

Kaiqi Zhao

School of Computer Science and Engineering,
Nanyang Technological University
Blk N4, B3c-14,
50 Nanyang Ave, Singapore 639798

Mobile Phone: +65 8510 9918
Email: kzhao002@e.ntu.edu.sg
<http://spatialkeyword.scse.ntu.edu.sg/~kzhao>

Research Interests

I am interested in data mining and machine learning. My current research focuses on mining geographical social media data including POI recommendation, geographical topic mining, and distributed training solution for geographical topic models.

Education

- 2013 - 2017 (Expected) Ph.D. in Computer Science, Nanyang Technological University, Singapore.
Advisor: Dr. Gao Cong.
Dissertation: *Effective and Efficient Topic Mining and Exploration from Geo-Textual Data*
- 2010 - 2013 M.Eng. in Computer Engineering, Shanghai Jiao Tong University, China.
Advisor: Dr. Kenny Q. Zhu.
- 2005 - 2009 B.Eng. in Software Engineering, Huazhong University of Science and Technology, China.
B.A. in Japanese Language, Huazhong University of Science and Technology, China.

Positions

- Aug 2017 - now Nanyang Technological University, Singapore
Research Associate. Advisor: Dr. Gao Cong
- Sep 2011 - Mar 2012 Microsoft Research Asia, Beijing, China
Research Intern. Mentor: Dr. Haixun Wang

Awards

- 2016 SIGMOD 2016 Student Travel Award
- 2015 ICDE 2015 Student Travel Award
- 2013 Ph.D. Scholarship, Nanyang Technological University
- 2006 1st Prize Scholarship, Huazhong University of Science and Technology
- 2006 Excellent Leadership in Student Union, Huazhong University of Science and Technology
- 2006 Freshmen Scholarship, Huazhong University of Science and Technology

Research Projects

- 2013 - 2017 **Spatiotemporal social media mining**
Nanyang Technological University

My research on spatiotemporal social media data includes user behavior modeling and efficient spatiotemporal topic mining.

User behavior modeling — We developed a probabilistic user behavior model over points of interest (POIs) data to cater the needs of various applications, such as POI recommendation, user recommendation and aspect analysis in regions. We further studied a user behavior model with a regression component for associating POI to geo-tagged tweets.

Efficient spatiotemporal topic mining — We proposed a framework to support online topic analysis within a user specified time interval and spatial region. We also proposed a distributed training solution for geographical topic models to make them applicable to big data.

Keywords: *Topic models, Sentiment analysis, Distributed machine learning, Big data*

2013 **Entity linking using concept level co-occurrence**

Shanghai Jiao Tong University

We studied the problem of linking the correct Wikipedia concept to each ambiguous term in a given plain text. We developed an efficient algorithm to identify the terms to be linked and then find the correct concept by considering the co-occurrences with the other concepts in the context of the term. Based on this entity linking technique, we further develop an algorithm to cluster and disambiguate images by linking each image's text context to Wikipedia concepts.

Keywords: *Entity linking, Clustering*

2011 - 2013 **Action conceptualization**

Shanghai Jiao Tong University & Microsoft Research Asia

The pairs of subject-verb or verb-object are known as “actions”. We developed an algorithm to abstract the subject and object arguments of a verb into a set of noun concepts, known as the “argument concepts”. The argument concepts can be used for various applications, e.g., verb similarity calculation, word sense disambiguation and verb argument identification.

Keywords: *Verb representation, Knowledgebase*

Publications

Book Chapter

1. **Kaiqi Zhao**, Quan Yuan, Gao Cong. Spatiotemporal Topic Detection from Social Media. *In Encyclopedia of Social Network Analysis and Mining, 2nd Edition (to be published soon)*.

Journal Article

1. Quan Yuan, Gao Cong, **Kaiqi Zhao**, Zongyang Ma, and Aixin Sun. Who, where, when and what: a non-parametric bayesian approach to context-aware recommendation and search for twitter users. *ACM Transactions on Information Systems (TOIS)*, 2015.

Conference Proceedings

1. Daniel Rugeles, **Kaiqi Zhao**, Gao Cong, Manoranjan Dash, Shonali Krishnaswamy. Biclustering: An application of Dual Topic Models. *SIAM International Conference on Data Mining (SDM)*, 2017.
2. **Kaiqi Zhao**, Gao Cong, Aixin Sun. Annotating Points of Interest with Microblogs. *Proceedings of ACM International Conference on Information and Knowledge Management (CIKM)*, 2016.
3. **Kaiqi Zhao**, Yiding Liu, Quan Yuan, Lisi Chen, Zhida Chen, Gao Cong. Towards Personalized Maps: Mining User Preferences from Geo-textual Data (Demo paper). *Proceedings of the VLDB Endowment (PVLDB)*, 2016.
4. Kaiyu Feng, **Kaiqi Zhao**, Yiding Liu, Gao Cong. A System for Region Search and Exploration (Demo paper). *Proceedings of the VLDB Endowment (PVLDB)*, 2016.
5. Gao Cong, Kaiyu Feng, **Kaiqi Zhao**. Querying and Mining Geo-textual Data for Exploration: Challenges and Opportunities. *Workshop on Keyword Search and Data Exploration on Structured Data (ICDE Workshop)*, 2016.

6. **Kaiqi Zhao**, Lisi Chen, Gao Cong. Topic Exploration in Spatio-Temporal Document Collections. *Proceedings of ACM International Conference on Management of Data (SIGMOD)*, 2016.
7. Yu Gong, **Kaiqi Zhao**, Kenny Q. Zhu. Representing Verbs as Argument Concepts. *In AAAI Conference on Artificial Intelligence (AAAI)*, 2016.
8. **Kaiqi Zhao**, Gao Cong, Quan Yuan, Kenny Q. Zhu. SAR: A Sentiment-Aspect-Region Model for User Preference Analysis in Geo-tagged Reviews. *Proceedings of IEEE International Conference on Data Engineering (ICDE)*, 2015.
9. **Kaiqi Zhao**, Zhiyuan Cai, Qingyu Sui, Enxun Wei, Kenny Q. Zhu. Clustering Image Search Results by Entity Disambiguation. *Joint European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD)*, 2014.
10. Zhiyuan Cai, **Kaiqi Zhao**, Kenny Q. Zhu, Haixun Wang. Wikification via link co-occurrence. *Proceedings of ACM International Conference on Information and Knowledge Management (CIKM)*, 2013.
11. **Kaiqi Zhao**, Enxun Wei, Qingyu Sui, Kenny Q. Zhu, Eric Lo. CISC: Clustered Image Search by Conceptualization (Demo paper). *Proceedings of International Conference on Extending Database Technology (EDBT)*, 2013.

Patents

- Kenny Q. Zhu, Zhiyuan Cai, Kaiqi Zhao. Semantic Disambiguation based on Link Co-occurrences in Online Encyclopedias. Chinese Patent No. 201310471554.5 (Pending).
- Kenny Q. Zhu, Kaiqi Zhao, Zhiyuan Cai, Qingyu Sui, Enxun Wei. An Entity Clustering System for Images Returned by Web Search Engines. Chinese Patent No. 201410554684.X (Pending).

Teaching Experience

- | | |
|-----------|---|
| Fall 2016 | Nanyang Technological University, Singapore
Teaching Assistant for “ <i>Introduction to Databases (CZ2007)</i> ”
Duties: lab supervision, grading |
| Fall 2015 | Nanyang Technological University, Singapore
Teaching Assistant for “ <i>Database System Principles (CZ4031)</i> ”
Duties: grading |
| Fall 2010 | Shanghai Jiao Tong University, Shanghai, China
Teaching Assistant for “ <i>Windows Internals (CS490)</i> ” (English)
Duties: lab supervision, grading |

Professional Services

Program Committee Member

- ADMA 2017, CIKM 2017.

Journal Reviewer:

- 2016: ACM Trans. on Knowledge Discovery from Data, IEEE Trans. on Big Data, World Wide Web Journal.

External Reviewer:

- 2017: WWW, DASFAA, WSDM, AAAI
- 2016: WWW, AAAI, CIKM, ICDM
- 2015: WWW, ACL, Trans. Big Data
- 2014: KDD, TKDE, CIKM, SDM, ICWSM, COLING
- Others: WWW 2013, EMNLP 2013, CIKM 2012, WWW 2011, ECML 2011