Synthesis Lectures on
Data Mining and Knowledge Discovery

SYNTHESIS
COLLECTION OF TECHNOLOGY

Chuan Shi · Xiao Wang ·
Cheng Yang

Advances in Graph Neural Networks

Springer

Synthesis Lectures on Data Mining and Knowledge Discovery

The series focuses on topics pertaining to data mining, web mining, text mining, and knowledge discovery, including tutorials and case studies. Potential topics include: data mining algorithms, innovative data mining applications, data mining systems, mining text, web and semi-structured data, high performance and parallel/distributed data mining, data mining standards, data mining and knowledge discovery framework and process, data mining foundations, mining data streams and sensor data, mining multi-media data, mining social networks and graph data, mining spatial and temporal data, pre-processing and post-processing in data mining, robust and scalable statistical methods, security, privacy, and adversarial data mining, visual data mining, visual analytics, and data visualization.

Chuan Shi · Xiao Wang · Cheng Yang

Advances in Graph Neural Networks

Springer

Chuan Shi
Beijing University of Posts
and Telecommunications
Beijing, China

Xiao Wang
Beijing University of Posts
and Telecommunications
Beijing, China

Cheng Yang
Beijing University of Posts
and Telecommunications
Beijing, China

ISSN 2151-0067 ISSN 2151-0075 (electronic)
Synthesis Lectures on Data Mining and Knowledge Discovery
ISBN 978-3-031-16173-5 ISBN 978-3-031-16174-2 (eBook)
https://doi.org/10.1007/978-3-031-16174-2

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Foreword

Relational structures are ubiquitous in the real world, such as social relations between people, transaction relations between companies, and biological relations between proteins. Graphs and networks are the most common way to characterize these structured data, where objects/relations are projected into nodes and edges, respectively. With the great success of machine learning and deep learning techniques, how to numerically represent a graph has become a fundamental problem in network analysis. In particular, graph representation learning, which aims to encode each node in a network into a low-dimensional vector, has attracted much attention during the last decade. More recently, representation learning methods based on graph neural networks (GNNs) show their superiority on various graph-based applications, and become the state-of-the-art paradigm for graph representation learning. GNNs work well for both node-level and graph-level tasks, and immensely contribute to the depth and breadth of the adoption of graph representation learning in real-world applications: ranging from classical graph-based applications such as recommender systems and social network analysis, to new frontiers such as combinational optimization, physics, and health care. The wide applications of GNNs enable diverse contributions and perspectives from disparate disciplines and make this research field truly interdisciplinary.

This book provides a comprehensive introduction to the foundations and frontiers of graph neural networks. It mainly consists of three parts: Fundamental Definitions and Development of GNNs in Part I (Chaps. 1 and 2); Frontier topics about GNNs in Part II (Chaps. 3 and 7), and Future Directions for GNNs in Part III (Chap. 8). The book starts from the basics of graph representation learning, and extensively introduces the cutting-edge research directions of GNNs, including heterogeneous GNNs, dynamic GNNs, hyperbolic GNNs, distilling GNNs, etc. The basic knowledge can help readers quickly understand the merits of GNNs, while the various topics of advanced GNNs are expected to inspire readers to develop their own models. Both beginners and experienced researchers from academia or industry are believed to benefit from the content of this book.

The authors of this book have worked on graph representation learning for years, and developed a series of fundamental algorithms. As my first visiting scholar from mainland China, Chuan Shi and I have built a close collaboration since 2010. He has done great work in heterogeneous information network analysis and promoted the development of this field. Xiao Wang and Cheng Yang, who published several top-cited papers in graph representation learning, are rising-star scholars in network analysis community. I know these excellent young researchers built a fast-rising laboratory focusing on Graph dAta Mining and MAchine learning (named GAMMA Lab, directed by Chuan Shi). This book systematically summarizes contributions of GAMMA Lab in the domain of graph neural networks. I hope you can learn from this book and enjoy it.

Philip S. Yu
UIC Distinguished Professor and Wexler
Chair in Information Technology
University of Illinois Chicago
Chicago, USA

Preface

In the era of big data, graph data has attracted considerable attention, ranging from social networks, biological networks to recommendation systems. For example, in social network, the user and their behavior can be modeled as graph; in chemistry, the molecular structure is naturally a graph; and in text analysis, the relations among words, sentences, and texts can also be modeled as a graph. Despite data may be generated from different fields with various modalities, they all can be considered as a graph, implying that graph will make a profound impact on every walk of life. Naturally, graph analysis is of great scientific and application values.

To bridge the gap between graph data with the real-world applications, one fundamental problem is the graph representation learning, i.e., how to learn the low-dimensional vector for nodes in a graph, so that the applications can be conducted on the new learned vectors instead of original graph structures. Deep learning, which has already well demonstrated their ability on other fields, e.g., computer vison, has also become a promising technique to deal with graph data. Different from previous graph representation learning which mainly focuses on preserving topology, graph neural networks learn the node representation by propagating node features along topology in a layer-by-layer manner. In this way, the learned representation naturally encodes the effective information from both of node feature and topology. To date, graph neural network has become a typical neural network in deep learning. Not surprisingly, we have witnessed the impressive performance of graph neural networks on various real-world applications, including but not limited to recommender systems and biological field. The increasing number of works on graph neural networks indicates a global trend in both academic and industrial communities. Therefore, there is a pressing demand for comprehensively summarizing and discussing the basic and advanced graph neural networks.

This book serves the interests of specific reader groups. Generally, the book is intended for anyone who wishes to understand the fundamental problems and techniques of graph neural networks. In particular, we hope that university students, researchers, and engineers in universities and IT companies will find this book inspiring. This book is divided into three parts, and the readers are able to quickly understand this field through the first part, deeply study the advanced topics of graph neural networks with the second part, and learn the future directions in the third part.

- In the first part (Chaps. 1–2), we first present an overview of the basic concepts of different graphs, and the development of graph neural networks, including several typical graph neural networks. This part will help readers rapidly understand the overall development of this field. In particular, in Chap. 1, the basic concepts and definitions, as well as the development of graph neural networks, will be summarized. The fundamental graph neural networks, including GCN, etc., will be introduced in Chap. 2.
- In the second part (Chaps. 3–7), we then provide an in-depth and detailed introduction of representative graph neural network techniques. This part will help readers understand the fundamental problems in this field, and illustrate how to design the advanced graph neural networks for these problems. In particular, the homogeneous graph neural networks are discussed in Chap. 3, including the multi-channel graph neural networks, etc. In Chap. 4, the heterogeneous graph neural networks are presented, mainly focusing on the heterogeneous graph propagation network, etc. After that, we introduce the dynamic graph neural networks in Chap. 5, which consider the temporal graph, the dynamic heterogeneous Hawkes process, and temporal heterogeneous graph. Then, in Chap. 6, we introduce the hyperbolic graph neural networks, covering the hyperbolic graph attention networks and Lorentzian graph convolutional neural networks, etc. Finally, the distilling graph neural networks are presented in Chap. 7, including the knowledge distillation of graph neural networks and adversarial knowledge distillation, etc.
- In the third part (Chap. 8), we make the conclusion and discuss the future research directions. Despite a large number of graph neural network methods being proposed, many important open problems are still not well explored, such as the robustness and the fairness of graph neural networks. When graph neural networks are applied to real-world applications, especially some risk-sensitivity areas, these problems need to be carefully considered. We summarize the future research directions here.

Writing a book always involves more people than just the authors. We would like to express our sincere thanks to all those who work with us on this book. In particular, this book is written in collaboration with BUPT GAMMA Lab, which is compiled by Chuan Shi, Xiao Wang, and Cheng Yang. The writers of each chapter include Mengmei Zhang, Meiqi Zhu, Deyu Bo, Ruijia Wang, Houye Ji, Nian Liu, Yugang Ji, Yuanfu Lu, Yiding Zhang, Jiawei Liu, Yuanxin Zhuang, Yuxin Guo, Tianyu Zhao, and Yaoqi Liu. In addition, the work is supported by National Natural Science Foundation of China (Nos. U20B2045, U1936220, 61772082, 61702296, 62002029, 62172052). We also thank the supports of these grants. Finally, we thank our families for their wholehearted support throughout this book.

Beijing, China

Chuan Shi
Xiao Wang
Cheng Yang

Contents

1 Introduction 1
1.1 Basic Concepts 1
1.1.1 Graph Definitions and Properties 1
1.1.2 Complex Graphs 3
1.1.3 Computational Tasks on Graphs 6
1.2 Development of Graph Neural Network 8
1.2.1 History of Graph Representation Learning 8
1.2.2 Frontier of Graph Neural Networks 9
1.3 Organization of the Book 10

2 Fundamental Graph Neural Networks 13
2.1 Introduction 13
2.2 Graph Convolutional Network 14
2.2.1 Overview 14
2.2.2 The GCN Method 15
2.3 Inductive Graph Convolution Network 17
2.3.1 Overview 17
2.3.2 The GraphSAGE Method 18
2.4 Graph Attention Network 20
2.4.1 Overview 21
2.4.2 The GAT Method 21
2.5 Heterogeneous Graph Attention Network 23
2.5.1 Overview 23
2.5.2 The HAN Method 24

3 Homogeneous Graph Neural Networks 27
3.1 Introduction 27
3.2 Adaptive Multi-channel Graph Convolutional Networks 28
3.2.1 Overview 28
3.2.2 Investigation 29

3.2.3 The AM-GCN Method 30
3.2.4 Experiments 34
3.3 Beyond Low-Frequency Information in Graph Convolutional Networks 36
3.3.1 Overview 36
3.3.2 Investigation 37
3.3.3 The FAGCN Method 38
3.3.4 Experiments 40
3.4 Graph Structure Estimation Neural Networks 42
3.4.1 Overview 42
3.4.2 The GEN Method 43
3.4.3 Experiments 49
3.5 Interpreting and Unifying GNNs with An Optimization Framework 50
3.5.1 Overview 50
3.5.2 Preliminary 51
3.5.3 The GNN-LF/HF Method 53
3.5.4 Experiments 56
3.6 Conclusion 58
3.7 Further Reading 58

4 Heterogeneous Graph Neural Networks 61
4.1 Introduction 61
4.2 Heterogeneous Graph Propagation Network 62
4.2.1 Overview 62
4.2.2 The HPN Method 63
4.2.3 Experiments 67
4.3 Heterogeneous Graph Neural Network with Distance Encoding 69
4.3.1 Overview 69
4.3.2 The DHN Method 70
4.3.3 Experiments 73
4.4 Self-supervised HGNN with Co-contrastive Learning 76
4.4.1 Overview 76
4.4.2 The HeCo Method 76
4.4.3 Experiments 81
4.5 Conclusion 84
4.6 Further Reading 85

5 Dynamic Graph Neural Networks 87
5.1 Introduction 87
5.2 Micro- and Macro-dynamics 88
5.2.1 Overview 88
5.2.2 The M^2DNE Method 89
5.2.3 Experiments 92

5.3 Heterogeneous Hawkes Process ... 94
5.3.1 Overview ... 94
5.3.2 The HPGE Method ... 95
5.3.3 Experiments ... 98
5.4 Dynamic Meta-Path ... 100
5.4.1 Overview ... 100
5.4.2 The DyMGNN Method ... 101
5.4.3 Experiments ... 104
5.5 Conclusion ... 107
5.6 Further Reading ... 107

6 Hyperbolic Graph Neural Networks ... 109
6.1 Introduction ... 109
6.2 Hyperbolic Graph Attention Network ... 110
6.2.1 Overview ... 110
6.2.2 The HAT Method ... 111
6.2.3 Experiments ... 114
6.3 Lorentzian Graph Convolutional Network ... 116
6.3.1 Overview ... 116
6.3.2 The LGCN Method ... 117
6.3.3 Experiments ... 120
6.4 Hyperbolic Heterogeneous Graph Representation ... 122
6.4.1 Overview ... 122
6.4.2 The HHNE Method ... 124
6.4.3 Experiments ... 126
6.5 Conclusion ... 128
6.6 Further Reading ... 129

7 Distilling Graph Neural Networks ... 131
7.1 Introduction ... 131
7.2 Prior-Enhanced Knowledge Distillation for GNNs ... 132
7.2.1 Overview ... 132
7.2.2 The CPF Method ... 132
7.2.3 Experiments ... 136
7.3 Temperature-Adaptive Knowledge Distillation for GNNs ... 137
7.3.1 Overview ... 137
7.3.2 The LTD Method ... 139
7.3.3 Experiments ... 142
7.4 Data-Free Adversarial Knowledge Distillation for GNNs ... 144
7.4.1 Overview ... 144
7.4.2 The DFAD-GNN Method ... 144
7.4.3 Experiments ... 147

7.5 Conclusion . . . 150
7.6 Further Reading . . . 151

8 Platforms and Practice of Graph Neural Networks . . . 153
8.1 Introduction . . . 153
8.2 Foundation . . . 154
8.2.1 Deep Learning Platforms . . . 154
8.2.2 Platforms of Graph Neural Networks . . . 159
8.2.3 GammaGL . . . 162
8.3 Practice of Graph Neural Networks on GammaGL . . . 165
8.3.1 Create Your Own Graph . . . 166
8.3.2 Create Message-Passing Network . . . 167
8.3.3 Advanced Mini-Batching . . . 168
8.3.4 Practice of GIN . . . 169
8.3.5 Practice of GraphSAGE . . . 171
8.3.6 Practice of HAN . . . 174
8.4 Conclusion . . . 177

9 Future Direction and Conclusion . . . 179
9.1 Future Direction . . . 179
9.1.1 Self-supervised Learning on Graphs . . . 179
9.1.2 Robustness . . . 180
9.1.3 Explainability . . . 180
9.1.4 Fairness . . . 181
9.1.5 Biochemistry . . . 182
9.1.6 Physics . . . 182
9.2 Conclusion . . . 183

References . . . 185