Peng Xu

Curriculum Vitae

Education

2011-2015(expected) **B.Eng. in Electronic Information Engineering**, *Beijing University of Posts and Telecommunications*, Beijing, *Major: 89.62/100, Overall: 87.06/100*.

Academic Interests

- Statistical Machine Learning
- Data Mining
- Parallel Computing

Research Experience

2014 - Present **High Performance Computing Laboratory**, *Tsinghua University*, Beijing, *Supervised by Zhihu Du, focus on parallel computing via GPU hardware and distributed systems. More specifically, efficient partical-mesh spreading on GPUs and handle large amount of images via MPI, Spark and Hadoop*.

2014 - Present **Knowledge Engineering Group**, *Tsinghua University*, Beijing, *Supervised by Jie Tang, focus on machine learning, data mining and information retrieve. More specifically, name disambiguation in large scale database and the incremental case of it via reinforcement learning and semi-supervised learning*.

Publications

2014 Xiangyu Guo, Xing Liu, **Peng Xu**, Zhihu Du, Edmond Chow, ***Efficient Particle-Mesh Spreading on GPUs (submitted)***,In 29th IEEE Internatioanal Parallel and Distributed Processing Symposium(IPDPS), 2015

Project Experience

2014 **Arnetminer II**, *python, scala*, Research Project

Arnetminer is a website that offers comprehensive search and mining services for academic community. I am working on the second version of it. My work mainly involves with the data, including cache, mongodb, redis and the corresponding computation. In addition, I am working on improve the performance of name disambiguation of the system and how to handle the new data properly.

+86 18810542627 • bly930725@gmail.com

2014 **Image Library and Process System**, *python, scala, C/C++*, Research Project
Handle large amount of images crawled online via parallel computing and distributed systems, more specifically, using MPI, Spark, Hadoop respectively and compare their performance in different situations.

2014 **Crawler for Google Crawler**, *python*, Research Project
Develop a web crawler to crawl the whole co-author network with all their papers on Google Scholar through bfs. In order to avoid blockage of Google, I also write several crawlers to obtain thousands of free proxies online.

2014 **Identify Experts in Baidu Baike**, *python*, Research Project
Develop a GUI to accept the input information of the experts entered or from the database. Use the given information to identify the experts in Baidu Baike, which is a website like Wikipedia in Chinese. And try various methods including SVM and Logisitic Regression to improve the accuracy and recall.

2014 **Efficient Particle-Mesh Spreading on GPUs**,*C, CUDA*,Research Project
Study various approaches for particle-mesh spreading on GPUs. A central concern is the use of atomic operations. We are also concerned with the case where spreading is performed multiple times using the same particle configuration, which opens the possiblity of preprocessing to accelerate the overall computation time. Experimental tests show which algorithm are best under different circumstances. Our work is submitted to IPDPS, 2015

2013 **"Social Networks and Algorithms Design" Curriculum Design**,*python, Matlab*, Course Project
Based on data provided by Alibaba, use logistic regression and SVM to predict the behaviors of the consumers.

Community

Homepage **billy-inn.github.io/Homepage/**
Blog **billyinn.wordpress.com** My blog(English)
GitHub **github.com/billy-inn**
Codeforces **codeforces.com/profile/billy_inn**,rating 1732

Awards

- 2014 Enter the quarter-final in **Microsoft Programming of Beauty Contest**, Rank 134 of 988
- 2013 Bronze Medalist in **ACM-ICPC Hunan Invitational Programming Contest**
- 2013 Gold Medalist in **Programming Contest of BUPT**

Skills and Qualifications

Languages C/C++, python, CUDA, Matlab, Shell, Scala
OS Microsoft Windows, Mac OS, Linux
Tools LaTeX, Git, LIBSVM, MongoDB, Scrapy, Redis, Spark
Coursera *Machine Learning, Probabilistic Graphical Models, Game Theory*

+86 18810542627 • bly930725@gmail.com