Qing Chen

qchen.cs3@gmail.com
(86) 15021347994

EDUCATION

Fudan University Shanghai, China 09/2013 - 01/2016
Master of Engineering, *Computer Science* GPA: 3.27/4.0 Supervisor: Prof. Zijing Tan
Thesis: Repairing Functional Dependency Violations in Horizontally Partitioned Data (in Chinese)

ZhengZhou University Zhengzhou, China 09/2009 - 07/2013
Bachelor of Engineering, *Computer Science* GPA: 3.5/4.0 Ranking: 3/200

RESEARCH EXPERIENCE

Research Assistant, Qatar Computing Research Institute Doha, Qatar 07/2015 - 06/2017
Graph Stream Summarization research project Superviosr: Dr. Nan Tang 07/2015 - 04/2016

- Participated in proposing and designing the graphical sketch TCM for graph streams[1].
- Carried out all the algorithms and experimental studies in C/C++ [1]

Malicious Domain Detections through Graph Analysis 05/2016 - 06/2017
In this project, I focused on coding and developing the detecting system. The system extracted domains and IPs from DNS logs to produce generated a bipartite graph. In the generated bipartite graph, there will be an undirected edge between one domain and one IP when the IP hosts that domain. We can construct one graph that called "domains to domains" wherein nodes are domains and two domains are connected by an undirected edge if they ever shared the same IP address. Given malicious domains in domains to domains graph, our goal is to infer maliciousness of unknown domains through their associations to known malicious domains.

- Participated in system design and graph algorithms tunning.
- ETL pipeline development in Python, Shell, Hadoop, PostgreSQL
- Participated in developed subsystem for the demonstration of the research project

Research Assistant, Fudan University Shanghai, China 09/2013 - 07/2015
Research projects on data repairing and data quality Superviosr: Zijing Tan

- Designed one message-passing distributed computing model to detect and repair violations of functional dependencies in horizontally partitioned data[2]
- Implemented message-passing framework and conducted experimental studies in Java[2]
- Participated in formulating data repairing diversification problem and carrying out experimental studies in [3]

TEACHING

Teaching Assistant, Fudan University Shanghai, China

- Introduction to the Internet 09/2013 – 01/2014
- Introduction to database management 02/2015 – 05/2014
- Visual Basic Programming 09/2014 – 01/2015

INDUSTRY EXPERIENCE

Data Engineer, PAYPAL Shanghi, China 05/2018 – now
Topic discovery and anomalious event detection in global customer emails of PayPal

- Developing machine learning pipelines across hetergeneous softwares: Spark, Hadoop, Teradata, Tensorflow
- Ultilizing topic models to discover 20+topics in massive customer emails from Customer Service Team
- Have discovered influential events such as protests against violent video game among PayPal Users

Protecting privacy of EU Users under EU General Data Protection Regulation (GDPR)
Identified all EU Users' private information, such as age, phone number, belief in Teradata. I have been responsible for this project.

- Designed the architecture of the privacy scanning system
- Optimizatized join computations among massive data
- The sysytem keeps scanning thousands of tables every day.

Software Engineer Intern, DELL EMC Shanghi, China 09/2014 – 06/2015
Developed system components that automatically installed OpenSuse Linux Virtual File into Vsphere Platform for the company's product "VSPEX BLUE Hyper-Converged Infrastructure Appliance".

- Understood Suse Linux, OpenSUSE open virtual files format and VSphere platform
- Developed system components in python, shell

PROGRAMMING AND SYSTEM SKILLS

- Proficient programming skills in: Python, Java, Shell, SQL
- Solid programming skills in: C\C++, Javascript, CSS, HTML
- Industrial experience in Tensorflow, Spark, Hadoop, Teradata, VSphere, Linux/Unix

PUBLICATIONS

[1]N. Tang, **Qing Chen**, Prasenjit Mitra. Graph Stream Summarization: From Big Bang to Big Crunch. SIGMOD 2016

[2]**Qing Chen**, Z.Tan, C. He, C. Sha, W W. Repairing Functional Dependency Violations in Distributed Data. DASFAA, 2015.

[3]Chu He, Zijing Tan, **Qing Chen**, Chaofeng Sha, Zhihui Wang, W Wang. Repair Diversification for Functional Dependency Violations. DASFAA 2014. (Best Paper Candidates)

AWARDS

- Huawei Scholarship in 2015
- Tung OOCL Scholarship in 2014
- Best Paper Runner at the International Conference on Database Systems for Advanced Applications 2014
- A Scholarship(top 3%) in 2010, 2011 and B Scholarship(top 5%) in 2012 of Zhengzhou University

SELF-LEARNING COURSES

Coursera: Neural Networks for Machine Learning. (Geoffrey Hinton)

Coursera: Machine Learning; Nerual Networks and Deep Learning; Imporving Deep Neural Networks: Hyperparamter tuning, Regularization and optimization; Structuring Machine Learning Projects; Convolutional Neural Networks. (Andrew Ng)