Chunqiu Zeng

Resume

11200 SW 8th. Street
Miami, 33199, FL
786-527-0301
grandzeng@gmail.com
Homepage: `http://users.cis.fiu.edu/~czeng001`,
GitHub: `https://github.com/grandzeng`

Interests

Large Scale Data Mining; Temporal Data Mining; Distributed Data Mining; Data Analysis

Experience

Industry Experience

Summer of 2015 — **Intern**, *DoubleClick Search Reporting Team*, **Google**, Kirkland , WA, USA.
Leveraged some data mining and statistics techniques to conduct change impact analysis, with the purpose to identify the correlation between advertising changes and performance metrics. Constructed a flume pipeline for change impact analysis over large scale data.

Summer of 2014 — **Intern**, *VideoAds Format Team*, **Google**, Mountain View , CA, USA.
Built a tool to integrate multi-sources Ads time series data on YouTube video. Developed an integrated interface to facilitate the users for Ads time series data querying, analysis and visualization. Explored the correlation between time series to identify the factors in Ads format which contributes to skipping rate of Ads.

Summer of 2013 — **Intern**, **IBM T.J Watson Research Center**, New York, NY, USA.
Facilitated problem diagnosis, determination and auto-resolution in a complex IT environment by designing and implementing a hierarchical multi-label classifier to analyze the incident tickets data of IBM Tivoli Monitoring System.

2009.7-2012.1 — **Senior Data Engineer**, *Data Warehouse Department*, **Alibaba Group**, Hangzhou, China.
- Designed and implemented the meta-data extraction and association rule mining algorithms for the *Meta-data Mining System*, a system that leverages data mining techniques to discover the latent relationship of 10000+ database tables from the heterogeneous sources such as DB2, Oracle, Greenplum, etc.
- Designed and implemented the scheduling algorithm for the internal data analysis job scheduling system. The implemented algorithm is able to schedule 10000+ ETL jobs per day on top of a heterogeneous architecture consisting of DB2, RAC, Greenplum, Hadoop/Hive, etc.
- Designed and implemented a distributed system called *Data Highway* to facilitate the high-speed data transmission between the heterogeneous data sources including DB2, HDFS, SQLServer, MySQL, Greenplum, RAC.
- Designed and implemented the *Automatic Deployment System* that is able to automatically deploy hundreds of projects per day.

2007.8-2008.3 — **Software Engineer Intern**, *MSGUI Group*, **Alcatel-Lucent**, Chengdu, China.
- Designed the conceptual, logical, and physical data model and implemented the data access interface for a network management system called MAPIM.
- Implemented the configuration module for a user graphic rendering service called GUI Server and the communication module between the GUI Server and the MAPIM.

Academic

2012.1-Present — **Research Assistant**, *Knowledge Discovery Research Group*, FIU, USA.
- Designed and implemented a Fast,Integrated,User-friendly system named FIU-Miner in support of running a number of data analysis tasks in distributed environment. FIU-Miner, a linearly extensible system, provides a user-friendly workflow to schedule data analysis tasks in full use of the available resources of distributed environment.[**online demo**]
- Leaded the design and implementation of manufacturing data analysis platform. This platform provides an integrated suit of solutions to analyze manufacturing data, especially the product data of Plasma Display Panel. The platform shows the entire solution of the data analysis from the initial data exploration to the final knowledge extraction.
[**online demo**]

Programming Skills

Languages:	Java (proficient), Python (familiar), Javascript (familiar), C/C++ (familiar)	Frameworks & Libraries:	Apache Hadoop, Flume, Thrift, Protocol Buffer, Stubby, Mahout, jQuery
DB:	Oracle, DB2, MySQL, MongoDB, PostgreSQL, HDFS, Hive, Big Table	Tools:	Git, SVN, make,

Selected Publications

1. Liang Tang, Yexi Jiang, Lei Li, **Chunqiu Zeng**, Tao Li."Personalized Recommendation via Parameter-Free Contextual Bandits", in Proceedings of the 38th Annual International ACM SIGIR Conference (SIGIR 2015), Santiago de Chile., Aug. 9-13, 2015.
2. **Chunqiu Zeng**, Liang Tang, Tao Li, Larisa Shwartz, Genady Ya. Graharnik."Mining Temporal Lag from Fluctuating Events for Correlation and Root Cause Analysis", The 10th International Conference on Network and Service Management(CNSM 2014).
3. **Chunqiu Zeng**, Hongtai Li, Huibo Wang, Yudong Guang, Chang Liu, Tao Li, Mingjin Zhang, Shu-Ching Chen, Naphtali Rishe."Optimizing Online Spatial Data Analysis with Sequential Query Patterns", The 15th IEEE international Conference on Inofrmation Integration and Reuse(IRI 2014).
4. Yexi Jiang, **Chunqiu Zeng**, Jian Xu, Tao Li. "Real time contextual collective anomaly detection over multiple data streams", ACM SIGKDD Workshop Outlier Detection & Description under Data Diversity (SIGKDD Workshop ODD$^\Lambda$2), 2014.
5. Li Zheng, **Chunqiu Zeng**, Lei Li, Yexi Jiang, Wei Xue, Jingxuan Li, Chao Shen, Wubai Zhou, Hongtai Li, Liang Tang, Tao Li, Bing Duan, Ming Lei, and Pengnian Wang."Applying Data Mining Techniques to Address Critical Process Optimization Needs in Advanced Manufacturing",in Proceedings of the 20th ACM Conference on Knowledge Discovery and Data Mining (KDD 2014), New York, USA, Aug. 2014.
6. **Chunqiu Zeng**, Tao Li, Larisa Shwartz, Genady Ya. Graharnik. "Hierarchical Multi-Label Classification over Ticket Data using Contextual Loss." in Proceedings of IEEE/IFIP Network Operations and Management Symposium (NOMS'2014), 2014.
7. **Chunqiu Zeng**, Yexi Jiang, Li Zheng, Jingxuan Li, Lei Li, Hongtai Li, Chao Shen, Wubai Zhou, Tao Li, Bing Duan, Ming Lei, Pengnian Wang. "FIU-Miner: A Fast, Integrated, and User-Friendly System for Data Mining in Distributed Environment." ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2013.
8. Li Zheng, Chao Shen, Liang Tang, **Chunqiu Zeng**, Tao Li, Steve Luis, Shu-Ching Chen. "Data Mining Meets the Needs of Disaster Information Management." In IEEE Transactions on Human-Machine Systems, 2013.
9. **Chunqiu Zeng**, Jie Zuo, Chuan Li, Kaikuo Xu, Shengqiao Ni, Liang Tang, Yue Zhang, Shaojie Qiao. "MPSQAR: Mining Quantitative Association Rules Preserving Semantics", in Proceedings of the International Conference on Advanced Data Mining and Applications (ADMA), pp. 572 - 580, 2008

Awards

12/2014	The Best Graduate Student Research Award in the School of Computing and Information Sciences (SCIS) at FIU for Year 2013-2014.
12/2006	The 2nd National Open Source Software Competition (China), Award finalists
06/2006	Amway College Student's computer works competition, Second class award
2004	Second Prize in Sichuan Contest District for National Undergraduate Mathematical Contest in Modeling

Education

2012-Now	**Ph.D. in Computer Science**, *Florida International University*, Miami, FL, USA. **GPA: 3.9**. Research Area: Large-scale Data Mining, Event Mining, Distributed Data Mining
2006-2009	**M.S. in Computer Science**, *Sichuan University*, Chengdu, Sichuan, China. GPA: 3.6. Research Area: Data Mining
2002-2006	**B.E. in Computer Science**, *Sichuan University*, Chengdu, Sichuan, China. GPA: 3.4