

Contents

Terminologies	xxx
1 Basic concepts of Graphs	xxx
2 Machine Learning on Graphs	xxx
3 Graph Neural Networks	xxx
Notations	xxx
Part I Introduction	
1 Representation Learning	3
Liang Zhao, Lingfei Wu, Peng Cui and Jian Pei	
1.1 Representation Learning: An Introduction	3
1.2 Representation Learning in Different Areas	5
1.2.1 Representation Learning for Image Processing	5
1.2.2 Representation Learning for Speech Recognition	8
1.2.3 Representation Learning for Natural Language Processing	10
1.2.4 Representation Learning for Networks	13
1.3 Summary	14
2 Graph Representation Learning	17
Peng Cui, Lingfei Wu, Jian Pei, Liang Zhao and Xiao Wang	
2.1 Graph Representation Learning: An Introduction	17
2.2 Traditional Graph Embedding	19
2.3 Modern Graph Embedding	20
2.3.1 Structure-Property Preserving Graph Representation Learning	20
2.3.2 Graph Representation Learning with Side Information	23
2.3.3 Advanced Information Preserving Graph Representation Learning	24
2.4 Graph Neural Networks	25
2.5 Summary	26

3	Graph Neural Networks	27
	Lingfei Wu, Peng Cui, Jian Pei, Liang Zhao and Le Song	
3.1	Graph Neural Networks: An Introduction	28
3.2	Graph Neural Networks: Overview	29
3.2.1	Graph Neural Networks: Foundations	29
3.2.2	Graph Neural Networks: Frontiers	31
3.2.3	Graph Neural Networks: Applications	33
3.2.4	Graph Neural Networks: Organization	35
3.3	Summary	36
Part II Foundations of Graph Neural Networks		
4	Graph Neural Networks for Node Classification	41
	Jian Tang and Renjie Liao	
4.1	Background and Problem Definition	41
4.2	Supervised Graph Neural Networks	42
4.2.1	General Framework of Graph Neural Networks	43
4.2.2	Graph Convolutional Networks	44
4.2.3	Graph Attention Networks	46
4.2.4	Neural Message Passing Networks	48
4.2.5	Continuous Graph Neural Networks	48
4.2.6	Multi-Scale Spectral Graph Convolutional Networks	51
4.3	Unsupervised Graph Neural Networks	54
4.3.1	Variational Graph Auto-Encoders	54
4.3.2	Deep Graph Infomax	57
4.4	Over-smoothing Problem	59
4.5	Summary	61
5	The Expressive Power of Graph Neural Networks	63
	Pan Li and Jure Leskovec	
5.1	Introduction	63
5.2	Graph Representation Learning and Problem Formulation	67
5.3	The Power of Message Passing Graph Neural Networks	70
5.3.1	Preliminaries: Neural Networks for Sets	70
5.3.2	Message Passing Graph Neural Networks	71
5.3.3	The Expressive Power of MP-GNN	72
5.3.4	MP-GNN with the Power of the 1-WL Test	75
5.4	Graph Neural Networks Architectures that are more Powerful than 1-WL Test	77
5.4.1	Limitations of MP-GNN	77
5.4.2	Injecting Random Attributes	79
5.4.3	Injecting Deterministic Distance Attributes	86
5.4.4	Higher-order Graph Neural Networks	92
5.5	Summary	97

6 Graph Neural Networks: Scalability	99
Hehuan Ma, Yu Rong, and Junzhou Huang	
6.1 Introduction	100
6.2 Preliminary	101
6.3 Sampling Paradigms	101
6.3.1 Node-wise Sampling	103
6.3.2 Layer-wise Sampling	106
6.3.3 Graph-wise Sampling	111
6.4 Applications of Large-scale Graph Neural Networks on Recommendation Systems	115
6.4.1 Item-item Recommendation	116
6.4.2 User-item Recommendation	116
6.5 Future Directions	118
7 Interpretability in Graph Neural Networks	121
Ninghao Liu and Qizhang Feng and Xia Hu	
7.1 Background: Interpretability in Deep Models	121
7.1.1 Definition of Interpretability and Interpretation	122
7.1.2 The Value of Interpretation	123
7.1.3 Traditional Interpretation Methods	124
7.1.4 Opportunities and Challenges	127
7.2 Explanation Methods for Graph Neural Networks	128
7.2.1 Background	128
7.2.2 Approximation-Based Explanation	130
7.2.3 Relevance-Propagation Based Explanation	134
7.2.4 Perturbation-Based Approaches	135
7.2.5 Generative Explanation	137
7.3 Interpretable Modeling on Graph Neural Networks	138
7.3.1 GNN-Based Attention Models	138
7.3.2 Disentangled Representation Learning on Graphs	141
7.4 Evaluation of Graph Neural Networks Explanations	143
7.4.1 Benchmark Datasets	143
7.4.2 Evaluation Metrics	145
7.5 Future Directions	146
8 Graph Neural Networks: Adversarial Robustness	149
Stephan Günnemann	
8.1 Motivation	149
8.2 Limitations of Graph Neural Networks: Adversarial Examples	152
8.2.1 Categorization of Adversarial Attacks	152
8.2.2 The Effect of Perturbations and Some Insights	156
8.2.3 Discussion and Future Directions	159
8.3 Provable Robustness: Certificates for Graph Neural Networks	160
8.3.1 Model-Specific Certificates	160
8.3.2 Model-Agnostic Certificates	163
8.3.3 Advanced Certification and Discussion	165

8.4	Improving Robustness of Graph Neural Networks	165
8.4.1	Improving the Graph	166
8.4.2	Improving the Training Procedure	167
8.4.3	Improving the Graph Neural Networks' Architecture	170
8.4.4	Discussion and Future Directions	171
8.5	Proper Evaluation in the View of Robustness	172
8.6	Summary	175

Part III Frontiers of Graph Neural Networks

9	Graph Neural Networks: Graph Classification	179
	Christopher Morris	
9.1	Introduction	179
9.2	Graph neural networks for graph classification: Classic works and modern architectures	180
9.2.1	Spatial approaches	181
9.2.2	Spectral approaches	184
9.3	Pooling layers: Learning graph-level outputs from node-level outputs	186
9.3.1	Attention-based pooling layers	187
9.3.2	Cluster-based pooling layers	187
9.3.3	Other pooling layers	188
9.4	Limitations of graph neural networks and higher-order layers for graph classification	189
9.4.1	Overcoming limitations	190
9.5	Applications of graph neural networks for graph classification	191
9.6	Benchmark Datasets	192
9.7	Summary	192
10	Graph Neural Networks: Link Prediction	195
	Muhan Zhang	
10.1	Introduction	195
10.2	Traditional Link Prediction Methods	197
10.2.1	Heuristic Methods	197
10.2.2	Latent-Feature Methods	200
10.2.3	Content-Based Methods	203
10.3	GNN Methods for Link Prediction	203
10.3.1	Node-Based Methods	203
10.3.2	Subgraph-Based Methods	206
10.3.3	Comparing Node-Based Methods and Subgraph-Based Methods	209
10.4	Theory for Link Prediction	211
10.4.1	γ -Decaying Heuristic Theory	211
10.4.2	Labeling Trick	217
10.5	Future Directions	220
10.5.1	Accelerating Subgraph-Based Methods	220

10.5.2	Designing More Powerful Labeling Tricks	221
10.5.3	Understanding When to Use One-Hot Features	222
11	Graph Neural Networks: Graph Generation	225
	Renjie Liao	
11.1	Introduction	225
11.2	Classic Graph Generative Models	226
11.2.1	Erdős–Rényi Model	226
11.2.2	Stochastic Block Model	228
11.3	Deep Graph Generative Models	229
11.3.1	Representing Graphs	230
11.3.2	Variational Auto-Encoder Methods	230
11.3.3	Deep Autoregressive Methods	236
11.3.4	Generative Adversarial Methods	244
11.4	Summary	250
12	Graph Neural Networks: Graph Transformation	251
	Xiaojie Guo, Shiyu Wang, Liang Zhao	
12.1	Problem Formulation of Graph Transformation	252
12.2	Node-level Transformation	253
12.2.1	Definition of Node-level Transformation	253
12.2.2	Interaction Networks	253
12.2.3	Spatio-Temporal Convolution Recurrent Neural Networks	254
12.3	Edge-level Transformation	256
12.3.1	Definition of Edge-level Transformation	256
12.3.2	Graph Transformation Generative Adversarial Networks	257
12.3.3	Multi-scale Graph Transformation Networks	259
12.3.4	Graph Transformation Policy Networks	260
12.4	Node-Edge Co-Transformation	261
12.4.1	Definition of Node-Edge Co-Transformation	261
12.4.2	Editing-based Node-Edge Co-Transformation	266
12.5	Other Graph-based Transformations	271
12.5.1	Sequence-to-Graph Transformation	271
12.5.2	Graph-to-Sequence Transformation	272
12.5.3	Context-to-Graph Transformation	273
12.6	Summary	275
13	Graph Neural Networks: Graph Matching	277
	Xiang Ling, Lingfei Wu, Chunming Wu and Shouling Ji	
13.1	Introduction	278
13.2	Graph Matching Learning	279
13.2.1	Problem Definition	280
13.2.2	Deep Learning based Models	282
13.2.3	Graph Neural Network based Models	284
13.3	Graph Similarity Learning	288
13.3.1	Problem Definition	288

13.3.2	Graph-Graph Regression Tasks	290
13.3.3	Graph-Graph Classification Tasks	293
13.4	Summary	295
14	Graph Neural Networks: Graph Structure Learning	297
	Yu Chen and Lingfei Wu	
14.1	Introduction	297
14.2	Traditional Graph Structure Learning	299
14.2.1	Unsupervised Graph Structure Learning	299
14.2.2	Supervised Graph Structure Learning	301
14.3	Graph Structure Learning for Graph Neural Networks	303
14.3.1	Joint Graph Structure and Representation Learning	304
14.3.2	Connections to Other Problems	317
14.4	Future Directions	319
14.4.1	Robust Graph Structure Learning	319
14.4.2	Scalable Graph Structure Learning	320
14.4.3	Graph Structure Learning for Heterogeneous Graphs	320
14.5	Summary	320
15	Dynamic Graph Neural Networks	323
	Seyed Mehran Kazemi	
15.1	Introduction	323
15.2	Background and Notation	325
15.2.1	Graph Neural Networks	325
15.2.2	Sequence Models	327
15.2.3	Encoder-Decoder Framework and Model Training	330
15.3	Categories of Dynamic Graphs	331
15.3.1	Discrete vs. Continues	331
15.3.2	Types of Evolution	333
15.3.3	Prediction Problems, Interpolation, and Extrapolation	334
15.4	Modeling Dynamic Graphs with Graph Neural Networks	335
15.4.1	Conversion to Static Graphs	335
15.4.2	Graph Neural Networks for DTDGs	337
15.4.3	Graph Neural Networks for CTDGs	340
15.5	Applications	343
15.5.1	Skeleton-based Human Activity Recognition	343
15.5.2	Traffic Forecasting	345
15.5.3	Temporal Knowledge Graph Completion	346
15.6	Summary	348
16	Heterogeneous Graph Neural Networks	351
	Chuan Shi	
16.1	Introduction to HGNNs	351
16.1.1	Basic Concepts of Heterogeneous Graphs	353
16.1.2	Challenges of HG Embedding	354
16.1.3	Brief Overview of Current Development	355

16.2	Shallow Models	356
16.2.1	Decomposition-based Methods	357
16.2.2	Random Walk-based Methods	358
16.3	Deep Models	360
16.3.1	Message Passing-based Methods (HGNNs)	360
16.3.2	Encoder-decoder-based Methods	363
16.3.3	Adversarial-based Methods	364
16.4	Review	366
16.5	Future Directions	367
16.5.1	Structures and Properties Preservation	367
16.5.2	Deeper Exploration	367
16.5.3	Reliability	368
16.5.4	Applications	369
17	Graph Neural Networks: AutoML	371
	Kaixiong Zhou, Zirui Liu, Keyu Duan and Xia Hu	
17.1	Background	372
17.1.1	Notations of AutoGNN	373
17.1.2	Problem Definition of AutoGNN	375
17.1.3	Challenges in AutoGNN	375
17.2	Search Space	376
17.2.1	Architecture Search Space	377
17.2.2	Training Hyperparameter Search Space	380
17.2.3	Efficient Search Space	381
17.3	Search Algorithms	382
17.3.1	Random Search	382
17.3.2	Evolutionary Search	382
17.3.3	Reinforcement Learning Based Search	383
17.3.4	Differentiable Search	385
17.3.5	Efficient Performance Estimation	386
17.4	Future Directions	387
18	Graph Neural Networks: Self-supervised Learning	391
	Yu Wang, Wei Jin, and Tyler Derr	
18.1	Introduction	392
18.2	Self-supervised Learning	393
18.3	Applying SSL to Graph Neural Networks: Categorizing Training Strategies, Loss Functions and Pretext Tasks	395
18.3.1	Training Strategies	396
18.3.2	Loss Functions	399
18.3.3	Pretext Tasks	402
18.4	Node-level SSL Pretext Tasks	403
18.4.1	Structure-based Pretext Tasks	403
18.4.2	Feature-based Pretext Tasks	404
18.4.3	Hybrid Pretext Tasks	406
18.5	Graph-level SSL Pretext Tasks	408

18.5.1	Structure-based Pretext Tasks	408
18.5.2	Feature-based Pretext Tasks	413
18.5.3	Hybrid Pretext Tasks	414
18.6	Node-graph-level SSL Pretext Tasks	417
18.7	Discussion	418
18.8	Summary	419

Part IV Broad and Emerging Applications with Graph Neural Networks

19	Graph Neural Networks in Modern Recommender Systems	423
	Yunfei Chu, Jiangchao Yao, Chang Zhou and Hongxia Yang	
19.1	Graph Neural Networks for Recommender System in Practice	423
19.1.1	Introduction	423
19.1.2	Classic Approaches to Predict User-Item Preference	428
19.1.3	Item Recommendation in user-item Recommender Systems: a Bipartite Graph Perspective	429
19.2	Case Study 1: Dynamic Graph Neural Networks Learning	431
19.2.1	Dynamic Sequential Graph	431
19.2.2	DSGL: Dynamic Sequential Graph Learning	432
19.2.3	Model Prediction	435
19.2.4	Experiments and Discussions	436
19.3	Case Study 2: Device-Cloud Collaborative Learning for Graph Neural Networks	438
19.3.1	The proposed framework	438
19.3.2	Experiments and Discussions	442
19.4	Future Directions	444
20	Graph Neural Networks in Computer Vision	447
	Siliang Tang, Wenqiao Zhang, Zongshen Mu, Kai Shen, Juncheng Li, Jiacheng Li and Lingfei Wu	
20.1	Introduction	448
20.2	Representing Vision as Graphs	448
20.2.1	Visual Node representation	448
20.2.2	Visual Edge representation	450
20.3	Case Study 1: Image	451
20.3.1	Object Detection	451
20.3.2	Image Classification	453
20.4	Case Study 2: Video	454
20.4.1	Video Action Recognition	454
20.4.2	Temporal Action Localization	456
20.5	Other Related Work: Cross-media	457
20.5.1	Visual Caption	457
20.5.2	Visual Question Answering	458
20.5.3	Cross-Media Retrieval	459
20.6	Frontiers for Graph Neural Networks on Computer Vision	460
20.6.1	Advanced Graph Neural Networks for Computer Vision	460

20.6.2	Broader Area of Graph Neural Networks on Computer Vision	461
20.7	Summary	462
21	Graph Neural Networks in Natural Language Processing	463
	Bang Liu, Lingfei Wu	
21.1	Introduction	463
21.2	Modeling Text as Graphs	466
21.2.1	Graph Representations in Natural Language Processing	466
21.2.2	Tackling Natural Language Processing Tasks from a Graph Perspective	468
21.3	Case Study 1: Graph-based Text Clustering and Matching	470
21.3.1	Graph-based Clustering for Hot Events Discovery and Organization	470
21.3.2	Long Document Matching with Graph Decomposition and Convolution	473
21.4	Case Study 2: Graph-based Multi-Hop Reading Comprehension	475
21.5	Future Directions	479
21.6	Conclusions	480
22	Graph Neural Networks in Program Analysis	483
	Miltiadis Allamanis	
22.1	Introduction	483
22.2	Machine Learning in Program Analysis	484
22.3	A Graph Representation of Programs	486
22.4	Graph Neural Networks for Program Graphs	489
22.5	Case Study 1: Detecting Variable Misuse Bugs	491
22.6	Case Study 2: Predicting Types in Dynamically Typed Languages	493
22.7	Future Directions	495
23	Graph Neural Networks in Software Mining	499
	Collin McMillan	
23.1	Introduction	499
23.2	Modeling Software as a Graph	500
23.2.1	Macro versus Micro Representations	501
23.2.2	Combining the Macro- and Micro-level	503
23.3	Relevant Software Mining Tasks	503
23.4	Example Software Mining Task: Source Code Summarization	504
23.4.1	Primer GNN-based Code Summarization	505
23.4.2	Directions for Improvement	510
23.5	Summary	512
24	GNN-based Biomedical Knowledge Graph Mining in Drug Development	517
	Chang Su, Yu Hou, Fei Wang	
24.1	Introduction	517

24.2	Existing Biomedical Knowledge Graphs	518
24.3	Inference on Knowledge Graphs	523
24.3.1	Conventional KG inference techniques	523
24.3.2	GNN-based KG inference techniques	524
24.4	KG-based hypothesis generation in computational drug development	528
24.4.1	A machine learning framework for KG-based drug repurposing	529
24.4.2	Application of KG-based drug repurposing in COVID-19	530
24.5	Future directions	531
24.5.1	KG quality control	532
24.5.2	Scalable inference	533
24.5.3	Coupling KGs with other biomedical data	533
25	Graph Neural Networks in Predicting Protein Function and Interactions	541
	Anowarul Kabir and Amarda Shehu	
25.1	From Protein Interactions to Function: An Introduction	541
25.1.1	Enter Stage Left: Protein-Protein Interaction Networks	542
25.1.2	Problem Formulation(s), Assumptions, and Noise: A Historical Perspective	543
25.1.3	Shallow Machine Learning Models over the Years	543
25.1.4	Enter Stage Right: Graph Neural Networks	544
25.2	Highlighted Case Studies	547
25.2.1	Case Study 1: Prediction of Protein-Protein and Protein-Drug Interactions: The Link Prediction Problem	547
25.2.2	Case Study 2: Prediction of Protein Function and Functionally-important Residues	549
25.2.3	Case Study 3: From Representation Learning to Multirelational Link Prediction in Biological Networks with Graph Autoencoders	553
25.3	Future Directions	555
26	Graph Neural Networks in Anomaly Detection	557
	Shen Wang, Philip S. Yu	
26.1	Introduction	557
26.2	Issues	561
26.2.1	Data-specific issues	561
26.2.2	Task-specific Issues	563
26.2.3	Model-specific Issues	563
26.3	Pipeline	564
26.3.1	Graph Construction and Transformation	564
26.3.2	Graph Representation Learning	565
26.3.3	Prediction	567
26.4	Taxonomy	568
26.5	Case Studies	568

26.5.1	Case Study 1: Graph Embeddings for Malicious Accounts Detection	569
26.5.2	Case Study 2: Hierarchical Attention Mechanism based Cash-out User Detection	570
26.5.3	Case Study 3: Attentional Heterogeneous Graph Neural Networks for Malicious Program Detection	572
26.5.4	Case Study 4: Graph Matching Framework to Learn the Program Representation and Similarity Metric via Graph Neural Networks for Unknown Malicious Program Detection	573
26.5.5	Case Study 5: Anomaly Detection in Dynamic Graph Using Attention-based Temporal GCN	575
26.5.6	Case Study 6: GCN-based Anti-Spam for Spam Review Detection	576
26.6	Future Directions	577
27	Graph Neural Networks in Urban Intelligence	579
	Yanhua Li, Xun Zhou, and Menghai Pan	
27.1	Graph Neural Networks for Urban Intelligence	580
27.1.1	Introduction	580
27.1.2	Application scenarios in urban intelligence	581
27.1.3	Representing urban systems as graphs	584
27.1.4	Case Study 1: Graph Neural Networks in urban configuration and transportation	586
27.1.5	Case Study 2: Graph Neural Networks in urban anomaly and event detection	589
27.1.6	Case Study 3: Graph Neural Networks in urban human behavior inference	590
27.1.7	Future Directions	592
	References	595