Jiawei Zhang, Ph.D.

Contact Information

1101 Dexter Ave N
Seattle, WA 98109

Mobile: +1-404-769-2414
E-mail: rivulet.zhang@gmail.com

Research Interest

Interactive machine learning and data mining, large-scale data management and visualization

Work Experience

Facebook Inc, Seattle, Washington

Research Scientist — Summer 2018 - present

Uber Technologies, San Francisco, California

Software engineering intern, Data Visualization Team — Fall 2017

- Mentor: Yang Wang
- Project: Interactive machine learning and scalable WebGL-powered visualization

Purdue/DHS Visual Analytics Center of Excellence, West Lafayette, Indiana

Graduate Research Assistant — 2013-2017

- Mentor: David Ebert
- Project: Real-time social media data mining and visualization

Education Background

Purdue University, West Lafayette, Indiana

Ph.D., Computer Engineering (GPA: 3.8/4.0) — 2013-2018

- Advisor: David Ebert
- Thesis: *Context-preserving visual analytics of multi-scale spatial clustering*

Zhejiang University, Zhejiang, China

B.Eng., Computer Science (GPA: 3.95/4.0) — 2009-2013

- Advisor: Wei Chen
- Thesis: *Visualizing large-scale graph based on line integral convolution*

Selected Publications

Jiawei Zhang, Yang Wang, Piero Molino, Lezhi Li, David Ebert. Manifold: A Model-Agnostic Framework for Interpretation and Diagnosis of Machine Learning Models. *IEEE Conference on Visual Analytics Science and Technology*, 2018 (conditionally accepted).

Jiawei Zhang, Chittayong Surakitbanharn, Niklas Elmqvist, Ross Maciejewski, Zhenyu Qian, David Ebert. TopoText: Context-Preserving Text Data Exploration Across Multiple Spatial Scales. *Proceedings of the ACM Conference on Human Factors in Computing Systems*, 2018 (Best Paper Honorable Mention, top 5%).

Jiawei Zhang, Junghoon Chae, Chittayong Surakitbanharn, David Ebert. SMART: Social Media Analytics and Reporting Toolkit. *IEEE Visualization Workshop on Visualization in Practice*, 2017.

Jiawei Zhang, Abish Malik, Benjamin Ahlbrand, Niklas Elmqvist, Ross Maciejewski, David Ebert. TopoGroups: Cotext-Preserving Visual Illustration of Multi-Scale Spatial Aggregates. *Proceedings of the ACM Conference on Human Factors in Computing Systems*, 2017.

Jiawei Zhang, Benjamin Ahlbrand, Abish Malik, Junghoon Chae, Zhiyu Min, Sungahn Ko, David Ebert. A Visual Analytics Framework for Microblog Data Analysis at Multiple Scales of Aggregation. *Computer Graphics Forum (Proceedings of EuroVis)*, 35(3):441-450, 2016.

Jiawei Zhang, Junghoon Chae, Shehzad Afzal, Abish Malik, Dennis Thom, Yun Jang, Thomas Ertl, Sorin Adam Matei, and David Ebert. Visual Analytics of User Influence and Location-Based Social Networks. *Transparency in Social Media*, Springer International Publishing, 223-237, 2015.

Jiawei Zhang, et al. Real-Time Identification and Monitoring of Abnormal Events Based on Microblog and Emergency Call Data Using SMART. *IEEE Conference on Visual Analytics Science and Technology (VAST Challenge)*, 393-394, 2014.

Selected Projects

WebGL-Powered Visualization and Interactive Machine Learning (Intern project at Uber)

- Applied interactive visualization techniques to enable agnostic, comparison and ensemble processes of multiple machine learning models.
- Contributed two highly reusable and scalable (WebGL-enabled) visualization layers to an open source library: deck.gl

Multi-Scale Spatial Data Clustering and Visualization

- Explored large-scale geospatial data at multiple spatial scales using hierarchical clustering.
- Designed polygon distortion algorithms to couple multi-scale spatial clusters in a holistic visual space for context preservation.
- Developed a text summarization method to identify top K representative documents from a large text corpus based on matrix reconstruction.

Real-Time Location-Based Social Media Analysis for Situational Awareness

- Led the design and implementation of interactive systems (both front and back end) to support visual analysis of massive real-time social media data. Demonstrated the capability and stability of the system through the usage in multiple nation-wide events by various law enforcement agencies (police departments, US Coast Guard, DHS fusion centers) for situational awareness and emergency management.
- Developed an anomaly detection method using topic modeling and seasonal-trend decomposition.
- Designed interactive interfaces to involve human knowledge in the evaluation and refinement of the topic classification process.

Massive Crowd Movement Analysis and Visualization

- Led the development of the back-end architecture including three microservices: kernel density spatial aggregation, trajectory data management and communication network management.
- Designed a scalable schemaless module for massive trajectory data based on geohash and sequence clustering to enable efficient nearest neighbor search and similarity search.

Honors and Awards

Best Paper Honorable Mention, ACM CHI Conference	2018
VAST Challenge Honorable Mention: *Compelling Narrative Debrief*, IEEE	2015
VAST Challenge Honorable Mention: *Sponsor's Award for Novel Visualization*, IEEE	2014
RCA Zworykin Scholarship, Purdue University	2013
Scholarship for Outstanding Merits, First Class (Top 5%), Zhejiang University	2009
Scholarship for Outstanding Students, Zhejiang University	2009

Technical Skills

- Programming: Java, Python, JavaScript (ES6), PHP, HTML/CSS, C++, C
- Web development: React/Redux, AngularJS, Node, Flask
- Visualization & Graphics: WebGL, SVG (D3.js), OpenGL, Processing
- Data management & Machine learning: Apache Kafka, Apache Solr (Lucene), MongoDB, SQL, scikit-learn, NLTK, Mallet