Qitong Wang

✉ qitongwang@hotmail.com · ☎ (+86) 185-0212-7069

Education

Fudan University, School of Computer Science   Sep 2016 – Jun 2019
Master in Data Science *Recommended without examinations* GPA 3.53/4.0

Fudan University, School of Computer Science   Sep 2012 – Jun 2016
Bachelor in Computer Science and Technology *GPA* 3.30/4.0

Stockholm University, Department of Computer and Systems Sciences   Sep 2014 – Feb 2015
Exchange Student with Full Scholarship GPA 3.40/4.0

Research Experience

Interactive Chinese Phrase Mining
Database and Massive Information Processing Laboratory at Fudan Univ.   Sep 2017 – Jun 2018

- Formalized problem, surveyed and experimented on related techniques, i.e., Chinese Word Segmentation and Phrase Mining
- Proposed an interactive Chinese phrase mining algorithm
- Implemented using Python/PyTorch, *SkipGram* for embedding, *LSTM* for labeling, and *active learning* for user interaction

Mining of Moving Companion Patterns
Database and Massive Information Processing Laboratory at Fudan Univ.   Sep 2016 – Aug 2017

- Characterized real-life data, formalized problem and surveyed related techniques
- Designed and implemented a distributed digest-based pruning algorithm, which pruned 99.5% candidates
- Optimized and refactored the proposed algorithm in Java using a two-level pruning structure with hash and digests, resulting 90% promotion of running time
- Implemented using *Java*, *Scala*, *Apache Spark*, *Locality Sensitive Hashing (LSH)*, *frequent pattern mining techniques*

Development of Columnar File Format for Time-Series Data (TsFile)
National Engineering Laboratory for Big Data Software at Tsinghua Univ.   May 2016 – Aug 2016

- Designed and developed the writing module of TsFile, a time series management system
- Studied literature of lossy compression encodings of data series, implemented and integrated into TsFile Run-Length Encoding (RLE), Delta Compression, Piecewise Linear Approximation (PLA), Swinging Door Algorithm (SDA), and Fast Fourier transform (FFT)
- Analyzed codebase of Apache Impala, based on which implemented the distributed writing module of TsFile
- Implemented using *C++*, *Java*, *Apache Thrift*, *Apache Impala*

Mining of Spatial Data
Database and Massive Information Processing Laboratory at Fudan Univ.   Nov 2015 – Apr 2016

- Designed and implemented a distributed map matching algorithm using grid index and hidden markov model
- Implemented using *Java*, *Apache Hadoop*, *Hidden Markov Model (HMM)*

Internships

Research Assistant, The University of Hong Kong
Advisor Prof. Reynold C.K. Cheng   Dec 2018 – Feb 2019

- Facilicate multimedia database queries with embedding techniques
- Implemented using *C++*, *Java*, *Python*, *Osmosis library of OpenStreetMap*

Research Assistant, Tsinghua University

Advisor Prof. Jianmin Wang and Chen Wang   May 2016 – Aug 2016

- Development of Columnar File Format for Time-Series Data (TsFile), deployed in Apache IoTDB
- Implemented using *C++*, *Java*, *Apache Thrift*, *Apache Impala*

PROFESSIONAL EXPERIENCE

Teaching Assistant, Fudan University

Probability and Statistics	Sep 2017 – Jan 2018
Discrete Mathematics	Sep 2017 – Jan 2018
Introduction to Databases	Mar 2018 – Jul 2018
New Query Approach in Big Data Era	Jul 2018

- Corrected homeworks, delivered recitations, arranged/designed/graded examinations

PUBLICATIONS

- **Qitong Wang**, Peng Wang, Yuliang Zhao and Wei Wang. An Efficient Approach to Mining Moving Companion Patterns. The 35th Chinese National Database Conference (NDBC), Dalian, China, 2018
- Yuliang Zhao, **Qitong Wang**, Peng Wang, Wei Wang, and Zhiguo Yan. A Distributed Approach of Accompany Vehicle Discovery. The 6th International Conference on Applications and Techniques in Cyber Security and Intelligence (ATCI), Shanghai, China, 2018
- Zhongsheng Li, Qiuhong Li, Wei Wang, **Qitong Wang**, Fengbin Qi, Yimin Liu, and Peng Wang. HDUMP: A Data Recovery Tool for Hadoop. The 23th International Conference on Database Systems for Advanced Applications (DASFAA), Gold Coast, Australia, 2018

AWARDS

- 2019 Outstanding Graduate of Shanghai City (*Top* 5%)
- 2017 Tung OOCL Award at Fudan University (*Top* 2%)
- 2016/2013 EMC Award at Fudan University (*Top* 5%)
- 2016 Outstanding Graduate of Fudan University (*Top* 10%)
- 2019/2018/2016/2014/2013 Outstanding Student Scholarship of Fudan University (*Top* 30%)

SKILLS

- Programming Skills: Java, Scala, Python, C/C++, SQL (PostgreSQL and MySQL)
- Open Source Library & Frameworks: PyTorch, Weka, Apache Hadoop, Apache Spark, Apache Impala, Apache Thrift
- Language Skills: English (CET-4/6, TOEFL-102, GRE-326), Chinese (Native)