Zhongfang Zhuang

70 Harley Dr. Apt 4, Worcester MA 01606

Email: zhuangzhongfang@gmail.com

Website: http://zzhuang.me

Cellphone: +1 (774) 315-8636

Objective

Seeking a full-time position that exploits challenging problems of innovative products and services.

Education

Ph.D. Candidate, Computer Science Expected Graduation: Jan 2019

Worcester Polytechnic Institute, Worcester MA, United States

- Dissertation: Deep Learning on Attributed Sequences
- Advisor: Dr. Elke Rundensteiner (rundenst@wpi.edu)
- Committee Members: Dr. Xiangnan Kong, Dr. Mohamed Eltabakh, and Dr. Philip Yu

Master of Engineering, Software Engineer Jul 2013

Beijing University of Posts and Telecommunications, Beijing, China

Bachelor of Engineering, Computer Science Jul 2011

Xi'an University of Posts and Telecommunications, Xi'an, China

Research in Deep Learning

Collaborated with R&D teams at Amadeus for fraud detection projects by analyzing user behaviors.

Identified four different scenarios and designed patentable innovative frameworks using neural networks.

Attention Network for Fraud Detection May 2018 – Oct 2018

- Designed and implemented a novel neural attention model for attributed sequence classification.
- Integrated conventional sequence attention model with the attributes from user profiles.
- Evaluated the proposed model and compared with state-of-the-art approaches to confirm its effectiveness.

Zhongfang Zhuang, Xiangnan Kong, and Elke Rundensteiner. "AMAS: Attention Model for Attributed Sequence Classification", SDM 2019.

Fraud Detection in One Shot Dec 2017 – Apr 2018

- Challenged by the real-world scenario that only one fraud case per fraud type is available.
- Designed a multimodal siamese neural network that is capable of generalizing from only one example.
- Studied and evaluated the proposed model in various real-world scenarios with diverse parameter settings.

Zhongfang Zhuang, Xiangnan Kong, Elke Rundensteiner, Aditya Arora and Jihane Zouaoui. "One-shot Learning on Attributed Sequences", IEEE Big Data 2018.

Incorporate User Feedback for Fraud Detection Mar 2017 – Dec 2017

- Identified the challenges of incorporating the feedback from human domain experts in fraud detection.
- Formulated the problem of deep metric learning on attributed sequences.
- Designed and implemented a deep learning framework to effectively learn from the human feedback.
- Evaluated the purposed model and confirmed it outperforms state-of-the-art in various mining tasks.

Zhongfang Zhuang, Xiangnan Kong, Elke Rundensteiner, Jihane Zouaoui, and Aditya Arora. "Deep Metric Learning on Attributed Sequences", in submission.

Unsupervised Attributed Sequence Embedding Jan 2016 – Feb 2017

– Proposed a new data model, the attributed sequence, for Amadeus application log files.
– Identified the challenges of using attributed sequences in fraud detection: attributed sequences are not represented as feature vectors that could be used directly by existing data mining algorithms.
– Designed a multimodal neural network model with a sequence network and an attribute network.
– Tailored an unsupervised training strategy to learn the information from attributed sequences.
– Evaluated the performance of the proposed neural network model in clustering and outlier detection tasks.
– Conducted case studies by using visualization tools and collaborating with domain experts.

Zhongfang Zhuang, Xiangnan Kong, Elke Rundensteiner, Jihane Zouaoui, and Aditya Arora. "Attributed Sequence Embedding", in submission.

PATENT APPLICATIONS IN DEEP LEARNING

Zhongfang Zhuang, Xiangnan Kong, and Elke Rundensteiner, Jihane Zouaoui, Aditya Arora. Machine Learning Systems and Methods for Attributed Sequences.
US Patent Application 16/057,025. French Patent Application FR1857430.

RESEARCH IN LARGE-SCALE DATA PROCESSING

Preference-Aware Recurring Query Optimization Oct 2014 – Dec 2015

– Formulated the problem of preference-aware recurring query optimization in the big data domain.
– Designed and implemented PRO, the first preference-aware optimizer for recurring queries on large-scale data processing platforms.
– Modeled the preference-aware recurring query optimization problem with an execution relation graph and tackled it as a pathfinding problem.
– Enabled big data processing platforms, such as Apache Hadoop and Apache Spark, dynamically optimize workload processing and maximally satisfying user preferences.

Zhongfang Zhuang, Chuan Lei, Elke Rundensteiner, and Mohamed Eltabakh. "PRO: Preference-aware Recurring Query Optimization," ACM CIKM 2016.

Zhongfang Zhuang, Chuan Lei, Elke Rundensteiner, and Mohamed Eltabakh. "Preference-aware Recurring Query Optimization," in Journal submission.

Redoop Infrastructure for Recurring Big Data Queries Jun 2014 – Aug 2014

– Developed the Redoop infrastructure, as the first full-fledged MapReduce framework, to support the processing of the recurring big data queries.
– Designed and developed a web-based interface for Redoop to visualize the performance at each stage in the job processing.

Chuan Lei, Zhongfang Zhuang, Elke Rundensteiner, and Mohamed Eltabakh. "Redoop Infrastructure for Recurring Big Data Queries," VLDB 2014.

Shared Execution of Recurring Query on Hadoop Sep 2013 – Aug 2014

– Developed Helix, the first scalable multi-query sharing engine for the recurring workloads in MapReduce.
– Helix exploits new sliced window-alignment techniques to create sharing opportunities among recurring queries without introducing additional I/O overheads or unnecessary data scans.

– Introduced a cost/benefit model for creating a sharing plan among the recurring queries, and a scheduling strategy for executing them to maximize the SLA satisfaction.

Chuan Lei, Zhongfang Zhuang, Elke Rundensteiner, and Mohamed Eltabakh. "Shared Execution of Recurring Workloads in MapReduce," VLDB 2015.

Technical Skills

Data Processing Systems: Apache Hadoop, Apache Spark

Deep Learning Tools: TensorFlow, Theano, Keras

Programming Languages: Python, Java, C++, Bash

Visualization Tools: matplotlib, plot.ly

Awards

WPI Graduate Research Innovation Exchange Finalist 2015

ACM CIKM 2016/SIGIR Travel Grants

IEEE Travel Award for Big Data 2018

Professional Service

External reviewer for EDBT 2014, 2017, VLDB 2015, ICDE 2016, SIGMOD 2015, 2017

References

Dr. Elke Rundensteiner
Founding Director, Data Science
Professor, Computer Science
Worcester Polytechnic Institute
Email: rundenst@wpi.edu

Dr. Xiangnan Kong
Assistant Professor
Computer Science Department
Worcester Polytechnic Institute
Email: xkong@wpi.edu

Dr. Mohamed Eltabakh
Associate Professor
Computer Science Department
Worcester Polytechnic Institute
Email: meltabakh@wpi.edu

Dr. Chuan Lei
Staff Member
IBM Research
Email: chuan.lei@ibm.com