

# ZHAOCHONG AN

Harbin Institute of Technology, Harbin, 150001, China, Tel: +86 18800415368

anzhaochong@outlook.com ◊ Personal Website: <https://1170801121.github.io> ◊ Github: 1170801121

## EDUCATION

---

### Harbin Institute of Technology

Aug.2017 – Jul.2021(*expected*)

*B.E. in Computer Science and Technology, specialized in Big Data Science* **GPA:** 92.62/100; **Ranking:** 13/266

#### Selected Courses:

Algebra and Geometry II (**100**); Calculus II (95); mathematical logic (**98.2**); Probability theory and Mathematical Statistics B (**99.5**); Modern algebra (**100**); Computing method (**97**); Digital logic and digital system design (96); Big Data Computing Foundation (95)

## PUBLICATIONS

---

**Zhaochong An**, Bozhou Chen, Houde Quan and Hongzhi Wang. EM-RBR: a reinforced framework for knowledge graph completion from reasoning perspective. NeurIPS 2020 (awaiting decision).

## RESEARCH EXPERIENCE

---

### EM-RBR: a Reinforced Framework for Knowledge Graph Completion from Reasoning Perspective

*Research Assistant (supervised: Hongzhi Wang) - Mass Data Research Center, HIT* Feb.2020 – May.2020

- Proposed EM-RBR (embedding and rule-based reasoning) framework, which can make knowledge graph embedding method capable of being used in the real and large-scale knowledge inference tasks.
- Conducted multi-relation path prediction using relational background knowledge contained in rules to explore relation between two entities in deeper context and achieve higher accuracy.
- Developed a scalable matrix approach that puts the structured information and embedding models into the same semantic space and designed a novel rating mechanism combined with reasoning process.
- EM-RBR is flexible and general enough to optimize any type of embedding models. In experiments, EM-RBR achieves better performance compared with previous models on FB15k, WN18 and new dataset FB15k-R.

### Profit Optimization for Taxi Driver and Passengers Using Large Dataset

Sept.2019

- Established decision-making model to help taxi driver select most profitable strategy by developing algorithms that take into account the costs of decisions and using Python crawler to collect relevant information to analyze the dependence of relevant inputs.
- Used probabilistic representations and analysis to address the uncertainty (particularly, those arising from human decision-makers), such as the passengers' boarding time, inherent in model based on queuing theory to make reasonable arrangements of driver and passengers for improving efficiency of taking rides.
- Introduced a priority factor to make above model more flexible and balanced. The regression curve from data performed well and showed the rationality of the model.

### One-stop Intelligent Class Assistant App

Nov.2019 – Apr.2020

- Designed smart terminals to conduct in-depth extraction and organization of the classroom's multiple sources of data, including key or difficult marks from student clients, course recordings, courseware, blackboards and photos.
- Performed speech transcription using a recognition framework based on DFCNN (Deep Fully Convolutional Neural Network) for recordings, applied TextRank and synonym extraction to extract keywords from texts, and proposed a n-gram-grammar based model to establish the relationship between keywords.
- Used backend IoT cloud platform to record marks from student clients, establish the mapping of structured knowledge elements to the original voice signal elements (audio sentences, etc.) and realize the schedule extraction models by syntactic parsing and mining the syntax rules.
- Developed the PowerPoint plug-in using C# and built-in interface of VSTO, which enabled obtain the timestamp of courseware, synchronize multiple client time and upload files to the remote server using *HttpWebRequest*.

### Construction Scheme of medical question answering system

*Apr.2019 – May.2019*

- Utilized Bert (Bidirectional Encoder Representations from Transformers) to encode problem and document as fixed dimension vectors.
- Employed Annoy (Approximate Nearest Neighbors Oh Yeah) to conduct Approximate nearest neighbor search.
- Used the question and document pairs in DuReader dataset and extracted the problem-document pairs to store them in a serialized file object.

### Effective unified index structure for multi-query on massive graph data (under distributed system environment)

*Nov.2019 - Dec.2019*

- Proposed construction algorithm of index structure and corresponding method based the index for queries (including shortest path query, reachability query and graph pattern matching) using Spark by Scala.
- Innovated the traditional 2-hop point selection scheme to use BFS search based on pruning to select the hop cover in index building algorithm.
- Decomposed the pattern graph to be matched into a single edge form, and joined the matching results of each edge, so as to obtain the graph matching results quickly and conveniently with the help of index.

### Intelligent Multi-functional Badminton Robot

*Oct.2017 – May.2018*

- Designed a multi-functional badminton court robot of serving and picking the ball by using a pneumatic control system that uses compressed gas as the working medium, and fluid transmission to transmit pressure of the gas.
- Developed sweeping services for robots with obstacle avoidance.
- Added fire monitoring service: through the temperature and humidity sensor, infrared sensor to detect the gas temperature value on the spot, turn on the early warning mode and send signals if reaching a certain standard.

## HONORS & AWARDS

---

**First-class People Scholarship**

*Oct.2017*

**Second-class People Scholarship**

*Oct.2018 & Oct.2019*

**Second-class Annual Innovation Projects Competition**

*Jun.2018*

**University-level Excellent Member**

*May.2017*

## EXTRA-CURRICULUM ACTIVITIES

---

**One Star Volunteer** , Heilongjiang Volunteer Association

Organized and participated in **18** different volunteering activities.

Total volunteer servicing time is up to **154** hours.

## SKILLS

---

**Programming Language:** Java, Python, C++, C, Scala, Matlab, Javascript , HTML, C#, Mysql, Cypher

**Operation System:** Linux (Ubuntu), Windows OS

**Machine Learning:** Python Numpy, Pandas, Sklearn

**Distributed System:** Hadoop, Spark, Pregel