Junbo Zhang

Department of Computer Science & Engineering
The Chinese University of Hong Kong

jbzhang@cse.cuhk.edu.hk
http://www.lucktroy.org

Education

9/2009 - present PhD Candidate in School of Information Science and Technology
Southwest Jiaotong University, Chengdu, China
Supervisors: Prof. Tianrui Li & Prof. Yi Pan

9/2005 - 7/2009 Bachelor of Telecommunication Engineering
Southwest Jiaotong University, Chengdu, China

Experience

5/2013 - present Research Assistant at The Chinese University of Hong Kong, Hong Kong
Research and develop large-scale deep learning and feature engineering algorithms.

5/2013 - present Research Intern at Huawei Noah's Ark Lab, Hong Kong
Research and develop large-scale deep learning and feature engineering algorithms.
Applied above techniques to telecommunication data mining and computer vision problems.

- Proposed a novel deep learning approach for learning high-level representation, which is presented at the 20th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (**KDD 2014**).
- Proposed a novel algorithm for mining discriminative and essential frequent sequential patterns via model-based search tree. It is the core component of the prediction model, which is used for the consumer complaint problem in the telecommunications industry. The related patent application has been submitted.
- Proposed a novel algorithm for mining discriminative and essential frequent sequential patterns via model-based search tree. It is the core component of the prediction model, which is used for the consumer complaint problem in the telecommunications industry. The related patent application has been submitted.
- Proposed a novel locally linear deep learning for large-scale stellar spectrum recognition.

2/2012 - 2/2013 Visiting PhD Student at Georgia State University, USA
Research and develop parallel algorithms, large-scale algorithms in cloud computing and GPU cluster.

- Participated as a mentor in 8-week NSF REU Undergraduate Summer Research Program.
- Proposed large-scale feature selection algorithms, provide HADOOP and SPARK implementations.
- Proposed a composite relation for data fusion & a novel parallel matrix algorithm using multi-GPU.

8/2011 - 9/2011 Intern at Belgian Nuclear Research Centre (SCK-CEN), Belgium
Research and develop incremenal and parallel algorithm for feature selection and knowledge acquisition.

Research Topics

- Deep Learning, Representation Learning, Feature Engineering
- Cloud Computing, Distributed Computing, High Performance Computing
- Big Data Mining
- Rough Sets/Fuzzy Sets

Publications

Journal Papers

1. **Junbo Zhang**, Yun Zhu, Yi Pan, Tianrui Li.
Efficient Parallel Boolean Matrix Based Algorithms for Computing Composite Rough Set Approximations.
Information Sciences, 1st revision.

2. **Junbo Zhang**, Jian-Syuan Wong, Yi Pan, Tianrui Li.
A Parallel Matrix-based Method for Computing Approximations in Incomplete Information Systems.
Accepted for publication in IEEE Transactions on Knowledge and Data Engineering (**TKDE**).

3. Dun Liu, Tianrui Li, **Junbo Zhang**.
Incremental Updating Approximations in Probabilistic Rough Sets Under the Variation of Attributes
Knowledge-Based Systems, vol. 73, pp. 81–96, 2015.

4. Anping Zeng, Tianrui Li, Dun Liu, **Junbo Zhang**, Hongmei Chen.
A Fuzzy Rough Set Approach for Incremental Feature Selection on Hybrid Information Systems.
Fuzzy Sets and Systems, vol. 258, pp. 39–60, 2015.

5. Dun Liu, Tianrui Li, **Junbo Zhang**.
A Rough Set-based Incremental Approach for Learning Knowledge in Dynamic Incomplete Information Systems.
International Journal of Approximate Reasoning, vol 55, no. 8, pp. 1764–1786, 2014.

6. **Junbo Zhang**, Tianrui Li, Yi Pan, Chuan Luo, Fei Teng.
A Parallel and Incremental Algorithm for Updating Knowledge Based on Rough Sets in Cloud Computing Platform.
Accepted for publication in **Journal of Software**. (in Chinese)

7. Anping Zeng, Tianrui Li, **Junbo Zhang**, Hongmei Chen.
Incremental Maintenance of Rough Fuzzy Set Approximations under the Variation of Object Set.
Fundamenta Informaticae, vol. 132, no. 3, pp. 401-422, 2014.

8. **Junbo Zhang**, Jian-Syuan Wong, Tianrui Li, Yi Pan.
A Comparison of Parallel Large-scale Knowledge Acquisition Using Rough Set Theory on Different MapReduce Runtime Systems.
International Journal of Approximate Reasoning, vol. 55, no. 3, pp. 896-907, 2014.

9. **Junbo Zhang**, Tianrui Li, Hongmei Chen.
Composite Rough Sets for Dynamic Data Mining.
Information Sciences, vol. 257, pp. 81-100, 2014.

10. **Junbo Zhang**, Dong Xiang, Tianrui Li, Yi Pan.
M2M: A Simple Matlab-to-MapReduce Translator for Cloud Computing.
Tsinghua Science and Technology, vol 18, no. 1, pp. 1-9, 2013.

11. Yi Pan, **Junbo Zhang**.
Parallel Programming on Cloud Computing Platforms: Challenges and Solutions.
KITCS/FTRA Journal of Convergence, vol. 3, no. 4, pp. 23-28,2012.

12. **Junbo Zhang**, Tianrui Li, Da Ruan, Zizhe Gao, Chengbing Zhao.
A Parallel Method for Computing Rough Set Approximations.
Information Sciences, vol. 194, pp. 209-223, 2012.

13. **Junbo Zhang**, Tianrui Li, Da Ruan, Dun Liu.
Rough Sets Based Matrix Approaches with Dynamic Attribute Variation in Set-valued Information Systems.
International Journal of Approximate Reasoning, vol. 53, no. 4, pp. 620-635, 2012.

14. **Junbo Zhang**, Tianrui Li, Da Ruan, Dun Liu.
Neighborhood Rough Sets for Dynamic Data Mining.
International Journal of Intelligent Systems, vol. 27, no. 4, pp. 317-342, 2012.

Conference Papers

1. **Junbo Zhang**, Guangjian Tian, Yadong Mu, Wei Fan.
Supervised Deep Learning with Auxiliary Networks.
Proceedings of the 20th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (**KDD 2014**), New York, USA, 2014, pp. 353-361.
(AR: 151/1036 = 14.6%)

2. Chuan Luo, Tianrui Li, Hongmei Chen, **Junbo Zhang**.
Dominance-based Rough Sets in Composite Ordered Information Systems.
In: Proceedings of the Joint Conference of 13th China Conference on Rough Sets and Soft Computing, 7th China Conference on Web Intelligence, 7th China Conference on Granular Computing (**CRSSC-CWI-CGrC 2013**), Zhangzhou, China, 2013. (in Chinese) (**Outstanding Student Paper Award**)

3. **Junbo Zhang**, Tianrui Li, Yi Pan.
PLAR: Parallel Large-scale Attribute Reduction on Cloud Systems.
The 14th International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT), Taipei, 2013.

4. **Junbo Zhang**, Chizheng Wang, Yi Pan, Tianrui Li.
Parallel Approaches to Neighborhood Rough Sets: Classification and Feature Selection.
The 8th International Conference on Intelligent Systems and Knowledge Engineering (ISKE), Shenzhen, China, 2013.

5. **Junbo Zhang**, Yun Zhu, Yi Pan, Tianrui Li.
A Parallel Implementation of Computing Composite Rough Set Approximations on GPUs.
The Eighth International Conference on Rough Sets and Knowledge Technology (RSKT), Halifax, Canada, pp. 240-250, 2013.

6. **Junbo Zhang**, Jian-Syuan Wong, Tianrui Li, Yi Pan.
H2T: A Simple Hadoop-to-Twister Translator for Cloud Computing.
2013 International Symposium on Biometrics and Security Technologies (ISBAST), Chengdu, China, pp. 180-186, 2013.

7. Hongmei Chen, Tianrui Li, **Junbo Zhang**, Chuan Luo.
Probabilistic Composite Rough Set and Attribute Reduction.
The 7th International Conference on Intelligent Systems and Knowledge Engineering (ISKE), Beijing, China, 2012.

8. **Junbo Zhang**, Tianrui Li, Hongmei Chen.
Composite Rough Sets.
Proceedings of the 4th International Conference on Artificial Intelligence and Computational Intelligence (AICI), Chengdu, China, pp. 150-159, 2012.

9. **Junbo Zhang**, Tianrui Li, Yi Pan.
Parallel Rough Set Based Knowledge Acquisition Using MapReduce from Big Data.
ACM SIGKDD 2012 Workshop on Big Data Mining (BigMine), Beijing, China, pp. 20-27,2012.

10. Anping Zeng, Tianrui Li, **Junbo Zhang**, Dun Liu.
An Incremental Approach for Updating Approximations of Rough Fuzzy Sets under the Variation of the Object Sets.
Proceedings of the 8th International Conference on Rough Sets and Current Trends in Computing (RSCTC), Chengdu, China, pp. 36-45 2012.

11. **Junbo Zhang**, Tianrui Li, Yan Yang, Wei Wang.
Neighborhood Rough Sets Based Matrix Approach for Calculation of the Approximations.
Proceedings of the 6th International Conference on Rough Sets and Knowledge Technology (RSKT), Banff, Canada, pp. 166-171, 2011.

12. **Junbo Zhang**, Tianrui Li, Da Ruan.
Rough Sets Based Incremental Rule Acquisition in Set-valued Information Systems.
Autonomous Systems: Developments and Trends, Mallorca, Spain, vol. 391, pp. 135-146, 2012.

13. Zizhe Gao, Tianrui Li,**Junbo Zhang**, Chengbing Zhao, Zhonggang Wang.
A Parallel Method for Unpacking Original High Speed Rail Data Based on MapReduce.
Proceedings of the 6th International Conference on Intelligent Systems and Knowledge Engineering (ISKE), Shanghai, China, vol. 124, pp. 59-68, 2011.

14. **Junbo Zhang**,Tianrui Li, Lei Wang.
A New Method for Calculation of the Approximations Under the Probabilistic Rough Sets.
Proceedings of 2010 IEEE International Conference on Intelligent Systems and Knowledge Engineering (ISKE), Hangzhou, China, pp. 73-76, 2010.

15. **Junbo Zhang**, Tianrui Li, Dun Liu.
An Approach for Incremental Updating Approximations in Variable Precision Rough Sets While Attribute Generalized.
Proceedings of 2010 IEEE International Conference on Intelligent Systems and Knowledge Engineering (ISKE), Hangzhou, China, pp. 77-81, 2010.

16. **Junbo Zhang**, Tianrui Li, Muhammad Zia ur Rehman, Shaoyong Li.
An Approach for Selection of the Proper Lever of Granularity in Granular Computing. Proceedings of the 9th International FLINS Conference on Foundations and Applications of Computational Intelligence (FLINS), Chengdu, China, pp. 764-769, 2010.

17. Dun Liu, **Junbo Zhang**, Tianrui Li.
A Probabilistic Rough Set Approach for Incremental Learning Knowledge on the Change of Attributes.
Proceedings of the 9th International FLINS Conference on Foundations and Applications of Computational Intelligence (FLINS), Chengdu, China, pp. 722-727, 2010.

18. Hongmei Chen, Tianrui Li, **Junbo Zhang**.
A Method for Incremental Updating Approximations when Objects and Attributes Vary with Time.
Proceedings of 2010 IEEE International Conference on Granular Computing (GrC), California, USA, pp. 90-95, 2010.

19. Hongmei Chen, Tianrui Li,**Junbo Zhang**.
A Method for Incremental Updating Approximations based on Variable Precision Set-valued Ordered Information Systems.
Proceedings of 2010 IEEE International Conference on Granular Computing (GrC), California, USA, pp. 96-101, 2010.

20. Shaoyong Li, Tianrui Li, **Junbo Zhang**.
Approach for Dynamically Updating Set Approximations Based Rough Sets While Condition Attributes

Value Coarsening and Refining.
Proceedings of 2010 IEEE International Conference on Intelligent Systems and Knowledge Engineering (ISKE), Hangzhou, China, pp. 27-32, 2010.

21. Shaoyong Li, Tianrui Li, **Junbo Zhang**.
Updating Approximations Dynamically in Dominance-based Rough Sets.
Proceedings of the International Conference on E-Business Intelligence (ICEBI), Kunming, China, pp. 534-540, 2010.

22. Dun Liu, Tianrui Li, **Junbo Zhang**.
An Incremental Approach for Rule Induction under Coarsening and Refining of Attribute Values in E-Business Systems.
Proceedings of the International Conference on E-Business Intelligence (ICEBI), Kunming, China, pp. 541-547, 2010.

Awards & Honors

- National Scholarship, China, 2013.
- National Scholarship, China, 2012.
- "Si Shi Yang Hua'' Medal in Southwest Jiaotong University, 2012. (Top 1/1000, the students' top honor of Southwest Jiaotong University).
- 2009-2014 Special Grade Scholarship for PhD Students, Southwest Jiaotong University.
- First Prize in the 9th "Huawei Cup" National Postgraduate Mathematical Contest in Modeling, China, 2012.
- Second Prize in the 8th National Postgraduate Mathematical contest in Modeling, China, 2011.
- Second Prize in the 2nd "Huawei Cup" Innovation Programming Contest, China, 2010.
- Second Prize in the 6th National Postgraduate Mathematical contest in Modeling, China, 2009.
- Second Prize in the 2nd Sichuan Provincial Programming Contest, China, 2008.
- Honorable Mention in the 33th ACM Asia Programming Contest, 2008.
- Second Prize in the 1st Sichuan Provincial Programming Contest, China, 2007.
- Second Prize (the 12th Place in person) in the TopCoder Sichuan Provincial Contest, China, 2007.
- Third Prize in National Electrical Engineering Association Mathematical Contest in Modeling, China, 2007.

- The President Scholarship of School of Information Science and Technology, Southwest Jiaotong University, 2006, 2007.

- 2005-2009 Undergraduate Comprehensive Scholarship of Southwest Jiaotong University.

Projects

- [Nov. 2012 - Nov. 2013] Research on dynamic knowledge discovery techniques and efficient algorithms under granular computing, the Fostering Foundation for the Excellent Ph.D. Dissertation of Southwest Jiaotong University, China. Leader and principal investigator.

- [Otc. 2012 - Otc. 2013] Dynamic knowledge discovery system based on rough sets in cloud computing environments, the Science and Technology Planning Project of Sichuan Province, China. Leader and principal investigator.

- [Otc. 2011 - Sep.2012]Research on dynamic knowledge discovery techniques under granular computing and probabilistic rough sets, and its fast algorithms based on cloud computing, the Doctoral Innovation Funding Project of Southwest Jiaotong University, China. Leader and principal investigator.

- [Otc. 2011 - Sep.2012] Research on dynamic knowledge discovery system based on cloud computing and rough sets, the Young Software Innovation Foundation of Sichuan Province, China. Leader and principal investigator.

- [May. 2012 - July.2012] 8-week NSF REU Undergraduate Summer Research Program hosted by the Department of Computer Science, Georgia State University, the National Science Foundation, USA. Participated as a mentor.
- [Jan. 2012 - Present] Research on dynamic updating knowledge theories and algorithms based on granular computing, the National Natural Science Foundation of China.
- [Jan. 2012 - Present] Research on composite rough set models and algorithms of knowledge discovery, the National Natural Science Foundation of China.
- [Jan. 2009 - Dec. 2011] Research on incremental learning theories and methods based on granular computing, the National Natural Science Foundation of China.

Professional Activities

Review for Journal

- ACM Transactions on Knowledge Discovery from Data (TKDD)
- IEEE Transactions on Computers (TOC)
- IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)
- IEEE Transactions on NanoBioscience
- World Wide Web Journal (WWWJ)
- International Journal of Cloud Computing (IJCC)
- International Journal of Bioinformatics Research and Applications (IJBRA)
- International Journal of Computational Intelligence Systems (IJCIS)

Extracurricular Activities

- IBM Cloud Academy CON, NC, Research Triangle Park (RTP), North Carolina, USA, Apr. 2012.
- 2010 Workshop on Frontiers of Data Management, Soochow University, Suzhou, China, Dec. 2010.
- 2010 Workshop on Massive Data Mining and Knowledge Discovery, Southwest Jiaotong University, Chengdu, China, Nov. 2010.
- 2010 China Computer Federation Advanced Disciplines Lectures (the 11th issue) - Massive Data Mining and Knowledge Discovery, Southwest Jiaotong University, Chengdu, China, Nov. 2010.
- Intel Software College - Multi-core Programming for Academia, Fudan University, Shanghai, China, 2010.
- 2010 National Graduate Summer School - Data Intensive Computing and Unstructured Data Management, Renmin University of China, Beijing, China, Jul. 2010.
- 2009 Workshop on Massive Data Mining and Knowledge Discovery, Southwest Jiaotong University, Chengdu, China, Dec. 2009.
- National Graduate Summer School (Dragon Star Plan) - Data Mining, Zhejiang University, Hangzhou, China, 2009.

Skills

Operating System Linux-based OS, Mac OS, Windows XP, Windows 7/8

Programming Language C, C++, CUDA C (GPU), Python, Java, Shell, Matlab (and various) languages with practical experiences.

Others Hadoop, Spark, Theano + Pylearn2 (GPU-based Deep Learning Packages), Weka, LaTeX, MS Office(Visio, Word, Excel, PowerPoint)