

*C02029: Doctor of Philosophy*

*CRICOS Code: 009469A*

*Subject Code: 33874*

*August 2019*

*Network Embedding Learning  
in Knowledge Graph*

---

*Zili Zhou*

School of Computer Science

Faculty of Engineering and Information Technology

University of Technology Sydney

NSW - 2007, Australia



---

---

# Network Embedding Learning in Knowledge Graph

---

---

*A thesis submitted in partial fulfilment of the requirements  
for the degree of*

Doctor of Philosophy

*in*

Analytics

*by*

**Zili Zhou**

*to*

School of Computer Science  
Faculty of Engineering and Technology  
University of Technology Sydney  
NSW - 2007, Australia

August 2019



## ABSTRACT

**K**nowledge Graph stores a large number of human knowledge facts in form of multi-relational network structure, is widely used as a core technique in real-world applications including search engine, question answering system, and recommender system. Knowledge Graph is used to provide extra info box for user query in Google search engine, the WolframAlpha site provides question answering service relying on Knowledge Graph, and the eBay uses Knowledge Graph as semantic enhance for their recommendation service.

Motivated by several characteristics of Knowledge Graph including incompleteness, structural inferability, and semantical application enhancement, a few efforts have been put into the Knowledge Graph analysis area. Some works contribute to Knowledge Graph construction and maintenance through crowdsourcing. Some previous network embedding learning models show good performance on homogeneous network analysis, while the performance of directly using them on Knowledge Graph is limited because the multiple relationship information of the Knowledge Graph is ignored. Then, the concept of Knowledge Graph embedding learning is given, by learning representation for Knowledge Graph components including entities and relations, the latent semantic information is extracted into embedding representation. And the embedding techniques are also utilized in collaborative learning for Knowledge Graph and external application scenarios, the target is to use Knowledge Graph as a semantic enhancement to improve the performance of external applications.

However, some problems still remain in Knowledge Graph completion, reasoning, and external application. First, a proper model is required for Knowledge Graph self-completion, and a proper integration solution is also required to add extra conceptual taxonomy information into the process of Knowledge Graph completion. Then, a framework to use sub-structure information of Knowledge Graph network into knowledge reasoning is needed. After that, a collaborative learning framework for knowledge graph completion and downstream machine learning tasks is needed to be designed. In this thesis, we take recommender systems as an example of downstream machine learning tasks.

To address the aforementioned research problems, a few approaches are proposed in the works introduced in this thesis.

- A bipartite graph embedding based Knowledge Graph completion approach for Knowledge Graph self-completion, each knowledge fact is represented in the form of bipartite graph structure for more reasonable triple inference.

- 
- An embedding based cross completion approach for completing the factual Knowledge Graph with additive conceptual taxonomy information, the components of factual Knowledge Graph and conceptual taxonomy, entities, relations, types, are jointly represented by embedding representation.
  - Two sub-structure based Knowledge Graph transitive relation embedding approaches for knowledge reasoning analysis based on Knowledge Graph sub-structure, the transitive structural information contained in Knowledge Graph network sub-structure is learned into relation embedding.
  - Two hierarchical collaborative embedding approaches for proper collaborative learning on Knowledge Graph and Recommender System through linking Knowledge Graph entities with Recommender items, then entities, relations, items, and users are represented by embedding in collaborative space.

The main contributions of this thesis are proposing a few approaches which can be used in multiple Knowledge Graph related domains, Knowledge Graph completion, reasoning and application. Two approaches achieve more accurate Knowledge Graph completion, other two approaches model knowledge reasoning based on network sub-structure analysis, and the other approaches apply Knowledge Graph into a recommender system application.

## CERTIFICATE OF ORIGINAL AUTHORSHIP

I, Zili Zhou declare that this thesis, is submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the School of Computer Science at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise reference or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution. I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of the requirements for a degree at any other academic institution except as fully acknowledged within the text. This thesis is the result of a Collaborative Doctoral Research Degree program with Shanghai University.

This research is supported by the Australian Government Research Training Program.

Production Note:

SIGNATURE: Signature removed prior to publication.

Zili Zhou

DATE: 21<sup>st</sup> August, 2019



## DEDICATION

*To my beloved wife ...*



## ACKNOWLEDGMENTS

First of all, I would like to thank Prof Guandong Xu, my supervisor, for his guidance, suggestion, and support throughout my doctoral program course at the University of Technology Sydney. Without his professional guidance and support, this work would not have been achieved.

I am also grateful to my co-supervisor Prof Jinyan Li for his constructive suggestions on my research works. I would like to pass my gratitude to my research mate and friend Dr. Shaowu Liu, for his help and suggestions on my works.

I am thankful to my beloved wife, Xiu Yan, and my baby sons, Yinuo Zhou and Yiheng Zhou, for their love and encouragement which helps me finish this thesis. I am also thankful to my parents and my parents in law for their support throughout my research studies.

Zili Zhou  
Sydney, Australia  
August, 2019



## LIST OF PUBLICATIONS

### RELATED TO THE THESIS :

1. Zhou, Z., Xu, G., Zhu, W., Li, J., & Zhang, W. (2017, May). Structure embedding for knowledge base completion and analytics. In 2017 International Joint Conference on Neural Networks (IJCNN) (pp. 737-743). IEEE.
2. Zhou, Z., Liu, S., Xu, G., Xie, X., Yin, J., Li, Y., & Zhang, W. (2018, June). Knowledge-Based Recommendation with Hierarchical Collaborative Embedding. In Pacific-Asia Conference on Knowledge Discovery and Data Mining (pp. 222-234). Springer, Cham.
3. Zhou, Z., Liu, S., Xu, G., & Zhang, W. (2019, July). On Completing Sparse Knowledge Base with Transitive Relation Embedding. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 33, pp. 3125-3132).
4. Zhou, Z., Liu, S., Xu, G. & Zhang, W. Meta-structure Transitive Relation Embedding for Knowledge Graph Completion. Prepared to be submitted as a Conference Paper.
5. Zhou, Z., Liu, S., Xu, G. & Zhang, W. Cross Completion for Factual Knowledge Graph and Conceptual Taxonomy. Prepared to be submitted as a Journal Paper.
6. Zhou, Z., Liu, S., Xu, G. & Zhang, W. Knowledge Graph based Collaborative Network Embedding for Recommender System. Prepared to be submitted as a Journal Paper.

### OTHERS :

7. Zhou, Z., Xu, G., Zhu, X., & Liu, S. (2017, October). Latent factor analysis for low-dimensional implicit preference prediction. In 2017 International Conference on Behavioral, Economic, Socio-cultural Computing (BESC) (pp. 1-2). IEEE.

- 
8. Liu, S., Xu, G., Zhu, X., & Zhou, Z. (2017, October). Towards simplified insurance application via sparse questionnaire optimization. In 2017 International Conference on Behavioral, Economic, Socio-Cultural Computing (BESC) (pp. 1-2). IEEE.
  9. Yin, J., Zhou, Z., Liu, S., Wu, Z., & Xu, G. (2018, June). Social Spammer Detection: A Multi-Relational Embedding Approach. In Pacific-Asia Conference on Knowledge Discovery and Data Mining (pp. 615-627). Springer, Cham.

# TABLE OF CONTENTS

<b>List of Publications</b>	<b>ix</b>
<b>List of Figures</b>	<b>xv</b>
<b>List of Tables</b>	<b>xvii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motivation . . . . .	2
1.2 Research Objectives . . . . .	4
1.3 Research Problems . . . . .	4
1.4 Proposed Approaches . . . . .	6
1.5 Thesis Outlines . . . . .	10
<b>2 Literature Review</b>	<b>13</b>
2.1 Knowledge Graph and Semantic Web . . . . .	14
2.2 Network Embedding Representation Learning . . . . .	17
2.3 Knowledge Graph Embedding . . . . .	19
2.3.1 Translational Distance Models . . . . .	22
2.3.2 Semantic Matching Models . . . . .	22
2.3.3 Relation-based Models . . . . .	24
2.4 Knowledge Graph based Application Systems . . . . .	24
<b>3 Bipartite Graph Embedding based Knowledge Graph Completion</b>	<b>27</b>
3.1 Introduction . . . . .	27
3.2 Preliminaries . . . . .	29
3.2.1 Structured Embedding (SE) and TransE . . . . .	29
3.2.2 Other Models . . . . .	29
3.3 Methodology . . . . .	30
3.3.1 Structured Embedding Framework . . . . .	30

## TABLE OF CONTENTS

---

3.3.2	Bipartite Graph Network Architecture . . . . .	32
3.3.3	Training . . . . .	34
3.3.4	Latent Information . . . . .	36
3.4	Experiment and Analysis . . . . .	37
3.5	Summary . . . . .	38
<b>4</b>	<b>Embedding based Cross Completion for Factual Knowledge Base and Conceptual Taxonomy</b>	<b>39</b>
4.1	Introduction . . . . .	39
4.2	Preliminaries . . . . .	41
4.2.1	Factual Knowledge Graph and Conceptual Taxonomy . . . . .	41
4.2.2	Structure Embedding based Factual Knowledge Graph Completion	42
4.2.3	Collaborative Filtering based Conceptual Taxonomy Completion .	42
4.2.4	Problem Definition . . . . .	43
4.3	Methodology . . . . .	43
4.3.1	Integration of Factual Knowledge Graph and Conceptual Taxonomy	43
4.3.2	Collaborative Embedding Model . . . . .	44
4.3.3	Adaptive True Knowledge Prediction . . . . .	47
4.4	Experiment and Analysis . . . . .	49
4.4.1	Dataset . . . . .	49
4.4.2	Baselines . . . . .	49
4.4.3	Comparison . . . . .	50
4.5	Summary . . . . .	51
<b>5</b>	<b>Sub-structure based Knowledge Graph Transitive Relation Embedding</b>	<b>53</b>
5.1	Introduction . . . . .	53
5.2	Preliminaries . . . . .	57
5.2.1	Knowledge Graph Completion . . . . .	57
5.2.2	Embedding-based Knowledge Graph Completion . . . . .	57
5.2.3	Relation-based Knowledge Graph Completion . . . . .	58
5.2.4	Limitations of Current Models . . . . .	58
5.3	On Completing Sparse Knowledge Base with Transitive Relation Embedding	58
5.3.1	Triangle Pattern . . . . .	59
5.3.2	Transitive Relation Embedding . . . . .	61
5.3.3	Training . . . . .	62
5.3.4	Joint Prediction Strategy . . . . .	63

5.3.5	Advantages of Proposed Model . . . . .	64
5.4	Meta-structure based Transitive Relation Embedding for Knowledge Graph Completion . . . . .	65
5.4.1	Meta-structure . . . . .	66
5.4.2	Transitive Relation Embedding . . . . .	67
5.4.3	Training and Prediction . . . . .	69
5.4.4	Joint Prediction Strategy . . . . .	71
5.4.5	Advantages of Proposed Model . . . . .	72
5.5	Experiment and Analysis . . . . .	73
5.5.1	Entity Link Prediction . . . . .	75
5.5.2	Relation Link Prediction . . . . .	77
5.5.3	Accurate Prediction on Extremely Sparse KG with Transitive Relation Embedding . . . . .	81
5.6	Summary . . . . .	82
<b>6</b>	<b>Hierarchical Collaborative Embedding for Knowledge Graph and Recommender System</b>	<b>85</b>
6.1	Introduction . . . . .	85
6.2	Preliminaries . . . . .	90
6.2.1	Implicit Feedback Recommendation . . . . .	90
6.2.2	Collaborative Filtering using Implicit Feedback . . . . .	91
6.2.3	Knowledge Graph . . . . .	91
6.3	Hierarchical Collaborative Embedding . . . . .	92
6.3.1	Knowledge Graph Structured Embedding . . . . .	92
6.3.2	Knowledge Conceptual Level Connection . . . . .	93
6.3.3	Collaborative Learning . . . . .	95
6.4	Collaborative Network Embedding with Different Measurement Scale . .	95
6.4.1	Proximity Probability Embedding on RS Bipartite Network . . . . .	97
6.4.2	Proximity Probability Embedding on KG Multi-relational Network	97
6.4.3	Collaborative Network Embedding . . . . .	98
6.4.4	Advantages of Proposed Model . . . . .	99
6.5	Experiment and Analysis . . . . .	100
6.5.1	Dataset . . . . .	100
6.5.2	Baselines . . . . .	102
6.5.3	Comparison . . . . .	103

TABLE OF CONTENTS

---

6.6 Summary . . . . .	105
<b>7 Conclusion and Future Work</b>	<b>107</b>
7.1 Contributions . . . . .	107
7.2 Possible Future Work . . . . .	110
<b>Bibliography</b>	<b>111</b>

## LIST OF FIGURES

<b>FIGURE</b>	<b>Page</b>
1.1 Approach Framework Structure . . . . .	7
3.1 Bipartite Graph Network Architecture . . . . .	32
3.2 The process of getting one unit value of predict right entity vector by using weight sum of all units in left entity vector, all the units in predict right entity vector are computed as the process described in figure. . . . .	33
3.3 The process of getting one unit value of predict left entity vector by using weight sum of all units in right entity vector, all the units in predict left entity vector are computed as the process described in figure. . . . .	34
4.1 Cross Completion System for Factual Knowledge Graph and Conceptual Taxonomy . . . . .	41
4.2 Completing Factual Knowledge Graph with information in Conceptual Taxonomy . . . . .	45
4.3 Completing Conceptual Taxonomy with information in Factual Knowledge Graph . . . . .	45
5.1 Basic Idea of Triangle Pattern . . . . .	54
5.2 Example of Meta-structure . . . . .	56
5.3 Meta-structure . . . . .	56
5.4 Triangle Pattern with Formulated Relation Definition . . . . .	59
5.5 Triangle Relation Pattern in Meta-structure . . . . .	67
5.6 Hit@10 Result Comparison between MSTRE and MSTRE+Baselines . . . . .	80
5.7 MRR Result Comparison between MSTRE and MSTRE+Baselines . . . . .	80
6.1 Conceptual Level of Hierarchical Collaborative Embedding . . . . .	87
6.2 Structure of Collaborative Network Embedding . . . . .	88



## LIST OF TABLES

TABLE	Page
2.1 List of popular Knowledge Graph Embedding methods . . . . .	21
3.1 Ranking results on FB200 (146 entities) . . . . .	37
3.2 Ranking results on FB500 (485 entities) . . . . .	37
4.1 Precision, Recall and F1-score results of Cross Completion on Factual Knowledge Graph and Conceptual Taxonomy . . . . .	49
5.1 Relation Inference Example . . . . .	63
5.2 Space and Time Complexity of TRE and baselines . . . . .	64
5.3 Space and Time Complexity of MSTRE and baselines . . . . .	72
5.4 Dataset Size of FB15k, WN18 and DBP . . . . .	74
5.5 Result of FB15K Entity Prediction with TRE . . . . .	76
5.6 Result of WN18 Entity Prediction with TRE . . . . .	76
5.7 Result of FB15K Entity Prediction with MSTRE . . . . .	76
5.8 Result of FB15K, WN18 and DBP Relation Prediction with TRE $(h,?,t)$ . . . . .	78
5.9 Result of FB15K, WN18 and DBpedia Relation Prediction with MSTRE $(h,?,t)$ . . . . .	78
5.10 Result of FB15K, WN18 and DBpedia Relation Prediction $(h,?,t)$ for Entity Pairs with MSTRE . . . . .	81
5.11 Result of Sparse FB15K Relation Prediction with TRE . . . . .	82
6.1 Dataset Size of Movielens and Book-crossing . . . . .	101
6.2 Mean MAP and Recall results of Hierarchical Collaborative Embedding on Github Dataset . . . . .	104
6.3 Precision and Recall results of Collaborative Network Embedding on Movielens Dataset . . . . .	104
6.4 Precision and Recall results of Collaborative Network Embedding on BookCrossing Dataset . . . . .	105

