

# Denghui Zhang

Institute of Computing Technology, Chinese Academy of Sciences, No.6 Kexueyuan South Road, Beijing, China  
☎ (+86) 13269365613 | ✉ zhangdenghui@ict.ac.cn | 🌐 zhangdenghui.tech

## Educations

**Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China** Sep. 2015 – Jun. 2018

- M.E. in Computer Science and Technology, at CAS Key Laboratory of Network Data Science and Technology
- Advisor: Prof. Jun Xu, and Prof. Yuanzhuo Wang
- Chinese Academy of Sciences Excellent Student Awards (**Top 5%**), Chinese Academy of Sciences Scholarships (Top 10%)

**University of Science and Technology Beijing, Beijing, China**

Sep. 2011 - Jun. 2015

- B.E. in the Communication Engineering, at School of Computer and Communication Engineering
- Major GPA: 3.83/4.0 (90.9/100), Overall GPA: 3.57/4.0 (86.7/100), WES GPA: 3.61/4.0
- National Scholarship (**Top 2%**), National Motivational Scholarship (Top5%), USTB Excellent Student Awards (Top 5%)

## Publications

AAAI'18 **Path-Based Attention Neural Model for Fine-Grained Entity Typing**

Denghui Zhang, Manling Li, Pengshan Cai, Yantao Jia, Yuanzhuo Wang.

The Thirty-Second AAAI Conference on Artificial Intelligence, 2018. (poster, accepted)

WI'17 **Efficient Parallel Translating Embedding For Knowledge Graphs.**

Denghui Zhang, Manling Li, Yantao Jia, Yuanzhuo Wang, Xueqi Cheng.

The IEEE/WIC/ACM International Conference on Web Intelligence, 2017. (regular paper, accepted)

IEEE Trans. **Link Prediction in Knowledge Graphs: A Hierarchy-Constrained Approach.**

Manling Li, Denghui Zhang, Yantao Jia, Yuanzhuo Wang, Xueqi Cheng.

IEEE Transaction on Big Data Special Issue on Knowledge Graphs: Techniques and Applications, 2017.  
(regular paper, under 2nd review)

## RESEARCH INTEREST

Natural Language Processing, Knowledge Graph and Machine Learning

## Research Projects

**Path-Based Attention Neural Model for Fine-Grained Entity Typing**

Jun. 2017 - Sep. 2017

NATURAL SCIENCE FOUNDATION OF CHINA (NSFC) NO.61572469

PI: Yuanzhuo Wang

- Problem: Fine-grained entity typing employs distant supervision to automatically generate training data. It labels entities with types in knowledge bases without considering the certain context of entities, and thus introduces noises.
- Came up with the idea to use attention mechanism to dynamically reduce the weights of wrong labeled sentences for each type during training.
- Proposed an end-to-end typing model, called the path-based attention neural model (PAN), by leveraging the distinct hierarchical structure of types to learn a noise robust performance. In this way, the attention learned for a type could assist the learning of the attention for its subtype.
- Experimented PAN by PyTorch, which can improve the fine-grained typing accuracy for 4 percent.
- The paper was accepted as a poster paper to **AAAI 2018** (Lead Author).

**Efficient Parallel Framework for Knowledge Graph Embedding**

Jan. 2017 - Apr. 2017

NATURAL SCIENCE FOUNDATION OF CHINA (NSFC) NO. 61402442

PI: Yantao Jia

- Problem: Knowledge graph embedding aims to embed entities and relations into low-dimensional spaces. However, a major limitation of these methods is the time consuming training process, which has difficulty in supporting large-scale knowledge graphs.
- Proposed a lock free framework for training knowledge graph embedding in parallel, called ParTrans-X.
- Proved the validity of ParTrans-X by formulating the training data of knowledge graph into hypergraphs.
- ParTrans-X can speed up the training process by more than an order of magnitude, with scaling performance along with increasing number of processors.
- The paper was accepted as a regular paper in **WI 2017** (Lead Author).

## Link Prediction Using Hierarchical Information in Knowledge Graphs

NATIONAL GRAND FUNDAMENTAL RESEARCH 973 PROGRAM OF CHINA No.2014CB340401

Dec. 2016 - Apr. 2017

Co-PI: Yuanzhuo Wang

- Problem: 50% triples in knowledge graphs are organized in hierarchical structures, which also contain rich inference patterns to predict links but do not be fully used.
- Divided the hierarchical structures into two categories, i.e., single-step and multi-step hierarchical structures.
- Participated in designing a link prediction method based on knowledge graph embedding, called **hTransM**. It can separate negative and positive examples by optimal margin, which is based on single-step and multi-step specific margin.
- Conducted the experiment on three datasets to demonstrate the effectiveness of hTransM.
- The paper was submitted to **IEEE Transaction on Big Data** (Second Author) and is under 2nd review now.

## Big Data Analysis Platform

KEY PROGRAM OF CAS KEY LABORATORY OF NETWORK DATA SCIENCE AND TECHNOLOGY

Sep. 2016 - Dec. 2016

PI: Jun Xu

- Objective: Providing a GUI web tool and Machine Learning library for data scientists to make workflows of various data mining applications fast and easily. The demo paper was accepted by **CIKM 2016**.
- Implemented GBDT (Gradient Boosting Decision Tree) algorithm on Spark.
- Optimized the GBDT algorithm on Spark. It ran faster than the official version of Spark MLlib, e.g. 2.4 times faster given 100GB data.
- Optimized the IO cost and storage cost of the workflow using Parquet file format.

## Projects

---

### Knowledge Graph Construction and Analysis for Videos

COLLABORATIVE PROJECT OF CHINESE ACADEMY OF SCIENCES AND HUAWEI INC.

Jan. 2017 - Apr. 2017

Advisor: Yantao Jia

- Objective: Constructing a knowledge graph for videos, including movies, series, etc. The knowledge graph can be enriched automatically from different data source, and can support a plenty of applications.
- Implemented the parallel version of key modules for knowledge graph construction, including ontology alignment algorithm, relation extraction algorithm and tag inference algorithm, which are based on knowledge graph embedding.
- Shortened the whole updating procedure to 5-6 hours from 3 days given 60k video entities, which was regarded as one of the main innovations by Huawei Inc.

### Commodity Demand Prediction

ALIBABA COMMODITY SUPPLY CHAIN PREDICTION COMPETITION

Apr. 2016 - Jun. 2016

Individual Work

- Objective: Predicting the future sales of certain products in TaoBao, using historical sales data and user behavior data.
- Designed and implemented the algorithm individually, which **ranked 13 out of 2807 teams**.
- Implemented GBDT algorithm using Java in MapReduce framework to make it run on the appointed platform ODPS.
- Designed the loss function to assign different loss to lower and higher prediction, which lead to a great improvement.

### Connect Six: A Computer Game

NATIONAL COMPUTER GAME TOURNAMENT

Jul. 2014 - Sep. 2014

Advisor: Ke Zhou

- Objective: Developing a computer program to play the board game Connect Six.
- Lead a team of three people and our team won the **national first prize**.
- Implemented Alpha-Beta Search algorithm to search the game tree of Connect Six.
- Optimized the search algorithm by adding VCF(Victory of Continuous Four) strategy.

## Honors

---

2016.10 CCF Agricultural Product Price Prediction Competition (**Rank: 2/547**)

2016.06 Alibaba Commodity Supply Chain Prediction Competition (**Rank: 13/2807**)

2014.08 National Computer Game Tournament (Connect Six) (**First prize**)

2013.10 National College Competition on Internet of Things (**Second prize**)

Scholarship National Scholarship (**Top 2%**, 2015), National Motivational Scholarship (**Top 5%**, 2014)

Chinese Academy of Sciences Scholarships (**Top 10%**, 2015)

Awards USTB Excellent Student Awards (**Top 5%**, 2012, 2013), USTB Excellent Student Leader Awards (**Top 10%**, 2013), Chinese Academy of Sciences Excellent Student Awards (**Top 5%**, 2016)

## Skills

---

Language Python, C/C++, Java, Scala, MySQL, Shell

Tools Sklearn, Spark MLlib, PyTorch, Tensorflow

English Tofel: 102, GRE: Verbal 153, Quantities 167, AW 3.5