









A Survey on Graph Neural Networks for Time Series: Forecasting, Classification, Imputation, and Anomaly Detection

Ming Jin^{*1}, Huan Yee Koh^{*1}, Qingsong Wen², Daniele Zambon³, Cesare Alippi³, Geoffrey I. Webb¹, Irwin King⁴, Shirui Pan⁵

> ¹ Monash University ² Alibaba Group ³ Università della Svizzera italiana ³ The Chinese University of Hong Kong ³ Griffith University

> > October 12th, 2023

Outline

- Introduction
- Framework and Categorization
- Graph Neural Networks for Time Series Forecasting
- Graph Neural Networks for Time Series Classification
- Graph Neural Networks for Time Series Anomaly Detection
- Graph Neural Networks for Time Series Imputation
- Practical Applications
- Future Directions



Part 1. Introduction & Taxonomy

- Time series and graph neural networks: background
- Graph neural networks for time series analysis: an overview
- Framework and categorization

What are time series? •



Time Series

['tīm 'sir-(,)ēz]

A sequence of data points that occur in successive order over some period





• What are time series?



• What are graphs?



Bibliography Networks



Knowledge Graphs



Traffic Networks

• What are graph neural networks (GNNs)?



• What are graph neural networks (GNNs)?



• Graph neural networks for time series analysis (GNN4TS)



Beyond the temporal dependencies

• Graph neural networks for time series analysis (GNN4TS)



• Graph neural networks for time series analysis (GNN4TS)



• Graph neural networks for time series analysis (GNN4TS)



In this example of wind farm, different analytical tasks can be categorized into time series **forecasting**, **classification**, **anomaly detection**, and **imputation**

- How to obtain the graph structures?
 - To employ GNNs for time series analysis, it is implied that a graph structure must be provided
 - However, not all time series data have readily available graph structures



Heuristic-based approaches



https://www.linkedin.com/pulse/future-forensics-heuristic-approach-vidhura-sethu/ https://www.kdnuggets.com/2019/08/neighbours-machine-learning-graphs.html

- Heuristic-based graphs
 - Spatial proximity: This approach defines the graph structure by considering the proximity between pairs of nodes based on, e.g., their geographical location.

$$\mathbf{A}_{i,j} = \begin{cases} \frac{1}{dij}, & \text{if } d_{ij} \neq 0, \\ 0, & \text{otherwise,} \end{cases} \quad d_{ij} \text{ denotes the shortest travel distance between node } i \text{ and } j \end{cases}$$

 Pairwise connectivity: The graph structure is determined by the connectivity between pairs of nodes, like that determined by transportation networks.

$$\mathbf{A}_{i,j} = \begin{cases} 1, & \text{if } v_i \text{ and } v_j \text{ are directly linked}, \\ 0, & \text{otherwise.} \end{cases}$$

- Heuristic-based graphs
 - Pairwise similarity: This method constructs the graph by connecting nodes with similar attributes.

$$\mathbf{A}_{i,j} = \frac{\mathbf{x}_i^\top \mathbf{x}_j}{\|\mathbf{x}_i\| \|\mathbf{x}_j\|} \qquad x_i \text{ and } x_j \text{ are two time series (variables), and } \|.\|$$
denotes the Euclidean norm

 Functional dependence: This approach defines the graph structure based on the functional dependence between pairs of nodes.

$$\mathbf{A}_{i,j} = \begin{cases} 1, & \text{if node } j \text{ Granger-causes} \\ & \text{node } i \text{ at a significance level } \alpha, \\ 0, & \text{otherwise.} \end{cases}$$

Other examples involve transfer entropy (TE) and discrete phase lag index (DPLI)

Learning-based graphs





Embedding-based (e.g., MTGNN)

Sampling-based (e.g., GTS)

Shang, C., Chen, J., & Bi, J. (2021). Discrete graph structure learning for forecasting multiple time series. In ICLR

Wu, Z., Pan, S., Long, G., Jiang, J., Chang, X., & Zhang, C. (2020). Connecting the dots: Multivariate time series forecasting with graph neural networks. In KDD.

Outline of GNN4TS



General pipeline for time series analysis using graph neural networks

Outline of GNN4TS



Methodology-oriented taxonomy of GNN4TS

Part 2. GNNs for Time Series Analysis

- Time series analytical tasks: an overview

- Graph neural networks for time series analysis: 4 categories

Time Series Analysis



GNNs for time series forecasting



GNNs for time series classification

Time Series Analysis



GNNs for time series anomaly detection

GNNs for time series imputation

- Modeling the inter-variable dependencies
 - Spectral GNNs, spatial GNNs, or a hybrid of both
- Modeling the inter-temporal dependencies
 - In the time or/and frequency domains
 - Recurrent models, convolution models, attention models, or hybrid models
- Forecasting architecture fusion
 - Discrete and continuous neural architectures
 - Spatial and temporal modules can be factorized or coupled



Approach	Year	Venue	Task	Architecture	Spatial Module	Temporal Module	Missing Values	Input Graph	Learned Relations	Graph Heuristics
DCRNN [71]	2018	ICLR	M-S	D-C	Spatial GNN	T-R	No	R	-	SP
STGCN [58]	2018	IJCAI	M-S	D-F	Spectral GNN	T-C	No	R	-	SP
ST-MetaNet [73]	2019	KDD	M-S	D-F	Spatial GNN	T-R	No	R	-	SP, PC
NGAR [74]	2019	IEEE IJCNN	S-S	D-F	Spatial GNN	T-R	No	R	-	-
ASTGCN [75]	2019	AAAI	M-S	D-F	Spectral GNN	T-H	No	R	-	SP, PC
ST-MGCN [46]	2019	AAAI	S-S	D-F	Spectral GNN	T-R	No	R	-	SP, PC, PS
Graph WaveNet [76]	2019	IJCAI	M-S	D-F	Spatial GNN	T-C	No	0	S	SP
MRA-BGCN [77]	2020	AAAI	M-S	D-C	Spatial GNN	T-R	No	R	-	SP
MTGNN [53]	2020	KDD	S-S, M-S	D-F	Spatial GNN	T-C	No	NR	S	-
STGNN* [78]	2020	WWW	M-S	D-C	Spatial GNN	T-H	No	R	-	SP
GMAN [79]	2020	AAAI	M-S	D-C	Spatial GNN	T-A	No	R	-	SP
SLCNN [80]	2020	AAAI	M-S	D-F	Hybrid	T-C	No	NR	S	-
STSGCN [81]	2020	AAAI	M-S	D-C	Spatial GNN	Т	No	R	-	PC
StemGNN [54]	2020	NeurIPS	M-S	D-F	Spectral GNN	F-C	No	NR	S	-
AGCRN [82]	2020	NeurIPS	M-S	D-C	Spatial GNN	T-R	No	NR	S	-
LSGCN [83]	2020	IJCAI	M-S	D-F	Spectral GNN	T-C	No	R	-	SP
STAR [84]	2020	ECCV	M-S	D-F	Spatial GNN	T-A	No	R	-	PC
GTS [56]	2021	ICLR	M-S	D-C	Spatial GNN	T-R	No	NR	S	-
GEN [85]	2021	ICLR	S-S	D-F	Spatial GNN	T-R	No	R	-	-
Z-GCNETs [86]	2021	ICML	M-S	D-C	Spatial GNN	T-C	No	NR	S	-
STGODE [70]	2021	KDD	M-S	C-F	Spatial GNN	T-C	No	R	-	SP, PS
STFGNN [49]	2021	AAAI	M-S	D-F	Spatial GNN	T-C	No	R	-	SP, PS
DSTAGNN [87]	2022	ICML	M-S	D-F	Spectral GNN	T-H	No	R	-	PC, PS
TPGNN [88]	2022	NeurIPS	S-S, M-S	D-F	Spatial GNN	T-A	No	NR	D	-
MTGODE [23]	2022	IEEE TKDE	S-S, M-S	C-C	Spatial GNN	T-C	No	NR	S	-
STG-NCDE [89]	2022	AAAI	M-S	C-C	Spatial GNN	T-C	Yes	NR	S	-
STEP [90]	2022	KDD	M-S	D-F	Spatial GNN	T-A	No	NR	S	-
Chauhan et al. [91]	2022	KDD	M-S	-	-	-	Yes	0	S	SP
RGSL [92]	2022	IJCAI	M-S	D-C	Spectral GNN	T-R	No	R	S	SP, PC
FOGS [93]	2022	IJCAI	M-S	-	-	-	No	NR	S	-
METRO [94]	2022	VLDB	M-S	D-C	Spatial GNN	Т	No	NR	D	-
SGP [95]	2023	AAAI	M-S	D-F	Spatial GNN	T-R	No	R	-	SP, PS
HiGP [96]	2023	arXiv	M-S	D-F	Spatial GNN	T-R	No	R	S	SP, PS
Jin et al. [29]	2023	arXiv	M-S, M-L	D-F	Spectral GNN	F-H	No	NR	S	-

- **Task**: "S" and "L" denote shortterm and long-term forecasting
- Architecture: "D" and "C" represent discrete and continuous; "C" and "F" stand for coupled and factorized
- **Temporal module**: "T" and "F" denote time and frequency domains; "R", "C", "A", and "H" are recurrent, convolutional, attentional, and hybrid models
- Input graph: "R", "O", and "NR" stand for required, optional, and not required
- Graph heuristics: "SP", "PC", "PS", and "FD" are spatial proximity, pairwise connectivity, pairwise similarity, and functional dependency₂₃

- Modeling the inter-variable dependencies
 - Spectral GNNs (e.g., StemGNN) & Spatial GNNs (e.g., DCRNN)
- Modeling the inter-temporal dependencies
 - Time domain (e.g., DCRNN) & Frequency domain (e.g., StemGNN)
 - Recurrent models (e.g., DCRNN) & Convolution models (e.g., STGCN) & Attention models (e.g., GMAN)
- Forecasting architecture fusion
 - Discrete factorized (e.g., MTGNN) & Continuous coupled (e.g., MTGODE)

• STGCN (Yu et al., 2018)



- Model architecture: discrete factorized
- Spatial module: ChebConv or GCN
- Temporal module: convolution-based

Yu, B., Yin, H., & Zhu, Z. (2018, July). Spatio-temporal graph convolutional networks: a deep learning framework for traffic forecasting. In Proceedings of the 27th International Joint Conference on Artificial Intelligence (pp. 3634-3640).

Monash University

• DCRNN (Li et al., 2018)



Model architecture: discrete coupled

- **Spatial module:** graph diffusion
- Temporal module: recurrent-based

Li, Y., Yu, R., Shahabi, C., & Liu, Y. (2018, February). Diffusion Convolutional Recurrent Neural Network: Data-Driven Traffic Forecasting. In International Conference on Learning Representations.

• GMAN (Zheng et al., 2020)



- Model architecture: discrete coupled
- Spatial module: GAT
- Temporal module: attention-based

Zheng, C., Fan, X., Wang, C., & Qi, J. (2020, April). Gman: A graph multi-attention network for traffic prediction. In Proceedings of the AAAI conference on artificial intelligence (Vol. 34, No. 01, pp. 1234-1241).

Monash University

• StemGNN (Cao et al., 2020)



- Model architecture: discrete factorized
- Spatial module: ChebConv
- Temporal module: convolution-based but in the frequency domain

• MTGNN (Wu et al., 2020)



- Model architecture: discrete factorized
- Spatial module: mix-hop MPNN
- **Temporal module:** convolution-based

Wu, Z., Pan, S., Long, G., Jiang, J., Chang, X., & Zhang, C. (2020, August). Connecting the dots: Multivariate time series forecasting with graph neural networks. In Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining (pp. 753-763).

• MTGODE (Jin et al., 2022)



- Model architecture: continuous coupled
- Spatial module: GDEs
- Temporal module: TDEs

- Univariate time series classification
 - Series-as-Graph: the core idea is to transform a univariate time series into a graph to identify unique patterns
 - Series-as-Node: each time series sample in a given dataset can be viewed as a node to construct a data graph
- Multivariate time series classification
 - This branch of methods maintains fundamental similarities with its univariate counterpart
 - However, it introduces an additional layer of complexity: the necessity to capture the inter-variable dependencies



Classification

Approach	Year	Venue	Task	Conversion	Spatial Module	Temporal Module	Missing Values	Input Graph	Learned Relations	Graph Heuristics
MTPool [171]	2021	NN	М		Spatial GNN	T-C	No	NR	S	-
Time2Graph+ [64]	2021	IEEE TKDE	U	Series-As-Graph	Spatial GNN	-	No	R	-	PS
RainDrop [37]	2022	ICLR	Μ	-	Spatial GNN	T-A	Yes	NR	S	-
SimTSC [65]	2022	SDM	U+M	Series-As-Node	Spatial GNN	T-C	No	R	-	PS
LB-SimTSC [66]	2023	arXiv	U+M	Series-As-Node	Spatial GNN	T-C	No	R	-	PS
TodyNet [172]	2023	arXiv	М	-	Spatial GNN	T-C	No	NR	D	-

- **Task**: "M" and "U" denote univariate and multivariate time series classification
- **Temporal module**: "T" and "F" denote time and frequency domains; "R", "C", "A", and "H" are recurrent, convolutional, attentional, and hybrid models
- Input graph: "R", "O", and "NR" stand for required, optional, and not required
- Learned relations: "S" and "D" denote static and dynamic graph structures
- **Graph heuristics**: "SP", "PC", "PS", and "FD" are spatial proximity, pairwise connectivity, pairwise similarity, and functional dependency

• Time2Graph+ (Cheng et al., 2021)



- Each time series is transformed into a graph where shapelets form nodes and transition probabilities create edges
- GAT is then leveraged along with a graph pooling operation to derive the global representation of the time series

Fig. 1. Illustration of *Time2Graph+* framework in the scenario of user electricity consuming. (a) shows a one-year electricity consumption of an emptynest elderly user, along with the assigned shapelet #40. After time-aware shapelets being extracted, it constructs the shapelet graph for each single sequence, and captures the evolutionary patterns of shapelets using graph attention networks (b), and (c) visualizes one typical shapelet #34 and its time-level attentions. Note that in (b), the node size is proportional to its in-degree, from which the node color is mapped; and the edge width is proportional to the attention score between two nodes.

Cheng, Z., Yang, Y., Jiang, S., Hu, W., Ying, Z., Chai, Z., & Wang, C. (2021). Time2Graph+: Bridging time series and graph representation learning via multiple attentions. IEEE Transactions on Knowledge and Data Engineering.

• SimTSC (Zha et al., 2022)



Figure 2: An overview of SimTSC framework. The graph is constructed based on the pair-wise similarities (e.g., DTW distances) of the time-series. Each time-series is processed by a backbone (e.g., ResNet) for feature extraction. The GNN module will aggregate the features and produce the final representations for classification.

- Time series nodes are connected using edges, which are defined by their pairwise DTW distance, to construct a graph
- A backbone network is initially employed to encode each time series into a feature vector
- Subsequently, a standard GNN operation is implemented to derive node (time series) representations, capturing the similarities between the series, for better classification

• RainDrop (Zhang et al., 2021)



- To classify irregularly sampled multivariate time series where subsets of variables have missing values at certain timestamps, Raindrop adaptively learns a "sensor graph"
- It then dynamically interpolates missing observations within the embedding space, based on any available recorded data

- Discrepancy framework for anomaly detection
 - Reconstrction discrepancy: the reconstructed error should be low during normal periods, but high during anomalous periods
 - Forecast discrepancy: the forecast error should be low during normal periods, but high during anomalous periods. Here the backbone is substituted with a GNN forecaster that is trained to predict a one-step-ahead forecast



Hybrid and other discrepancy



Anomaly Detection

Approach	Year	Venue	Strategy	Spatial Module	Temporal Module	Missing Values	Input Graph	Learned Relations	Graph Heuristics
CCM-CDT [52]	2019	IEEE TNNLS	RC	Spatial GNN	T-R	No	R	-	PC, FD
MTAD-GAT [39]	2020	IEEE ICDM	FC+RC	Spatial GNN	T-A	No	NR	-	-
GDN [40]	2021	AAAI	FC	Spatial GNN	-	No	NR	S	-
GTA [147]	2021	IEEE IoT	FC	Spatial GNN	T-H	No	NR	S	-
EvoNet [148]	2021	WSDM	CL	Spatial GNN	T-R	No	R	-	PS
Event2Graph [149]	2021	arXiv	RL	Spatial GNN	T-A	No	R	-	PS
GANF [61]	2022	ICLR	RC+RL	Spatial GNN	T-R	No	NR	S	-
Grelen [150]	2022	IJCAI	RC+RL	Spatial GNN	T-H	No	NR	D	-
VGCRN [151]	2022	ICML	FC+RC	Spatial GNN	T-R	No	NR	S	-
FuSAGNet [60]	2022	KDD	FC+RC	Spatial GNN	T-R	No	NR	S	-
GTAD [152]	2022	Entropy	FC+RC	Spatial GNN	T-C	No	NR	-	-
HgAD [153]	2022	IEEE BigData	FC	Spatial GNN	-	No	NR	S	
HAD-MDGAT [154]	2022	IEEE Access	FC+RC	Spatial GNN	T-A	No	NR	-	-
STGAN [154]	2022	IEEE TNNLS	RC	Spatial GNN	T-R	No	R	-	SP
GIF [155]	2022	IEEE IJCNN	RC	Spatial GNN	-	No	R	-	SP, PC, FD
DyGraphAD [156]	2023	arXiv	FC+RL	Spatial GNN	T-C	No	R	-	PS
GraphSAD [157]	2023	arXiv	CL	Spatial GNN	T-C	No	R	-	PS, PC
CST-GL [158]	2023	arXiv	FC	Spatial GNN	T-C	No	NR	S	-

• **Strategy**: "RC", "FC", "RL", and "CL" indicate the reconstruction, forecast, relational, and class discrepancies

• GDN (Deng & Hooi, 2021)



Deng, A., & Hooi, B. (2021, May). Graph neural network-based anomaly detection in multivariate time series. In Proceedings of the AAAI conference on artificial intelligence (Vol. 35, No. 5, pp. 4027-4035).

• VGCRN (Chen et al., 2022)



Chen, W., Tian, L., Chen, B., Dai, L., Duan, Z., & Zhou, M. (2022, June). Deep variational graph convolutional recurrent network for multivariate time series anomaly detection. In International Conference on Machine Learning (pp. 3621-3633). PMLR.

• CST-GL (Zheng et al., 2023)



Zheng, Y., Koh, H. Y., Jin, M., Chi, L., Phan, K. T., Pan, S., ... & Xiang, W. (2023). Correlation-aware Spatial-Temporal Graph Learning for Multivariate Time-series Anomaly Detection. arXiv preprint arXiv:2307.08390.

GNNs for Time Series Imputation

Task categorization

- In-sample imputation: this involves filling in missing values within the given time series data
- Out-of-sample imputation: this predicts missing values in disjoint sequences



- Methodology categorization
 - Deterministic imputation: this provides a single best estimate for the missing values
 - Probabilistic imputation: this accounts for the uncertainty in the imputation and provides a distribution of possible values

GNNs for Time Series Imputation

Approach	Year	Venue	Task	Туре	Spatial Module	Temporal Module	Inductiveness	Input Graph	Learned Relations	Graph Heuristics
IGNNK [187]	2021	AAAI	Out-of-sample	Deterministic	Spectral GNN	-	Yes	R	-	SP, PC
GACN [188]	2021	ICANN	In-sample	Deterministic	Spatial GNN	T-C	No	R	-	PC
SATCN [189]	2021	arXiv	Out-of-sample	Deterministic	Spatial GNN	T-C	Yes	R	-	SP
GRIN [41]	2022	ICLR	Both	Deterministic	Spatial GNN	T-R	Yes	R	-	SP
SPIN [190]	2022	NIPS	In-sample	Deterministic	Spatial GNN	T-A	No	R	-	SP
FUNS [191]	2022	ICDMW	Out-of-sample	Deterministic	Spatial GNN	T-R	Yes	R	-	-
AGRN [192]	2022	ICONIP	In-sample	Deterministic	Spatial GNN	T-R	No	NR	S	-
MATCN [193]	2022	IEEE IoT-J	In-sample	Deterministic	Spatial GNN	T-A	No	R	-	-
PriSTI [42]	2023	arXiv	In-sample	Probabilistic	Spatial GNN	T-A	No	R	-	SP
DGCRIN [194]	2023	KBS	In-sample	Deterministic	Spatial GNN	T-R	No	NR	D	-
GARNN [195]	2023	Neurocomputing	In-sample	Deterministic	Spatial GNN	T-R	No	R	-	PC
MDGCN [196]	2023	Transp. Res. Part C	In-sample	Deterministic	Spatial GNN	T-R	No	R	-	SP, PS

• Inductiveness: this indicates whether a method can generalize to unseen nodes



• PriSTI (Liu et al., 2023)



PriSTI leverages the diffusion model for spatio-temporal imputation

Probabilistic imputations

Part 3. Applications & Future Directions

• Smart Transportation



• Environment & Sustainable Energy





Monash University

Fraud Detection



Future Directions

• Pre-training, transfer learning, and large models



Future Directions

• Uncertainty quantification



Wen, H., Lin, Y., Xia, Y., Wan, H., Zimmermann, R., & Liang, Y. (2023). Diffstg: Probabilistic spatio-temporal graph forecasting with denoising diffusion models. *arXiv preprint arXiv:2301.13629*.

Monash University

Future Directions

• Scalability



Released Source	Dataset	Nodes	Edges	Degree	Meta	Time Range	Frames	Data Points
Vu at al [24]	PeMSD7(M)	228	1,664	7.3	6	05/01/2012 - 06/30/2012	12,672	2.89 <mark>M</mark>
Iu et al. [34]	PeMSD7(L)	1,026	14,534	14.2	0	05/01/2012 - 06/30/2012	12,672	13.00 <mark>M</mark>
Listal [10]	METR-LA	207	1,515	7.3	3	03/01/2012 - 06/27/2012	34,272	7.09 <mark>M</mark>
Li et al. [19]	PEMS-BAY	325	2,369	7.3	3	01/01/2017 - 06/30/2017	52,116	16.94 <mark>M</mark>
	PEMS03	358	546	1.5	1	09/01/2018 - 11/30/2018	26,208	9.38 <mark>M</mark>
Song et al. [20]	PEMS04	307	338	1.1	0	01/01/2018 - 02/28/2018	16,992	5.22 M
3011g et al. [30]	PEMS07	883	865	1.0	0	05/01/2017 - 08/06/2017	28,224	24.92 <mark>M</mark>
	PEMS08	170	276	1.6	0	07/01/2016 - 08/31/2016	17,856	3.04 <mark>M</mark>
	CA	8,600	201,363	23.4	9	01/01/2017 - 12/31/2021	525,888	4.52 B
LargeST (ours)	GLA	3,834	98,703	25.7	9	01/01/2017 - 12/31/2021	525,888	2.02 B
Larges 1 (ours)	GBA	2,352	61,246	26.0	9	01/01/2017 - 12/31/2021	525,888	1.24 B
	SD	716	17,319	24.2	9	01/01/2017 - 12/31/2021	525,888	0.38 <mark>B</mark>

(a) Overview of the LargeST dataset

(b) Fine-grained distribution of sensors

Figure 1: An illustration of the LargeST benchmark dataset.

Part 4. Conclusion

Conclusion

- What we have covered?
 - Taxonomies of graph neural networks for time series analysis (GNN4TS)



Conclusion

- What we have covered?
 - An overview of graph neural networks for time series analysis (GNN4TS)



Fig. 4: Four categories of graph neural networks for time series analysis. For the sake of simplicity and illustrative purposes, we assume the graph structures are fixed in all subplots.

Conclusion

- What we have covered?
 - Applications & Future directions of GNN4TS













Monash University



Thank you

Paper: https://arxiv.org/pdf/2307.03759.pdf

GitHub Page: https://github.com/KimMeen/Awesome-GNN4TS

