Natural and Artificial Dynamics in GNNs: A Tutorial

Dongqi Fu
University of Illinois Urbana-Champaign
Illinois, USA
dongqif2@illinois.edu

Zhe Xu
University of Illinois Urbana-Champaign
Illinois, USA
zhexu3@illinois.edu

Hanghang Tong
University of Illinois Urbana-Champaign
Illinois, USA
htong@illinois.edu

Jingrui He
University of Illinois Urbana-Champaign
Illinois, USA
jingrui@illinois.edu

ABSTRACT

In the big data era, the relationship between entities becomes more complex. Therefore, graph (or network) data attracts increasing research attention for carrying complex relational information. For a myriad of graph mining/learning tasks, graph neural networks (GNNs) have been proven as effective tools for extracting informative node and graph representations, which empowers a broad range of applications such as recommendation, fraud detection, molecule design, and many more. However, real-world scenarios bring pragmatic challenges to GNNs. First, the input graphs are evolving, i.e., the graph structure and node features are time-dependent. Integrating temporal information into the GNNs to enhance their representation power requires additional ingenious designs. Second, the input graphs may be unreliable, noisy, and suboptimal for a variety of downstream graph mining/learning tasks. How could end-users deliberately modify the given graphs (e.g., graph topology and node features) to boost GNNs' utility (e.g., accuracy and robustness)? Inspired by the above two kinds of dynamics, in this tutorial, we focus on topics of *natural dynamics* and *artificial dynamics* in GNNs and introduce the related works systematically. After that, we point out some promising but under-explored research problems in the combination of these two dynamics. We hope this tutorial could be beneficial to researchers and practitioners in areas including data mining, machine learning, and general artificial intelligence.

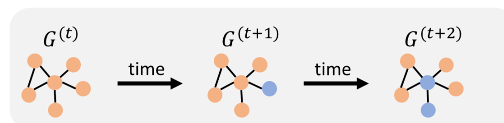
CCS CONCEPTS

• **Mathematics of computing → Graph algorithms**; • **Computing methodologies → Neural networks**; **Learning latent representations**.

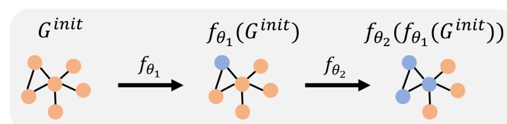
KEYWORDS

Graph Neural Networks; Temporal Graphs; Graph Augmentation

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
WSDM '23, February 27-March 3, 2023, Singapore, Singapore
© 2023 Association for Computing Machinery.
ACM ISBN 978-1-4503-9407-9/23/02…$15.00
https://doi.org/10.1145/3539597.3572726



(a) Natural Dynamics



(b) Artificial Dynamics

Figure 1: Natural and Artificial Dynamics w.r.t Node Feature Changes.

ACM Reference Format:
Dongqi Fu, Zhe Xu, Hanghang Tong, and Jingrui He. 2023. Natural and Artificial Dynamics in GNNs: A Tutorial. In *Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining (WSDM '23), February 27-March 3, 2023, Singapore, Singapore.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3539597.3572726

1 INTRODUCTION

Graph neural networks (GNNs), as one kind of deep learning model, attract much research attention for effectively leveraging entity relational information and providing informative representation vectors for numerous application domains [20]. In the complex real world, the adaption of GNNs, at least, faces two general challenges. First, given an input graph, its topological structure and node (or graph) features can be dependent on time, i.e., they are evolving with time. A typical example is the World Wide Web. The resulting problems to GNNs include but are not limited to ignoring entity temporal correlation, overlooking causality discovery, computation inefficiency, non-generalization, etc. Second, given an input graph, its initial topological structure and node (or graph) features may be imperfect (e.g., having construction errors, sampling noises, missing features, etc.). The corresponding problems for GNNs include but are not limited to non-robustness, indiscriminative representation vectors, non-generalization, etc.

Studying natural and artificial dynamics in graphs is the research to solve the challenges brought by temporally evolving and imperfect input graphs [6]. To be specific, natural dynamics and artificial dynamics in graphs are two general terms. *Natural dynamics in graphs* can illustrate that the input graphs themselves are evolving, i.e., the topology structures, the node-level, edge-level, and (sub)graph-level features, and labels are dependent on time [1, 11]. As for *artificial dynamics in graphs*, we can use this term to describe that end-users *change the existing* or *construct the non-existing* graph-related elements (e.g., graph topology, node/graph attributes/labels, GNN layer connections, GNN gradients, etc.) to achieve certain performance boosting in graph mining [7] and graph representation learning tasks [5, 26]. In Figure 1, the difference between natural and artificial dynamics is shown by setting the feature changes as an example. In Figure 1(a), the node features are evolving along with time, where a GNN-based graph representation learning model should consider the different node features at different timestamps when producing the node representation vectors [4]. In Figure 1(b), when the input graphs are imperfect to the downstream tasks like node classification, a learning-based model θ can be introduced to augment the node feature to improve the GNN-based classifier accuracy [22].

In this tutorial, we narrow the scope of natural and artificial dynamics in GNNs by (1) using "**natural dynamics in GNNs**" to focus on temporal graph neural networks (TGNNs) and introduce how related works leverage time information during the message-passing process of GNNs; (2) and using "**artificial dynamics in GNNs**" to target how related works modify (i.e., corrupt or augment) the input graph structure and features [2] to enhance GNNs performance. To introduce natural and artificial dynamics in GNNs, we will cover some classic and nascent related works, and this tutorial is organized into four parts. In Part I, we begin with introducing the preliminaries of GNNs, in terms of background knowledge, motivations, and problem settings. In Part II, we first discuss the key challenges brought by evolving input graphs to GNNs, and then we introduce different works in effectively dealing with time information during the representation learning process of GNNs. In Part III, the key challenges brought by imperfect input graphs to GNNs are mentioned first, and then we introduce the different strategies from end-users to change the input graphs and boost the GNNs performance. In Part IV, considering the imperfect and evolving graphs, we discuss future research opportunities and provide some insights to the research community.

2 TUTORIAL DETAILS

The tutorials will be 3 hours, and the outline is as follows.

2.1 Outline of Tutorial

- Part I: Introduction
  - Background and Motivations
  - Problem Definitions and Settings
- Part II: Natural Dynamics in GNNs
  - Key Challenges from Involving Input Graphs
  - Incorporating GNNs with Convolutional Operations
  - Incorporating GNNs with Recurrent Units
  - Incorporating GNNs with Time Attention
  - Incorporating GNNs with Time Point Process
  - Incorporating GNNs with Time Kernel
- Part III: Artificial Dynamics in GNNs
  - Key Challenges from Imperfect Input Graphs
  - Heuristic-based Artificial Dynamics
  - Data-driven Artificial Dynamics
- Part IV: Open Questions and Challenges
  - Key Challenges from Imperfect Evolving Graphs
  - Data-driven Artificial Dynamics on Temporal graphs
  - Transferable Artificial Dynamics on Temporal Graphs
  - Interpretable Artificial Dynamics on Temporal Graphs

The details for Parts II – IV are introduced as follows.

2.2 Part II: Natural Dynamics in GNNs

For dealing with graphs having evolving natures, temporal graph neural networks (TGNNs) are proposed. The general principle to TGNNs is that the input graphs have evolving components, e.g., the graph structure and/or node attributes, etc. Since the TGNNs take the graphs as input and the graph topological information is also called spatial information in some applications (e.g., traffic modeling and sensor modeling), TGNNs are also called spatial-temporal graph neural networks (STGNNs) in some works [20]. Here, we use the term temporal graph neural networks (TGNNs).

To appropriately consider the time information during the GNNs' information aggregation process, many TGNN works utilize different temporal models. For example, (1) in [23, 25], authors apply the convolutional operations from convolutional neural networks (CNNs) on graphs to capture temporal features. (2) In [9, 13, 15], authors insert the recurrent units (from various RNNs such as LSTM and GRU) into GNNs to preserve the temporal dependency during the GNNs' representation learning process. (3) In [17], authors propose to use the self-attention mechanism on time features to learn the temporal correlations along with node representations. Also, (4) authors in [18] utilize Time Point Process to model time, and (5) authors in [21] use Time Kernel to model time, respectively, in order to obtain time domain features.

2.3 Part III: Artificial Dynamics in GNNs

In this part, we will introduce two lines of artificial dynamics in GNNs. Overall, we categorize artificial dynamics into heuristic-based and data-driven dynamics. For the heuristic-based artificial dynamics, we introduce 4 common applications on graph machine learning, and their representative works include: (1) improving GNNs' message passing by *inserting a super node* into the given graphs [8], (2) improving the generalization ability of GNNs by *randomly dropping edges* [16] during the message passing , (3) *adding features* from the network topology [12] to break the GNNs' expressiveness limitation, and (4) *randomly perturbing the given graphs* to generate corrupted views for graph contrastive learning [24].

For the data-driven artificial dynamics, our introduction is still around 3 typical applications which include: (1) *optimizing the given topology and node features* for graph denoising [22], (2) *learning an adversarial graph view* for better contrastive learning samples [3], and (3) *condensing a given graph into a small and informative graph* [10] to speed up the training of GNNs.

2.4 Part IV: Open Questions and Challenges

In this part, we will introduce several under-explored questions with a special focus on the interaction between natural and artificial dynamics, i.e., augmenting the temporal graphs. In general, research about data augmentation on temporal graphs [14, 19] is rare, and most, if not all, of them are based on heuristics. We will point out some promising open questions, including (1) data-driven augmentation on temporal graphs and scalable solutions to handle tremendous search space, (2) transferable and generalizable augmentation techniques across multiple distinct temporal graphs, and (3) interpretable augmentation on temporal graphs where new methods are needed to understand the augmentations (e.g., the uncertainty of shifting the timestamps).

3 COVERED WORKS IN THIS TUTORIAL

For space, we only list part of representative works here. This is not an exhaustive list of papers that are relevant to the topic.

- Part II: Natural Dynamics in GNNs
  - CNN-based Temporal Graph Neural Networks
    - Bing Yu, Haoteng Yin, Zhanxing Zhu. Spatio-Temporal Graph Convolutional Networks: A Deep Learning Framework for Traffic Forecasting. IJCAI 2018.
  - RNN-based Temporal Graph Neural Networks
    - Yaguang Li, Rose Yu, Cyrus Shahabi, Yan Liu. Diffusion Convolutional Recurrent Neural Network: Data-Driven Traffic Forecasting. ICLR 2018.
    - Ehsan Hajiramezanali, Arman Hasanzadeh, Nick Duffield, Krishna R. Narayanan, Mingyuan Zhou, Xiaoning Qian. Variational Graph Recurrent Neural Networks. NeurIPS 2019.
  - Attention-based Temporal Graph Neural Networks
    - Aravind Sankar, Yanhong Wu, Liang Gou, Wei Zhang, Hao Yang. DySAT: Deep Neural Representation Learning on Dynamic Graphs via Self-Attention Networks. WSDM 2020.
  - Time Point Process-based Temporal Graph Neural Networks
    - Rakshit Trivedi, Mehrdad Farajtabar, Prasenjeet Biswal, Hongyuan Zha. DyRep: Learning Representations over Dynamic Graphs. ICLR 2019.
  - Time Kernel-based Temporal Graph Neural Networks
    - Da Xu, Chuanwei Ruan, Evren Körpeoglu, Sushant Kumar, Kannan Achan. Inductive representation learning on temporal graphs. ICLR 2020.
- Part III: Artificial Dynamics in GNNs
  - Heuristic-based Artificial Dynamics
    - Yu Rong, Wenbing Huang, Tingyang Xu, Junzhou Huang. DropEdge: Towards Deep Graph Convolutional Networks on Node Classification. ICML 2020.
    - Justin Gilmer, Samuel S. Schoenholz, Patrick F. Riley, Oriol Vinyals, George E. Dahl. Neural Message Passing for Quantum Chemistry. ICML 2017.
    - Pan Li, Yanbang Wang, Hongwei Wang, Jure Leskovec. Distance Encoding: Design Provably More Powerful Neural Networks for Graph Representation Learning. NeurIPS 2020.
    - Yuning You, Tianlong Chen, Yongduo Sui, Ting Chen, Zhangyang Wang, Yang Shen. Graph Contrastive Learning with Augmentations. NeurIPS 2020.
  - Data-driven Artificial Dynamics
    - Zhe Xu, Boxin Du, Hanghang Tong. Graph Sanitation with Application to Node Classification. WWW 2022.
    - Shengyu Feng, Baoyu Jing, Yada Zhu, Hanghang Tong. Adversarial Graph Contrastive Learning with Information Regularization. WWW 2022.
    - Wei Jin, Lingxiao Zhao, Shichang Zhang, Yozen Liu, Jiliang Tang, Neil Shah. Graph Condensation for Graph Neural Networks. ICLR 2022.
- Part IV: Open Questions and Challenges
  - Augmenting Temporal Graphs for TGNNs
    - Yiwei Wang, Yujun Cai, Yuxuan Liang, Henghui Ding, Changhu Wang, Siddharth Bhatia, Bryan Hooi. Adaptive Data Augmentation on Temporal Graphs. NeurIPS 2021.

4 RELATED TUTORIALS

- **Graph Representation Learning: Foundations, Methods, Applications and Systems**
  - **Presenters**: Wei Jin, Yao Ma, Yiqi Wang, Xiaorui Liu, Jiliang Tang, Yukuo Cen, Jiezhong Qiu, Jie Tang, Chuan Shi, Yanfang Ye, Jiawei Zhang, and Philip S. Yu.
  - **Conference**: KDD' 2021, August 14–18, 2021, Singapore
  - **Connection**: This tutorial introduces the graph neural network (GNN) models systematically, which paves the way for our tutorial for GNNs performance improvement through natural dynamics and artificial dynamics.
  - **Differences**: This tutorial introduces the graph neural networks in the fundamental theories, applications, and systems of GNNs. Our tutorial mainly focuses on the scenario where the input graphs for GNNs are evolving and imperfect and how the related works are proposed to address corresponding challenges.
- **Mining Temporal Networks**
  - **Presenters**: Polina Rozenshtein and Aristides Gionis.
  - **Conference**: KDD' 2019, August 4-8, 2019, USA
  - **Connection**: This tutorial introduces how to mine the temporal graphs to obtain knowledge to serve application tasks. The temporal graphs in this tutorial stand for the natural dynamics of our tutorial. Moreover, the goal of graph neural networks (GNNs) extracting meaningful node and graph hidden representations is to serve various graph mining tasks.
  - **Differences**: This tutorial focuses on graph mining tasks, especially when the input graphs are evolving (i.e., input graphs have natural dynamics). In our tutorial, we target the mining tools only at graph neural networks, and half of our tutorial introduces how GNNs effectively deal with evolving input graphs. We extend our tutorial with artificial dynamics to introduce how GNNs deal with imperfect graphs by dropping and augmenting.
- **Data Efficient Learning on Graphs**
  - **Presenters**: Chuxu Zhang, Jundong Li, Meng Jiang.
  - **Conference**: KDD' 2021, August 14–18, 2021, Singapore
  - **Connection**: This tutorial introduces self-supervised learning, which is an important application of artificial dynamics in GNNs.
  - **Differences**: This tutorial focuses on efficiently using the limited labeled samples on graph data, where self-supervised learning uses graph augmentation techniques to generate corrupted samples. Our tutorial covers this application in the artificial dynamics part and introduces the graph augmentation techniques systematically.

5 PRESENTER BIOGRAPHY

Dongqi Fu. He is currently a fourth-year Ph.D. student at the Department of Computer Science, University of Illinois Urbana-Champaign. His research interests focus on graph mining, graph representations, and graph neural networks, especially studying natural dynamics and artificial dynamics in graphs. He has published several research papers at top-tier conferences (e.g., KDD, SIGIR, CIKM, etc.) and served as the program committee member for many top-tier conferences (e.g., Top 8% reviewer in NeurIPS 2022) and as the reviewer for many journals (e.g., TKDD and TIST). For more details, please refer to his personal website at https://dongqifu.github.io/.

Zhe Xu. He is a Ph.D. student at the Department of Computer Science at the University of Illinois Urbana-Champaign. Prior to that, he obtained his M.S. degree in Computer Science from the University of Illinois Urbana-Champaign in 2019 and his B.E. degree in Electronic Engineering from Fudan University. His research interests lie in graph machine learning with a focus on graph data augmentation. His research works have been published at several major conferences in data mining and machine learning (e.g., TheWebConf, ICDM, CIKM, etc.). He also served as a reviewer in many top-tier data mining and machine learning venues. For more information, please check his personal website at https://pricexu.github.io/.

Hanghang Tong. He is currently an associate professor at Department of Computer Science at University of Illinois at Urbana-Champaign. Before that he was an associate professor at School of Computing, Informatics, and Decision Systems Engineering (CIDSE), Arizona State University. He received his M.Sc. and Ph.D. degrees from Carnegie Mellon University in 2008 and 2009, both in Machine Learning. His research interest is in large scale data mining for graphs and multimedia. He has received several awards, including SDM/IBM Early Career Data Mining Research award (2018), NSF CAREER award (2017), ICDM 10-Year Highest Impact Paper award (2015), four best paper awards (TUP 2014, CIKM 2012, SDM 2008, ICDM 2006), seven 'bests of conference', 1 best demo, honorable mention (SIGMOD 2017), and 1 best demo candidate second place (CIKM 2017). He has published over 200 refereed articles. He is the Editor-in-Chief of SIGKDD Explorations (ACM), and an associate editor of Knowledge and Information Systems (Springer) and Computing Surveys (ACM); and has served as a program committee member in multiple data mining, database and artificial intelligence venues (e.g., SIGKDD, CIKM, SIGMOD, AAAI, WWW, etc.). He has given several tutorials at top-tier conferences, such as IEEE Big Data 2015, SDM 2016, WSDM 2018, KDD 2018, CIKM 2020, etc. For more information, please refer to his personal website at http://tonghanghang.org/.

Jingrui He. She is currently an Associate Professor at the School of Information Sciences, University of Illinois at Urbana-Champaign. She also has a courtesy appointment with the Computer Science Department. She received her Ph.D. from Carnegie Mellon University in 2010. Her research focuses on heterogeneous machine learning, rare category analysis, active learning and semi-supervised learning, with applications in security, social network analysis, healthcare, and manufacturing processes. She is the recipient of the 2016 NSF CAREER Award, the 2020 OAT Award, three times recipient of the IBM Faculty Award in 2018, 2015 and 2014 respectively, and is selected as IJCAI 2017 Early Career Spotlight. She has more than 130 publications at major conferences (e.g., IJCAI, AAAI, KDD, ICML, NeurIPS) and journals (e.g., TKDE, TKDD, DMKD), and is the author of two books. Her papers have received the Distinguished Paper Award at FAccT 2022, as well as Bests of the Conference at ICDM 2016, ICDM 2010, and SDM 2010. For more details, please refer to her personal website at https://www.hejingrui.org/.

ACKNOWLEDGEMENT

This work is supported by National Science Foundation under Award No. IIS-1947203, IIS-2117902, and IIS-2137468. The views and conclusions are those of the authors and should not be interpreted as representing the official policies of the funding agencies or the government.

REFERENCES

[1] Charu C. Aggarwal and Karthik Subbian. 2014. Evolutionary Network Analysis: A Survey. *ACM Comput. Surv.* (2014).
[2] Kaize Ding, Zhe Xu, Hanghang Tong, and Huan Liu. 2022. Data Augmentation for Deep Graph Learning: A Survey. *CoRR* (2022).
[3] Shengyu Feng, Baoyu Jing, Yada Zhu, and Hanghang Tong. 2022. Adversarial Graph Contrastive Learning with Information Regularization. In *WWW 2022*.
[4] Dongqi Fu, Liri Fang, Ross Maciejewski, Vetle I. Torvik, and Jingrui He. 2022. Meta-Learned Metrics over Multi-Evolution Temporal Graphs. In *KDD 2022*.
[5] Dongqi Fu and Jingrui He. 2021. SDG: A Simplified and Dynamic Graph Neural Network. In *SIGIR 2021*.
[6] Dongqi Fu and Jingrui He. 2022. Natural and Artificial Dynamics in Graphs: Concept, Progress, and Future. *Frontiers in Big Data* (2022).
[7] Dongqi Fu, Dawei Zhou, and Jingrui He. 2020. Local Motif Clustering on Time-Evolving Graphs. In *KDD 2020*.
[8] Justin Gilmer, Samuel S. Schoenholz, Patrick F. Riley, Oriol Vinyals, and George E. Dahl. 2017. Neural Message Passing for Quantum Chemistry. In *ICML 2017*.
[9] Ehsan Hajiramezanali, Arman Hasanzadeh, Krishna R. Narayanan, Nick Duffield, Mingyuan Zhou, and Xiaoning Qian. 2019. In *NeurIPS 2019*.
[10] Wei Jin, Lingxiao Zhao, Shichang Zhang, Yozen Liu, Jiliang Tang, and Neil Shah. 2022. Graph Condensation for Graph Neural Networks. In *ICLR 2022*.
[11] Seyed Mehran Kazemi, Rishab Goel, Kshitij Jain, Ivan Kobyzev, Akshay Sethi, Peter Forsyth, and Pascal Poupart. 2020. Representation Learning for Dynamic Graphs: A Survey. *J. Mach. Learn. Res.* (2020).
[12] Pan Li, Yanbang Wang, Hongwei Wang, and Jure Leskovec. 2020. Distance Encoding: Design Provably More Powerful Neural Networks for Graph Representation Learning. In *NeurIPS 2020*.
[13] Yaguang Li, Rose Yu, Cyrus Shahabi, and Yan Liu. 2018. Diffusion Convolutional Recurrent Neural Network: Data-Driven Traffic Forecasting. In *ICLR 2018*.
[14] Xu Liu, Yuxuan Liang, Yu Zheng, Bryan Hooi, and Roger Zimmermann. 2021. Spatio-Temporal Graph Contrastive Learning. *CoRR* (2021).
[15] Aldo Pareja, Giacomo Domeniconi, Jie Chen, Tengfei Ma, Toyotaro Suzumura, Hiroki Kanezashi, Tim Kaler, Tao B. Schardl, and Charles E. Leiserson. 2020. EvolveGCN: Evolving Graph Convolutional Networks for Dynamic Graphs. In *AAAI 2020*.
[16] Yu Rong, Wenbing Huang, Tingyang Xu, and Junzhou Huang. 2020. DropEdge: Towards Deep Graph Convolutional Networks on Node Classification. In *ICLR 2020*.
[17] Aravind Sankar, Yanhong Wu, Liang Gou, Wei Zhang, and Hao Yang. 2020. DySAT: Deep Neural Representation Learning on Dynamic Graphs via Self-Attention Networks. In *WSDM 2020*.
[18] Rakshit Trivedi, Mehrdad Farajtabar, Prasenjeet Biswal, and Hongyuan Zha. 2019. DyRep: Learning Representations over Dynamic Graphs. In *ICLR 2019*.
[19] Yiwei Wang, Yujun Cai, Yuxuan Liang, Henghui Ding, Changhu Wang, Siddharth Bhatia, and Bryan Hooi. 2021. Adaptive Data Augmentation on Temporal Graphs. In *NeurIPS 2021*.
[20] Zonghan Wu, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, and Philip S. Yu. 2021. A Comprehensive Survey on Graph Neural Networks. *IEEE Trans. Neural Networks Learn. Syst.* (2021).
[21] Da Xu, Chuanwei Ruan, Evren Körpeoglu, Sushant Kumar, and Kannan Achan. 2020. Inductive representation learning on temporal graphs. In *ICLR 2020*.
[22] Zhe Xu, Boxin Du, and Hanghang Tong. 2022. Graph Sanitation with Application to Node Classification. In *WWW 2022*.
[23] Sijie Yan, Yuanjun Xiong, and Dahua Lin. 2018. Spatial Temporal Graph Convolutional Networks for Skeleton-Based Action Recognition. In *AAAI, 2018*.
[24] Yuning You, Tianlong Chen, Yongduo Sui, Ting Chen, Zhangyang Wang, and Yang Shen. 2020. Graph Contrastive Learning with Augmentations. In *NeurIPS 2020*.
[25] Bing Yu, Haoteng Yin, and Zhanxing Zhu. 2018. Spatio-Temporal Graph Convolutional Networks: A Deep Learning Framework for Traffic Forecasting. In *IJCAI 2018*.
[26] Lecheng Zheng, Dongqi Fu, Ross Maciejewski, and Jingrui He. 2022. Deeper-GXX: Deepening Arbitrary GNNs. *CoRR* (2022).