Jiawei Zhang, Ph.D.

CONTACT
INFORMATION

1101 Dexter Ave N
Seattle, WA 98109

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

WORK
EXPERIENCE

Facebook Inc, Seattle, Washington   Summer 2018 - present

Research Scientist

Uber Technologies, San Francisco, California   Fall 2017

Software engineering intern

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

Purdue/DHS Visual Analytics Center of Excellence, West Lafayette, Indiana   2013-2017

Graduate Research Assistant

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

EDUCATION
BACKGROUND

Purdue University, West Lafayette, Indiana   2013-2018

Ph.D., Computer Engineering (GPA: 3.8/4.0)

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

Zhejiang University, Zhejiang, China   2009-2013

B.Eng., Computer Science (GPA: 3.95/4.0)

- Advisor: Wei Chen
- Thesis: *Analyzing large-scale network data using 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.

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, Abish Malik, Benjamin Ahlbrand, Niklas Elmqvist, Ross Maciejewski, David Ebert. TopoGroups: Context-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.

SELECTED
PROJECTS

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

- Created an interactive and generic framework to enable agnostic, comparison and ensemble processes of multiple machine learning models for classification and regression tasks.
- Contributed two highly reusable and scalable (WebGL-enabled) visualization layers to an open source library: deck.gl
- Showed exceptional performance and yielded a first-author full paper in IEEE VAST conference.

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 perceptual 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 real-time analysis of massive 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 topic/event detection method using LDA topic modeling and time-series analysis (Seasonal and Trend decomposition using Loess).
- Explored neural embeddings of short text (i.e., Twitter) and applied recurrent neural networks for disaster-related content classification and sentiment analysis.

Massive Crowd Movement Analysis (IEEE VAST Data Challenge 2015)

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

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, C++
- Data management: Apache Kafka, Apache Solr (Lucene), MongoDB, SQL
- Maching learning: PyTorch, scikit-learn, NLTK, Mallet
- Visualization & Web: WebGL/OpenGL, SVG (D3.js), React/Redux, Node