XIN HUANG

+1(213)421-2739 ⋄ huan233@usc.edu ⋄ hxin18.github.io

EDUCATION

University of Southern California (USC), Los Angeles, USA — Jan. 2018 - May. 2020*(excepted)*
Master of Science in Computer Science — GPA: 4.0/4.0

Telecom Paristech(ENST), Paris, France — Sept. 2016 - Feb. 2017
Exchange Student, Majoring in Networking and Data Science — GPA: 3.8/4.0

Shanghai Jiao Tong University (SJTU), Shanghai, China — Sept. 2012 - Aug. 2016
Bachelor of Science in Information Engineering — GPA: 3.6/4.0

SKILLS

Programming Languages	C/C++, Python, HTML/CSS/JavaScript, R, SQL, Java, Matlab, Solidity
Platforms & FrameWork	Hadoop, Spark, Linux, Tensorflow, Spring, Git
Languages	English (TOEFL 100), French(TCF B2), Chinese(Mandarin&Cantonese)

EXPERIENCE

Research Programmer at Information Science Institute, USC — Jun. 2018 - Present
Entity Resolution and Knowledge Graph Construction — *Advisor: Mayank Kejriwal*

- Designed and built up an entity resolution pipeline for LORELEI from scratch, including preprocessing text, extracting string features, implementing blocking algorithm for scalability and defining matching rules for instances. It obtained the performance of 0.96 in precision and 0.9 in F-measure.
- Designed, implemented and modularized entity resolution pipeline for GAIA, including matching extracted instances from picture and text and designing high performance entity resolution framework based on min hashing. It executed four times faster than the baseline.
- Proposed, built up and visualized the human-terrain knowledge graph for capturing and interpreting events in Nepal earthquake twitter dataset.

Research Assistant at Intelligent IoT Center, Shanghai Jiao Tong University Jun. 2014 - Nov. 2017
Information Retrieval and Topic-based Recommendation for Academic Data — *Advisor: Xinbing Wang*

- Proposed a scalable generative model (able to process datasets containing over millions of papers) that is compatible to all topic modeling models and effectively generated the importance of papers in topic spaces.
- Achieved an improvement over 50% in recall, precision and mean reciprocal rank compared with other topic-based citation prediction methods.
- Implemented a topic sensitive academic recommender system based on the learnt parameters of the model.
- First-authored paper accepted by PAKDD 2018 (**Tier A Conferences in Data Mining**).

Research Assistant at DBGroup, Telecom ParisTech — Sept. 2016 - Feb. 2017
Large Graph Indexing and Compression — *Advisor: Mauro Sozio*

- Designed a novel vertices ordering algorithm(implemented in C++) for BV-Graph based on local sensitive hashing.
- Achieved around 3% improvement on compressing social network than state-of-art techniques.

SELECTED PROJECTS

Distributed Lookup Services — C++,Socket Programming

- Implemented a distributed lookup system from scratch. It contains three backend databases, an intermediate server, a client and a monitor and established stream sockets to transmit queries and results among all devices.
- Implemented the methods of exact matching, prefix matching and suffix matching for backend server.

House Searching System*(Light version of Airbnb)* — Java, MySQL, Spring, JQuery, Elasticsearch

- Implemented a house searching system with Spring framework. It is consisted of a backend management system, a house retrieval system combining search engine and baidu map and a leasing management system.
- Implemented the search engine with ElasticSearch and Kafka, the backend secure data transmission with MySQL, Spring Security and Spring Data JPA, the front-end with thymeleaf, bootstrap and jquery.

PUBLICATION

[1] **X.Huang**, C.Chen, C.Peng, X.Wu, L.Fu, X.Wang *Topic-sensitive Influential Paper Discovery in Citation Network.* Accepted by PAKDD 2018 *(Acceptance Rate: 20.6%)*