



Wentao Huang

✉ wentaohuang@ruc.edu.cn

🌐 <https://github.com/fukien>

📍 Room 500, Information Building, Renmin University of China,
No. 59 Zhongguancun Street, Beijing, P.R.China, 100872

Education

Sep 2017 – Present

Renmin University of China

- M.Sc in Computer Science
- Supervisor: Ju Fan
- GPA: 3.74/4.00

Sep 2013 – Jun 2017

Shandong University

- B.Eng in Electronic Engineering
- GPA: 90.22/100.00
- Rank: 3/39

Internship

Jul 2018 – Feb 2019

Advanced Digital Sciences Center, Singapore

- Mentor: Yuchen Li
- Team: LS Heterogeneous Graphs project team
- Title: Research Engineer

Research Experience

Dec 2019 – Present

Rumor Detection

- The explosive growth of rumors has become a major issue concerning the public and increases the demand for research on rumor detection. In this work, we utilize machine learning techniques and propose a novel detection framework to deal with the rumor detection problem on the online news platform of ByteDance.

Sep 2019 – Present

Entity Matching for Scattered Tables

- Entity matching finds data instances that refer to the same real-world entity. In this work, we aim to find the matching entities from scattered tables. We utilize LSH methods to reduce the number of unnecessary comparisons and leverage NLP methods to directly extract features to find matching data instances.

Nov 2018 – Aug 2019

Bipartite Attributed Network Embedding

- Network embedding transforms complex network data into a low-dimensional vector space and shown great performance in many real-world scenarios. We proposed a novel model to deal with bipartite attributed network embedding problems, which incorporates both attributed and structural information to enhance the learning processes. Besides, we devised an efficient sampling strategy to make it more scalable on large-scale networks. A paper about this work was written and was accepted to SIGIR-2020.

Publication

- **Wentao Huang**, Yuchen Li, Yuan Fang, Ju Fan, Hongxia Yang. "BiANE: Bipartite Attributed Network Embedding", 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval (**SIGIR**), 2020.

Project Experience

Jul 2018 – Sep 2018

Apache SINGA

- Apache SINGA is an Apache Incubating project for developing an open-source machine learning library. I took part in the project of the implementation of GANs.
- Source code: <https://github.com/apache/incubator-singa>

Course Projects

Sep 2017 – Dec 2017

CarrotSQL

- CarrotSQL is an implementation of the relational database on Linux. It supports storage management, index, query parsing, and query optimization.
- Source code: <https://github.com/fukien/carrot>

Sep 2017 – Dec 2017

Pard

- Pard is a distributed database management system based on PostgreSQL and MySQL with P2P architecture. It supports data fragmentation, query decomposition, query localization, and query optimization on different sites.
- Source code: <https://github.com/dbiir/pard>

Awards & Prizes

- ACM SIGMOD 2020 Programming Contest Top30.
- The Postgraduate Scholarship of Renmin University of China.
- The Undergraduate Scholarship of Shandong University.
- Third Prize in Mathematics Competitions of Chinese College Student, China.
- First Prize in iCAN, Shandong Province.
- First Prize in Innovation and Entrepreneurship Competition, Shandong Province.
- Second Prize in "Challenge Cup", Shandong Province.
- Second Prize in National Undergraduate Electronic Design, Shandong Province.
- Gold Award in "Challenge Cup", Shandong University.
- First Prize in National Innovation Design Competition of Intelligent Mechatronical Systems for College Students, Shandong University.
- First Prize in National University Student Social Practice And Science Contest On Energy Saving & Emission Reduction, Shandong University.
- Champion of Football Cup, Shandong University.

Activities

- Moderator of Technical Q&A, TensorFlow China Community.
- Teaching Assistant for Linear Algebra, Renmin University of China.
- Minister of English Broadcasting Station, Shandong University.
- SSI certificated Open Water Diver.