Qing Chen

(+86) 150 2134 7994
qchen.cs3@gmail.com
qingchen3.github.io

Education

2013–2016 **M.Eng Computer Tech**, *Fudan University*, China.
GPA: 3.27/4.0

2009–2013 **B.Eng Computer Science**, *Zhengzhou University*, China.
GPA: 3.5/4.0 (Rank: 3/200)

Work Experience

2018/05 – now **Data Engineer**, *PayPal*, Shanghai, China.

2015/07 – 2017/06 **Research Assistant**, *Qatar Computing Research Institute (QCRI)*, Doha, Qatar.

2014/09 – 2015/06 **Software Engineer Intern**, *DELL EMC*, Shanghai, China.

Conference Publications

[1] Nan Tang, **Qing Chen**, Prasenjit Mitra. *Graph Stream Summarization: From Big Bang to Big Crunch.* International Conference on Management of Data (SIGMOD)(acceptance rate 19%), San Francisco, USA, 2016.

[2] **Qing Chen**, Zijing Tan, Chu He, Chaofeng Sha, Wei Wang. *Repairing Functional Dependency Violations in Distributed Data.* International Conference on Database Systems for Advanced Applications(DASFAA) (acceptance rate 22%), Hanoi, Vietnam, 2015.

[3] Chu He, Zijing Tan, **Qing Chen**, Chaofeng Sha, Zhihui Wang, Wei Wang. *Repair Diversification for Functional Dependency Violations.* International Conference on Database Systems for Advanced Applications(DASFAA) (acceptance rate 24%), Bali, Indonesia, 2014 (**Best Paper Runner-up)**.

Journal Publications

[4] Chu He, Zijing Tan, **Qing Chen**, Chaofeng Sha. *Repair diversification: A new approach for data repairing.* Information Sciences(Inf.Sci.). Vol.346-347 pp. 90-105, 2016.

Projects

Research Projects

2015/07–2016/04 **Graph Stream Summarization**, QCRI, Qatar.
data streams
Supervisor: Dr. Nan Tang and Prof. Prasenjit Mitra
Achievements: With co-workers, proposed TCM (SIGMOD 2016) which is the state-of-art graphical sketch for graph streams[1]. As TCM keeps internal connections of graph streams, TCM supports richer graph queries than existing sketches.

- Participated in designing TCM sketch and algorithms.
- Evaluated the performance, efficiency and effectiveness of TCM.
- Carried out experimental studies and implemented algorithms in C++[1].

2013–2016 **Data quality projects**, Fudan University.
data quality Supervisor: Prof. Zijing Tan and Prof. Chaofeng Sha
Achievements: With co-workers, proposed an efficient framework to repair violations of functional dependencies in horizontally partitioned relational database (DASFAA 2015)[2]. Proposed a novel repair diversification methodology (DASFAA 2014)[3] (Inf.Sci.)[4]
- Designed efficient message-passing distributed computing model[1].
- Implemented message-passing framework and conducted experimental studies in Java[1]
- Participated in formulating data repairing diversification problem and carrying out experimental studies in[3][4].

Industrial Projects

2018/05–now **Machine Learning System & Data System**, *PayPal*, Shanghai, China.
As a full stack data engineer, I'm working with a wide array of technologies ranging from big data technologies such as Spark, Hadoop to machine learning/deep learning paradigm and python web frameworks.
Achievements: Developed disruptive products, experimented with emerging technologies, and promoted lean innovation.
- Designed, implemented and evaluated a text summary tool for PayPal. I'm developing and evaluating machine learning models to improve the quality of auto text summary.
- Designed, implemented and evaluated a machine learning system to detect anomalous events for PayPal. Developed and evaluated topic models and sequence models for anomaly detection. This system has reported several influential events such as large scale phishing and protests against violence.
- Designed, implemented and evaluated a data privacy scanning system within Teradata. The system is scanning thousands of tables daily for PayPal and reporting privacy data for PayPal users.

2016/06–2017/06 **Malicious Domain Detection System through graph analysis**, QCRI, Qatar.
Achievements: With co-workers, developed an efficient and large-scale system to report malicious domains, experimented with emerging technologies to support research team.
- Designed and evaluated graph algorithms for a large-scale data.
- Designed, implemented and evaluated the detecting system.
- Design and developed a demo system for the research team.

2014/09–2015/06 **Developing tools for Hyper-converged infrastructure product**, DELL EMC, Shanghai, China.
Achievements: Developed system tools that automatically installed Linux System for company Hyper-converged infrastructure Product **EMC VSPEX-BLUE**.
- Investigated the VSphere platform and Linux system.
- Designed, implemented tools that read Linux system files and installed into Hyper-converged infrastructure through VSphere.

Computer Skills

Languages Java, Python, C++, Javascript
Framework Tensorflow, Flask, Pyramid
RDBS MySQL, PostgreSQL, Oracle
Big data Spark, Hadoop, Hive, Giraph
System Linux, Teradata, VSphere

Scholarships and Awards

05/2014 **Best Paper Runner-up**, *International Conference on Database Systems for Advanced Applications(DASFAA)*, Bali, Indonesia.
The award includes an honorarium and is based on both the paper and the presentation at the conference.

2015 **Huawei Scholarship**, *Fudan University*, Shanghai, China.
8,000 RMB is granted to students who have done outstanding research.

2014 **Tung OOCL Scholarship**, *Fudan University*, Shanghai, China.
6,000 RMB is granted to students who have done outstanding research.

2011, 2010 **First Class Scholarship**, *Zhengzhou University*, Henan, China.
3,000 RMB is granted to students who are among top 3 percents in academic performance.

2012 **Second Class Scholarship**, *Zhengzhou University*, Henan, China.
2,000 RMB is granted to students who are among top 5 percents in academic performance.

Teaching Activities

2013–2014 **Teaching Assistant**, Fudan University.
2013 Fall, Introduction to Computer Network
2014 Spring, Introduction to Database
2014 Spring, Introduction to Programming Lauguage
Assisted professors in the teaching classes every week, taught students at their coding classes every week, and graded students' assignments.