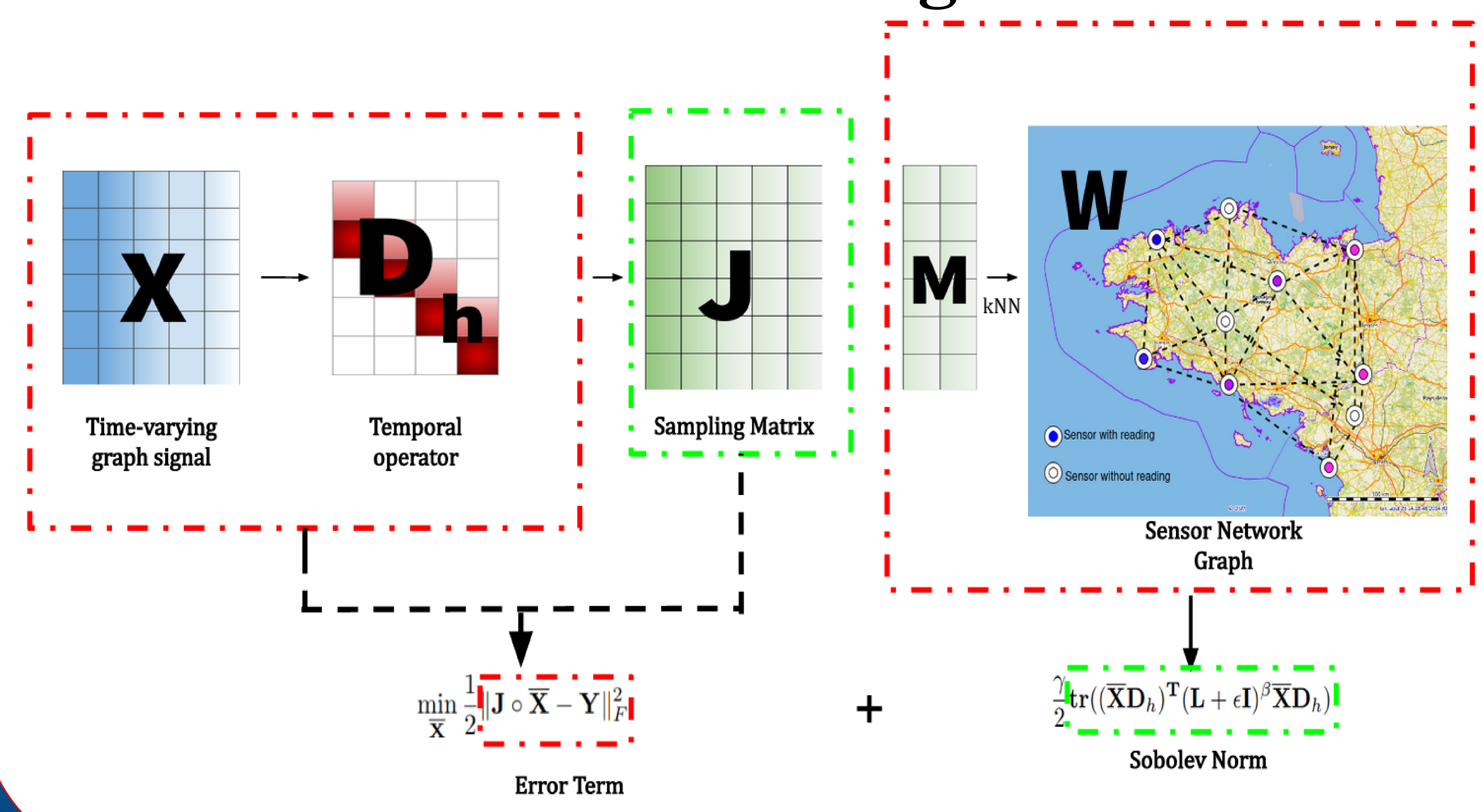
Contributions

We introduce the concepts of reconstruction of graph signals from GSP for recovering missing data in wireless sensor networks.

The Sobolev algorithm shows significant performance improvement over state-of-the-art (SoTA) algorithms dealing with the problem of missing sensor data.

We test the studied method on several publicly available datasets from diverse environments (indoor and outdoor) to demonstrate its versatility.

Schematic Diagram



Comparison Results

	Sampling Density	kNN [2], [3]		EM [4]		LRMC [5], [6]		PMF [7]		Proposed	
		RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
Molene Dataset	0.1	8.05	6.75	7.34	5.68	3.47	2.35	4.75	2.56	**2.40**	**1.65**
	0.3	5.45	4.68	5.20	3.93	2.82	1.35	3.42	1.64	**1.60**	**1.02**
	0.5	4.20	3.04	4.14	2.95	1.96	1.02	2.45	1.09	**1.20**	**0.74**
	0.7	3.40	2.13	3.19	1.82	1.02	0.81	1.22	0.95	**0.97**	**0.65**
Intel Dataset	0.1	9.70	5.34	7.78	5.63	4.68	2.94	5.34	3.98	**3.04**	**1.50**
	0.3	6.30	3.52	6.45	4.96	3.40	2.15	3.95	2.37	**2.01**	**0.97**
	0.5	4.60	2.98	5.17	3.81	2.06	1.21	2.75	1.65	**1.57**	**0.70**
	0.7	3.75	1.78	4.63	2.57	1.45	0.75	1.97	0.99	**1.30**	**0.55**

References

Giraldo, Jhony H., and Thierry Bouwmans. "On the minimization of Sobolev norms of time-varying graph signals: Estimation of new coronavirus disease 2019 cases." In 2020 IEEE 30th International Workshop on Machine Learning for Signal Processing (MLSP), pp. 1-6. IEEE, 2020.

Qiu, K., Mao, X., Shen, X., Wang, X., Li, T., & Gu, Y. (2017). Time-varying graph signal reconstruction. IEEE Journal of Selected Topics in Signal Processing, 11(6), 870-88.