

# NANYANG TECHNOLOGICAL UNIVERSITY

## CV for Academic Staff

**Name:** Cong Gao

**School:** School of Computer Science and Engineering (SCSE)

Date of CV Revision: 20 April 2024

---

### Academic Qualifications

- |                   |                                                                                                           |
|-------------------|-----------------------------------------------------------------------------------------------------------|
| 01/2000 – 01/2004 | <b>Doctor of Philosophy</b> in Computer Science,<br>School of Computing, National University of Singapore |
| 09/1996 – 02/1999 | <b>Master of Engineering</b> (in Information Systems), Tianjin University, China                          |
| 09/1992 – 07/1996 | <b>Bachelor of Engineering</b> (in Information Systems), Tianjin University, China                        |

### Working Experience

- |                    |                                                                                                                                                                                   |
|--------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 09/2018 -- Present | <b>Professor</b><br>School of Computer Science and Engineering, Nanyang Technological University, Singapore. Cross appointment with School of Physical and Mathematical Sciences. |
| 03/2015 – 08/2018  | <b>Associate Professor</b><br>School of Computer Engineering, Nanyang Technological University, Singapore                                                                         |
| 03/2010 – 02/2015  | <b>Assistant Professor</b><br>School of Computer Engineering, Nanyang Technological University, Singapore                                                                         |
| 04/2008 -- 02/2010 | <b>Assistant Professor</b><br>Department of Computer Science, Aalborg University, Denmark                                                                                         |
| 2006 – 2008        | <b>Researcher</b> at Microsoft Research Asia                                                                                                                                      |
| 2004 – 2006        | <b>Postdoc Research Fellow</b><br>Database group, School of Informatics, the University of Edinburgh, UK                                                                          |
| 2004               | <b>Pre-doctor Researcher</b><br>School of Computing, National University of Singapore                                                                                             |
| 1999               | <b>Engineer</b> in Shenzhen Huawei Company, China                                                                                                                                 |

---

## RESEARCH SUMMARY

---

### Key Areas of Research

Data Management, Spatial Data Mining and Smart City, Recommendation Systems

Specifically, his current research interests include:

- Spatial +X Data Systems
- Machine Learning for Data Management
- Spatial-temporal Data Mining and Smart City
- Recommendation Systems

### Research Award

- Best paper runner up award, the 15th annual ACM International Conference on Web Search and Data Mining (WSDM) 2022 (Tier-1 conference)

- One of best papers by ACM SIGSPATIAL 2021
- Best paper runner up award, the 13th annual ACM International Conference on Web Search and Data Mining (WSDM) 2020 (acceptance rate 15%, 91 out of 615 submissions)
- Best paper runner-up award The 11th International Conference on Web Information System Engineering (WISE 2010)
- Top 2% Scientists Worldwide by Stanford University

### **Invited Talks, Papers and Tutorials**

- SIGMOD 2024 Tutorial: Machine Learning for Databases: Foundations, Paradigms, and Open Problems, in Proc. of the 2024 ACM SIGMOD International Conference on Management of Data, 2024
- Invited Keynote Talk at 5<sup>th</sup> International Conference on Spatial Data and Intelligence (SpatialDI) 2024: Geospatial entity representation: a step towards city foundation models
- Invited Talk at Huawei LLM security workshop: Machine Learning for Database Systems (2024)
- Invited Keynote Talk at 2023 Australasian Database Conference (ADC 2023): Empowering Database Systems with Machine Learning 2023
- Invited Talk at Oceanbase subforum of the 2023 China National Computer Congress (CNCC): Empowering Database Systems with Machine Learning 2023
- Invited Talk at Huawei Sentosa Software Technology Summit: Geospatial entity representation: a step towards city foundation models (2023)
- Invited Tutorial at CIRES PhD School: Geospatial entity representation: a step towards city foundation models 2023
- Invited Keynote Talk at The second spatial temporal data consortium of China 2021.
- Invited talk at Sino-Singapore forum 2020
- Invited Keynote Talk at The 12th ACM International WSDM Conference Cup 2019
- Invited Tutorial for Very Large Data Bases (VLDB) summer school, Beijing, China, 2017
- Invited Keynote Talk at 12th International Conference on Advanced Data Mining and Applications (ADMA), Australia, 2016
- Invited Keynote Talk at 2016 International Workshop on Mathematical Issues in Information Sciences (MIIS), Shenzhen.
- SIGMOD 16 Tutorial: Querying Geo-Textual Data: Spatial Keyword Queries and Beyond, in Proc. of the 2016 ACM SIGMOD International Conference on Management of Data (SIGMOD) 2016
- Invited keynote talk at the 4th Workshop on Keyword Search and Data Exploration on Structured Data (In conjunction with ICDE'16), 2016
- Invited keynote talk at AI Forum'16, Taiwanese Association for Artificial Intelligence, 2016: Querying, Exploring, and Mining Geospatial Social Media Data
- Invited Distinguished Young Lecturer keynote talk at the 15th International Conference on Web-Age Information Management (WAIM), Macau, 2014
- Invited talk at the summer school of WAIM, Macau, 2014
- Invited paper " Spatial Keyword Querying" by the 31st International Conference on Conceptual Modeling (ER) 2012
- Invited talk at workshop: The Second International Workshop on Spatial Information Modeling, Management and Mining (SIM3-2012)
- Invited talk at Workshop: Databases in Networked Information Systems (DNIS) 2007

### **Research Funding**

As a Principal Investigator, he has secured 42.4M to set up the Singtel-NTU corporate research lab (Scale@NTU), 3M Tier-2 grant from Ministry of Education (MOE), S\$1M from Roll-Royce, S\$1.123M grant from MDA, and S\$81K from Microsoft.

### **External Grants**

1. **Principal Investigator.** Machine Learning Enhanced Multidimensional Indexes for Modern Database Systems. Ministry of Education (MOE) Tier-2 Project. 850k. 08/2024—07/2027

2. **Principal Investigator.** Machine Learning for Parametric Query Optimization OceanBase Company Singapore. 200k. 02/2024—01/2026
3. **Principal Investigator.** Machine Learning Driven Query Optimizers for Big Database Systems. Ministry of Education (MOE) Tier-2 Project. 740k. 08/2022—07/2025
4. **Co-PI.** Towards Semantic-Aware Multimodal and Multilingual Deep Learning Systems for E-Commerce Applications. 1.7M, National Research Foundation. 2023-2026
5. **Co-PI.** Novel context-aware multivariate time series modelling for underground transportation infrastructure monitoring and management. AISG Technology Challenge (Open-Theme). 3.3M, 05/2022 –04/2024
6. **Co-PI.** TradeMaster: Reinforcement Learning-based Quantitative Trading Toolkit. 1 July 2022 to 30 June 2025. 4.7M, National Research Foundation
7. **Principal Investigator.** Collaborative AI: Spatio-temporal Network Learning for Smart Cities, Knowledge Representation with Graph Neural Networks? AISG 480k 10-Aug-2021--09-Aug-2025
8. **Lead Principal Investigator.** AI and Data Science. RIE SGUnited Jobs NRF, 945,000.00 07-Aug-2020 31-Mar-2024
9. **Principal Investigator.** Geospatial Knowledge Graphs for South Eastern Asia and its Applications in Logistics. Alibaba. 130K 3/2021—3/2022
10. **Principal Investigator.** Systems for Mining Attributed Bipartite Networks. Alibaba. 130K 3/2020—9/2021
11. **Leading Principal Investigator.** Singtel Cognitive and Artificial Intelligence Lab for Enterprises (SCALE@NTU). S\$42.4M, 12/2017—11/2022
12. **Principal Investigator.** Deep Trajectory Representation meets Big Trajectory Querying and Mining. Ministry of Education (MOE) Tier-2 Project MOE2019-T2-2-181 S\$544K 11/2020—10/2023
13. **Principal Investigator.** “Towards Intelligent Personalized Map: Data Driven Human Preference Mining,” AcRF Tier-2, MOE, S\$468K, 2/2017—1/2020
14. **Principal Investigator.** “Digital Events --- Semantic Miner,” Rolls-Royce, S\$482K, 04/2016—09/2018
15. **Principal Investigator.** “Digital Events --- Data Manager,” Rolls-Royce, S\$447K, 04/2016—09/2018
16. **Principal Investigator.** Data Driven Human Preference Analytics for Urban Computing. S\$25,000, Microsoft, 5/2015 – 8/2018
17. **Principal Investigator.** “Enhancing Location-based Services via Queries and Mobile Activity Recognition,” AcRF Tier-2, MOE, S\$431K, 12/2012—4/2016
18. **Principal Investigator.** “Build a Knowledge Base for Health Care Workers and Patients,” Microsoft Research, S\$56K, 10/2011—10/2018 (Competitive grant and the theme of the call-for-proposal is “healthcare”)
19. **Project Principal Investigator.** “Consumer behavior-based real-time contextual analytic engine” under “Rapid Rich Object Search (ROSE) Lab”, MDA project, S\$500K , 08/2012—07/2017
20. **Project Principal Investigator.** “Productive aging services” under “Lily: Active Living for the Elderly Lab,” MDA project , S\$623K, 2012—2018
21. **Collaborator.** “Herbalomic: Novel approach for fingerprinting and authentication of herbal products,” NRF, around S\$10 million , 2012—2017 (PI: Prof Tam, School of Biological Science, NTU)
22. **Associate partner.** EU project GEOCROWD (3. 57 million Euro) <http://www.geocrowd.eu/partners/>

## Internal Grants

1. **Principal Investigator.** Quantum Algorithm for Graph database. Ministry of Education (MOE) Themed Tier-1 Project 160k. 04/2024—03/2026
2. **Principal Investigator.** “Detecting and Tracking Volcanic Ash Using Social Media Data,” NTU ACE, S\$19.8K, 4/2019—3/2021

3. **Principal Investigator.** Deep Learning for spatial-temporal data mining Cong Gao School of Computer Science and Engineering Ministry of Education (MOE) Tier -1 ProjectRG114/19 (S) S\$99,000 2/2020 –2/2022
4. **Principal Investigator.** Novel Embedding Models for Representing Spatial Items. S\$80,000, AcRF Tier-1, MOE, 11/2017 – 10/2019
5. **Principal Investigator.** A Data Driven Approach to Mining Human Preferences for Urban Informatics. S\$100,000, AcRF Tier-1, MOE, 11/2015 – 5/2017
6. **Principal Investigator.** “Bridging Massive Geo-textual Data and Human Mobility Data for Data Analytics,” Tier-1 grant, MOE, S\$99.8K, 02/2013—02/2015
7. **Principal Investigator.** “Automatic Question-Answering using the Web,” Tier-1 grant MOE, S\$50K, 1/2011—2/2013
8. **Principal Investigator.** “Mining influential users in social networks,” Start-up grant of NTU, S\$100K, 08/2010—06/2014
9. **Co-Investigator.** “How Do Investors Perceive The Conflict Of Interest Of Compensation Consultants,” Tier-1 grant, MOE, S\$81.4K, 03/2014—02/2016, (PI: Zhang Huai, Nanyang Business School, NTU. No amount is allocated to Dr Cong)

### Citation Summary

Database	Citation Count		H-index
	without self-citations	with self-citations	
Google Scholar		18000+	71

### Publications

\* indicates students (including PhD students and visiting students) supervised by him

\*\* indicates research staffs supervised by him

### Journal Papers

1. Haitao Yuan\*\*, **Gao Cong**, Guoliang Li. Nuhuo: An Effective Estimation Model for Traffic Speed Histogram Imputation on A Road Network. Proceedings of the VLDB Endowment 2024 (PVLDB 2024)
2. Yanchuan Chang, Egemen Tanin, Gao Cong, Christian S. Jensen, Jianzhong Qi. Trajectory Similarity Measurement: An Efficiency Perspective. Proceedings of the VLDB Endowment 2024
3. Yue Zhao\*, Zhaodonghui Li\*, **Gao Cong**. A Comparative Study and Component Analysis of Query Plan Representation Techniques in ML4DB Studies, PVLDB 2024
4. Yile Chen\*, **Gao Cong**, Cuauhtemoc Anda. TERI: An Effective Framework for Trajectory Recovery with Irregular Time Intervals. Proceedings of the VLDB Endowment 2024 (PVLDB 2024)
5. Haicang Zhou\*, Tiantian He, Yew-Soon Ong, Gao Cong, Quan Chen. Differentiable Clustering for Graph Attention. In IEEE Transactions on Knowledge and Data Engineering, 2024.
6. Jiangneng Li\*, Zheng Wang\*, **Gao Cong**, Cheng Long, Han Mao Kiah, Bin Cui: Towards Designing and Learning Piecewise Space-Filling Curves. Proc. VLDB Endow. 16(9): 2158-2171 (2023)
7. Jiachen Shi\*, **Gao Cong**, Xiaoli Li. Learned Index Benefits: Machine Learning Based Index Performance Estimation. Proceedings of the VLDB Endowment (PVLDB) 2023 ##
8. Zhida Chen\*, **Gao Cong**, Walid G. Aref: STAR: A Cache-based Stream Warehouse System for Spatial Data. ACM Trans. Spatial Algorithms Syst. 9(4): 28:1-28:27 (2023)

9. Zheng Wang\*, Cheng Long, Gao Cong: Similar Sports Play Retrieval With Deep Reinforcement Learning. *IEEE Trans. Knowl. Data Eng.* 35(4): 4253-4266 (2023)
10. Liang Zhang, Cheng Long, Gao Cong. Region Embedding with Intra and Inter-View Contrastive Learning. In *IEEE Transactions on Knowledge and Data Engineering*, 35(9): 9031-9036 (2023)
11. Yue Chen\*, Kaiyu Feng\*\*, **Gao Cong**. Example-based Spatial Pattern Matching. *Proceedings of the VLDB Endowment (PVLDB) 2022. ##*
12. Junghoon Kim\*, Kaiyu Feng\*\*, **Gao Cong**, Diwen Zhu, Wenyuan Yu, Chunyan Miao. ABC: Attributed Bipartite Co-clustering. *Proceedings of the VLDB Endowment (PVLDB) 2022. ##*
13. Yue Zhao\*, **Gao Cong**, Jiachen Shi\*, Chunyan Miao. QueryFormer: A Tree Transformer Model for Query Plan Representation. *Proceedings of the VLDB Endowment (PVLDB) 2022. ##*
14. Yile Chen\*, Xiucheng Li\*\*, **Gao Cong**, Cheng Long, Zhifeng Bao, Shang Liu, Wanli Gu, Fuzheng Zhang. Points-of-Interest Relationship Inference with Spatial-enriched Graph Neural Networks, *Proceedings of the VLDB Endowment (PVLDB) 2022. ##*
15. Yuxing Han, Ziniu Wu, Peizhi Wu\*\*, Rong Zhu, Jingyi Yang\*, Liang Wei Tan\*, Kai Zeng, Gao Cong, Yanzhao Qin, Andreas Pfadler, Zhengping Qian, Jingren Zhou, Jiangneng Li, Bin Cui: Cardinality Estimation in DBMS: A Comprehensive Benchmark Evaluation, *Proceedings of the VLDB Endowment (PVLDB) 2022 ##*
16. Sheng Wang, Zhifeng Bao, J. Shane Culpepper, Gao Cong: A Survey on Trajectory Data Management, Analytics, and Learning. *ACM Comput. Surv.* 54(2): 39:1-39:36 (2022)
17. David A. Tedjopurnomo\*, Xiucheng Li\*, Zhifeng Bao, **Gao Cong**, Farhana Choudhury, Kai Qin. Similar Trajectory Search with Spatio-temporal Deep Representation Learning, *ACM Trans. Intell. Syst. Technol.* (ACM TIST) 2022. ##
18. Hui Luo, Zhifeng Bao, Gao Cong, J. Shane Culpepper, Nguyen Lu Dang Khoa: Let Trajectories Speak Out the Traffic Bottlenecks. *ACM Trans. Intell. Syst. Technol.* 13(1): 8:1-8:21 (2022) ##
19. Guangtao Wang\*\*, **Gao Cong**, Ying Zhang, Zhen Hai, Jieping Ye: A Synopsis Based Approach for Itemset Frequency Estimation over Massive Multi-Transaction Stream. *ACM Trans. Knowl. Discov. Data* 16(2): 29:1-29:30 (2022) ##
20. Shengnan Guo\*, Youfang Lin, Huaiyu Wan, Xiucheng Li, **Gao Cong**. Learning Dynamics and Heterogeneity of Spatial-Temporal Graph Data for Traffic Forecasting. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*. 34(11): 5415-5428 (2022) 2022 ##
21. Di Yao\*, **Gao Cong**, Chao Zhang, Xuying Meng, Rongchang Duan, Jingping Bi. A Linear Time Approach to Computing Time Series Similarity based on Deep Metric Learning. *IEEE Transactions on Knowledge and Data Engineering (TKDE)* 34(10): 4554-4571 (2022) ##
22. Daniel Rugeles\*, Zhen Hai, Manoranjan Dash, **Gao Cong**. Deterministic Inference of Topic Models via Maximal Latent State Replication. *IEEE Transactions on Knowledge and Data Engineering (TKDE)* 34(4): 1684-1695 (2022) ##
23. Kaiqi Zhao\*, **Gao Cong**, Xiucheng Li\*: PGeoTopic: A Distributed Solution for Mining Geographical Topic Models. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 34(2): 881-893 (2022) ##
24. Zhida Chen\*, Lisi Chen\*, Gao Cong, Christian S. Jensen. Location and Keyword Based Querying of Geo-Textual Data: A Survey. *VLDB J.* 30(4): 603-640 (2021) ##
25. Yile Chen\*, Xiucheng Li\*, **Gao Cong**, Cheng Long, Zhifeng Bao, Shang Liu, Wanli Gu, Fuzheng Zhang: Points-of-Interest Relationship Inference with Spatial-enriched Graph Neural Networks. *Proc. VLDB Endow.* 15(3): 504-512 (2021)
26. Yuxing Han, Ziniu Wu, Peizhi Wu\*\*, Rong Zhu, Jingyi Yang\*, Liang Wei Tan\*, Kai Zeng, Gao Cong, Yanzhao Qin, Andreas Pfadler, Zhengping Qian, Jingren Zhou, Jiangneng Li\*, Bin Cui: Cardinality Estimation in DBMS: A Comprehensive Benchmark Evaluation. *Proc. VLDB Endow.* 15(4): 752-765 (2021)
27. Kaiyu Feng\*, Tao Guo\*, **Gao Cong**, Sourav S. Bhowmick, Shuai Ma. SURGE: Continuous Detection of Bursty Regions Over a Stream of Spatial Objects. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, *IEEE Trans. Knowl. Data Eng.* 32(11): 2254-2268 (2020) ##
28. Yue Chen\*, Zhida Chen\*, Gao Cong, Walid G. Aref: SSTD: A Distributed System on Streaming Spatio-Textual Data, *Proceedings of the VLDB Endowment (PVLDB)* 13(11): 2284-2296 (2020) ##
29. Zheng Wang\*, Cheng Long, Gao Cong and Yiding Liu\*: Efficient and Effective Similar Subtrajectory Search with Deep Reinforcement Learning, *Proceedings of the VLDB Endowment (PVLDB) 2020 ##*

30. Kaiyu Feng\*, Gao Cong, Christian S. Jensen, Tao Guo\*. Finding attribute-aware similar regions for data analysis, Proceedings of the VLDB Endowment (PVLDB) 12(11): 1414-1426 (2019) ##
31. Kaiqi Zhao\*, Gao Cong, Jin-Yao Chin\*, Rong Wen. Exploring Market Competition over Topics in Spatio-Temporal Document Collections. VLDB Journal, 28(1): 123-145 (2019) 2019##
32. Yixiang Fang, Yun Li, Reynold Cheng, Nikos Mamoulis, Gao Cong: Evaluating pattern matching queries for spatial databases. VLDB J. 28(5): 649-673 (2019) ##
33. Sheng Wang, Zhifeng Bao, J. Shane Culpepper, Timos Sellis, Gao Cong. Reverse k Nearest Neighbor Search over Trajectories. IEEE Transactions on Knowledge and Data Engineering (TKDE) 30(4): 757-771 (2018) ##
34. Jialong Han, Aixin Sun, Gao Cong, Wayne Xin Zhao, Zongcheng Ji, Minh C. Phan. Linking Fine-Grained Locations in User Comments. Transactions on Knowledge and Data Engineering (TKDE). 30(1): 59-72 (2018) ##
35. Miao Xie, Sourav S. Bhowmick, Hao Su, Gao Cong, Wook-Shin Han PANDA: A System for Partial Topology-based Search on Large Networks. Proceedings of the VLDB Endowment (PVLDB) 11(12): 1966-1969 (2018) (Demo)
36. Yiding Liu\*, Tuan-Anh Nguyen Pham\*, **Gao Cong**, Quan Yuan\*. An Experimental Evaluation of Point-of-interest Recommendation in Location-based Social Networks. Proceedings of the VLDB Endowment (PVLDB), 10(10): 1010-1021 (2017) ##
37. Zhen Hai\*, **Gao Cong**, Kuiyu Chang, Peng Chen, Chunyan Miao. Analyzing Sentiments in One Go: A Supervised Joint Topic Modeling Approach. IEEE Transactions on Knowledge and Data Engineering (TKDE) 29(6): 1172-1185 (2017) ##
38. Long Guo, Dongxiang Zhang, Gao Cong, Wei Wu, Kian-Lee Tan. Influence maximization in trajectory databases. IEEE Transactions on Knowledge and Data Engineering (TKDE). 29(3): 627-641 (2017) ##
39. Miao Xie\*, Sourav S Bhowmick, Gao Cong, Qing Wang. PANDA: Towards Partial Topology-based Search on Large Networks in a Single Machine. The VLDB Journal, Springer-Verlag. 26(2): 203-228 (2017) ##
40. Xutao Li\*\*, Michael Kwok-Po Ng, Gao Cong, Yunming Ye, Qingyao Wu. MR-NTD: Manifold Regularization Nonnegative Tucker Decomposition for Tensor Data Dimension Reduction and Representation, IEEE Transactions on Neural Networks and Learning Systems, 28(8): 1787-1800 (2017) ##
41. Tuan-Anh Nguyen Pham\*, Xutao Li\*\*, **Gao Cong**, Zhenjie Zhang. A General Recommendation Model for Heterogeneous Networks. IEEE Transactions on Knowledge and Data Engineering (TKDE), 28(12): 3140-3153 (2016) #
42. Wei Wei\*\*, **Gao Cong**, Chunyan Miao, Feida Zhu and Guohui Li. Learning to Find Topic Experts in Twitter via Different Relations. IEEE Transactions on Knowledge and Data Engineering (TKDE), 28(7) pp1764-1778, 2016 #
43. Kaiqi Zhao\*, Yiding Liu\*, Quan Yuan\*, Lisi Chen\*, Zhida Chen\*, **Gao Cong**. Towards Personalized Maps: Mining User Preferences from Geo-textual Data. Proceedings of the VLDB Endowment (PVLDB) 9(13):1545-1548 (2016) (Demonstration) #
44. Kaiyu Feng\*, Kaiqi Zhao\*, Yiding Liu\*\*, **Gao Cong**. A System for Region Search and Exploration. PVLDB 9(13): 1549-1552 (2016) (Demonstration) #
45. Shuai Ma, Kaiyu Feng\*, Jianxin Li, Haixun Wang, Gao Cong, Jin-Peng Huai. Proxies for Shortest Path and Distance Queries. IEEE Transactions on Knowledge and Data Engineering (TKDE), 28(7): 1835-1850 2016 #
46. Yifeng Zeng, Xuefeng Chen, Gao Cong, Shengchao Qin, Jing Tang, Yanping Xiang. Maximizing Influence under Influence Loss Constraint in Social Networks Expert Systems With Applications. Expert Systems With Applications, Volume 55, Pages 255–267, August 2016 #
47. Quan Yuan\*, **Gao Cong**, Kaiqi Zhao\*, Zongyang Ma, and Aixin Sun. Who, where, when and what: a nonparametric bayesian approach to context-aware recommendation and search for twitter users. ACM Transactions on Information Systems (TOIS), 33(1): 2:1-2:33 2015 #
48. Xin Cao\*, **Gao Cong**; Tao Guo\*; Christian Jensen, Beng Chin Ooi. Efficient Processing of Spatial Group Keyword Queries. ACM Transactions on Database Systems (TODS) 40(2): 13 2015 #
49. Zhen Hai\*, Kuiyu Chang, Gao Cong, Christopher C. Yang. An Association-based Unified Framework for Mining Features and Opinion Words. ACM Transactions on Intelligent Systems and Technology (ACM TIST) 6(2): 26 (2015) #

50. Xin Cao\*, **Gao Cong**, Christian Jensen, Man Lung Yiu. Retrieving Regions of Interest for User Exploration. Proceedings of the VLDB Endowment (PVLDB), Volume 7, No. 9, pp.733-744, 2014 #
51. Lisi Chen\*, Yan Cui\*, **Gao Cong**, Xin Cao\*. SOPS: A System for Efficient Processing of Spatial-Keyword Publish/Subscribe. Proceedings of the VLDB Endowment (PVLDB), Volume 7, No. 13, pp.1601-1604, 2014 (Demonstration) #
52. Ying Lu\*, Jiaheng Lu, Gao Cong, Wei Wu, Cyrus Shahabi. Efficient algorithms and cost models for reverse spatial-keyword k-nearest neighbor search. ACM Transactions on Database Systems (TODS), Volume 39, No. 2 (42 pages), 2014. #
53. Bo Liu\*\*, Quan Yuan\*, **Gao Cong**, Dong Xu. Where your photo is taken: Geolocation prediction for social images. Journal of the American Society for Information Science and Technology (JASIST). Volume 65, No. 6, pp. 1232-1243, 2014. #
54. Bo Liu\*\*, **Gao Cong**, Yifeng Zeng, Dong Xu, Yeow Meng Chee. Influence Spreading Path and its Application to the Time Constrained Social Influence Maximization Problem and Beyond. IEEE Transactions on Knowledge and Data Engineering (TKDE), Volume 26, No. 8, pp.1904-1917, 2014. #
55. Yu Gu, Chunpeng Gao, Gao Cong, Ge Yu. Effective and Efficient Clustering Methods for Correlated Probabilistic Graphs. IEEE Transactions on Knowledge and Data Engineering (TKDE), Volume 26, No. 5, pp. 1117-1130, 2014. #
56. Kelvin Sim\*, Vivekanand Gopalkrishnan, Clifton Phua, Gao Cong. 3D Subspace Clustering for Value Investing. IEEE Intelligent Systems, Volume 29, No. 2, pp. 52-59, 2014. #
57. Lisi Chen\*, **Gao Cong**, Christian S. Jensen, and Dingming Wu. Spatial Keyword Query Processing: An Experimental Evaluation , Proceedings of the VLDB Endowment (PVLDB), Volume 6, No 3. Pp.217-228, Jan, 2013 #
58. Zongyang Ma, Aixin Sun, Gao Cong. On Predicting the Popularity of Newly Emerging Hashtags in Twitter. Journal of the American Society for Information Science and Technology (JASIST). 64(7): 1399-1410, 2013 #
59. Kelvin Sim\*, Ghim-Eng Yap, David R. Hardoon, Vivekanand Gopalkrishnan, Gao Cong, Suryani LUKMAN. Centroid-based Actionable 3D Subspace Clustering. IEEE Transactions on Knowledge and Data Engineering (TKDE), 25(6): 1213-1226 (2013) #
60. Kelvin Sim\*, Vivekanand Gopalkrishnan, Arthur Zimek, Gao Cong. A Survey on Enhanced Subspace Clustering. Data Mining and Knowledge Discovery (DMKD), 26(2): 332-397 (2013) #
61. Xin Cao\*, Lisi Chen\*, **Gao Cong**, Xiaokui Xiao. Keyword-aware Optimal Route Search. Proceedings of the VLDB Endowment (PVLDB) 5(11): 1136-1147, 2012 #
62. Lingkun Wu, Xiaokui Xiao, Dingxiong Deng, Gao Cong, Diwen Zhu, Shuigeng Zhou. Shortest Path and Distance Queries on Road Networks: An Experimental Evaluation. Proceedings of the VLDB Endowment (PVLDB). 5(5): 406-417, 2012 #
63. Xin Cao\*, **Gao Cong**, Christian S. Jensen, Jun Jie Ng, Beng Chin Ooi, Phan NhanTue, Dingming Wu. SWORS: A System for the Efficient Retrieval of Relevant Spatial Web Objects. Proceedings of the VLDB Endowment (PVLDB), 5(12):1914-1917, 2012 (Demonstration) #
64. Siqiang Luo, Yifeng Luo, Shuigeng Zhou, Gao Cong, Jihong Guan. DISKs: A System for Distributed Spatial Group Keyword Search on Road Networks. Proceedings of the VLDB Endowment (PVLDB), 5(12): 1966-1969, 2012 (Demonstration) #
65. Dingming Wu, Gao Cong, Christian S Jensen. A Framework for Efficient Spatial Web Object Retrieval, The VLDB Journal, 21(6): 797-822, 2012 #
66. **Gao Cong**, Wenfei Fan, Anastasios Kementsietsidis, Jianzhong Li, and Xianmin Liu. Partial Evaluation for Distributed XPath Query Processing and Beyond. ACM Transactions on Database Systems (TODS), 37(4): 32 (45 pages), 2012. #
67. Bo Qu, Gao Cong, Cuiping Li, Aixin Sun, Hong Chen. An Evaluation of Classification Models for Question Topic Categorization. Journal of the American Society for Information Science and Technology (JASIST), 63(5): 889-903, 2012 #
68. Xin Cao\*, **Gao Cong**, Bin Cui, Christian S. Jensen, Quan Yuan. Approaches to Exploring Category Information for Question Retrieval in Community Question-Answer Archives. ACM Transactions on Information Systems (TOIS), 30(2):7 (30+ pages), 2012 #
69. Dingming Wu, Man Lung Yiu, Gao Cong, Christian S. Jensen. Joint Top-K Spatial Keyword Query Processing. IEEE Transactions on Knowledge and Data Engineering (TKDE), 24(10): 1889-1903, 2012 #

70. **Gao Cong**, Wenfei Fan, Floris Geerts, Jianzhong Li, and Jizhou Luo. On the Complexity of View Update Analysis and its Application to Annotation Propagation. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 24 (3): 506-519, 2012 #
71. Zhe Zhao, Bin Cui, Gao Cong, Zi Huang, and Heng Tao Shen. Extracting Representative Motion Flows for Effective Video Retrieval. *Multimedia Tools and Applications (MTAP)*, 58(3): 687-711, 2012
72. Junjie Yao, Bin Cui, Gao Cong, Y Huang: Evolutionary taxonomy construction from dynamic tag space. *World Wide Web*, 15(5-6): 581-602, 2012
73. **Gao Cong**, Yingxin Li. Updating XML views and querying XML views with update syntax. *International Journal of Computational Science and Engineering (IJCSE)*, 5(2):118-126, 2010 [invited paper]
74. Xin Cao\*, **Gao Cong**, Christian S. Jensen. Retrieving Top-k Prestige-based Relevant Spatial Web Objects. *Proceedings of the VLDB Endowment (PVLDB)*, 3(1):373-384, 2010 #
75. Xin Cao\*, **Gao Cong**, Christian S. Jensen. Mining Significant Semantic Locations from GPS Data. *Proceedings of the VLDB Endowment (PVLDB)*, 3(1): 1009-1020 2010. #
76. **Gao Cong**, Christian S. Jensen, Dingming Wu. Efficient Retrieval of the Top-k Most Relevant Spatial Web Objects. *Proceedings of the VLDB Endowment (PVLDB)*, 2(1):337-348, 2009. [Acceptance rate 16.7%] #
77. Byron Choi, **Gao Cong**, Wenfei Fan, Stratis D. Viglas. Updating Recursive XML Views of Relations. *Journal of Computer Science and Technology*, 23(4):516-537, 2008.

## Conference Papers

### —2024—

1. Jingyi Yang\*, **Gao Cong**. PLATON: Top-down R-tree Packing with Learned Partition Policy. *SIGMOD 2024*.
2. Zizhong Meng\*, Xin Cao, **Gao Cong**. Selectivity Estimation for Queries Containing Predicates over Set-Valued Attributes. *SIGMOD 2024*.
3. Songsong Mo\*, Yile Chen\*, Hao Wang\*\*, **Gao Cong**, Zhifeng Bao. Lemo: A Cache-Enhanced Learned Optimizer for Concurrent Queries. *SIGMOD 2024*.
4. Jiayi Xie\*, Shang Liu\*, **Gao Cong**, Zhenzhong Chen. UnifiedSSR: A Unified Framework of Sequential Search and Recommendation. *the Web Conf 2024*
5. Yang Xu\*\*, **Gao Cong**, Lei Zhu, Lizhen Cui. MMPOI: A Multi-Modal Content-Aware Framework for POI Recommendations. *the Web Conf 2024*
6. Zheng Wang\*, Cheng Long, Gao Cong, and Christian S. Jensen. Collectively Simplifying Trajectories in a Database: A Query Accuracy Driven Approach. In *IEEE International Conference on Data Engineering, 2024. ICDE'24*
7. Tangwen Qian\*, Yile Chen\*, **Gao Cong**, Yongjun Xu, Fei Wang. AdapTraj: A Multi-Source Domain Generalization Framework for Multi-Agent Trajectory Prediction. *IEEE International Conference on Data Engineering (ICDE 2024)*
8. Yue Jiang\*, Xiucheng Li\*, Yile Chen\*, Shuai Liu\*, Weilong Kong\*\*, Antonis F. Lentzakis, **Gao Cong**<sup>^</sup>. SAGDFN: A Scalable Adaptive Graph Diffusion Forecasting Network for Multivariate Time Series Forecasting. *IEEE International Conference on Data Engineering (ICDE 2024)*
9. Xiaofan Li\*\*, **Gao Cong**, Rui Zhou. Quantum Algorithm for Maximum K-Plex Problem. *IEEE International Conference on Data Engineering (ICDE 2024)*
10. Kethmi Hirushini Hettige, Jiahao Ji, Shili Xiang, Cheng Long, Gao Cong, and Jingyuan Wang. AirPhyNet: Harnessing Physics-Guided Neural Networks for Air Quality Prediction. In *The Twelfth International Conference on Learning Representations, 2024. ICLR'24*

### —2023—



11. Gu Tu\*, Kaiyu Feng\*, **Gao Cong**, Cheng Long, Zheng Wang, Sheng Wang. The RLR-Tree: A Reinforcement Learning Based R-Tree for Spatial Data. SIGMOD 2023
12. Pasquale Balsebre\*, Weiming Huang\*\*, Dezhong Yao\*\*, **Gao Cong**, Zhen Hai. Mining Geospatial Relationships from Text. SIGMOD 2023
13. Shang Liu\*, **Gao Cong**, Kaiyu Feng\*\*, Wanli Gu, Fuzheng Zhang. Effectiveness Perspectives and a Deep Relevance Model for Spatial Keyword Queries. SIGMOD 2023
14. Yufan Sheng, Xin Cao, Yixiang Fang, Kaiqi Zhao, Jianzhong Qi, Gao Cong, Wenjie Zhang. WISK: A Workload-aware Learned Index for Spatial Keyword Queries. SIGMOD 2023
15. Yi Li\*, Weiming Huang\*\*, **Gao Cong**, Hao Wang, Zheng Wang. Urban Region Representation Learning with OpenStreetMap Building Footprints. KDD 2023
16. Qianru Zhang, Zheng Wang\*, Cheng Long, Chao Huang, Siu-Ming Yiu, Yiding Liu, Gao Cong, Jieming Shi. Online Anomalous Subtrajectory Detection on Road Networks with Deep Reinforcement Learning. IEEE ICDE 2023
17. Shuai Liu\*, Xiucheng Li\*, **Gao Cong**, Yue Jiang, Yile Chen, Multivariate Time-series Imputation with Disentangled Temporal Representations. ICLR 2023

—2022—

18. Zizhong Meng\*, **Gao Cong**, Peizhi Wu, Rong Zhu, Shuai Ma. Unsupervised Selectivity Estimation by Integrating Gaussian Mixture Models and Autoregressive Model", EDBT 2022
19. Dezhong Yao\*\*, Yuhong Gu, **Gao Cong**, Hai Jin, Xinqiao Lv. Entity Resolution with Hierarchical Graph Attention Networks, Proc. of ACM International Conference on Management of Data (SIGMOD). 2022
20. Jingyi Yang\*, Peizhi Wu\*, **Gao Cong**, Xiao He, Tieying Zhang. SAM: Database Generation from Query Workload with Supervised Autoregressive Model, Proc. of ACM International Conference on Management of Data (SIGMOD). 2022
21. Junghoon Kim\*, Siqiang Luo, **Gao Cong**, DMCS : Density Modularity based Community Search, Proc. of ACM International Conference on Management of Data (SIGMOD). 2022
22. Pasquale Balsebre\*, Dezhong Yao\*\*, **Gao Cong**, Zhen Hai. Geospatial Entity Resolution, The Web Conference (WWW) 2022
23. Jinyao Chin\*, Yile Chen\*, **Gao Cong**: The Datasets Dilemma: How Much Do We Really Know About Recommendation Datasets? WSDM 2022
24. Di Yao, Haonan Hu, Lun Du, Gao Cong, Shi Han, Jingping Bi. TrajGAT: A Graph-based Long-term Dependency Modeling Approach for Trajectory Similarity Computation," KDD 2022

—2021—

25. Peizhi Wu\*\*, **Gao Cong**, A Unified Deep Model of Learning from both Data and Queries for Cardinality Estimation. Proc. of ACM International Conference on Management of Data (SIGMOD). 2021
26. Zhida Chen\*, **Gao Cong**, Walid G. Aref. STAR: A Cache-based Distributed Warehouse System for Spatial Data Streams. ACM SIGSPATIAL 2021
27. Xiucheng Li\*\*, Jin Yao Chin\*, Yile Chen\*, **Gao Cong**. Sinkhorn Collaborative Filtering, The Web Conference (WWW) 2021.
28. Zheng Wang\*, Cheng Long, Gao Cong. Error-Bounded Online Trajectory Simplification with Multi-Agent Reinforcement Learning, KDD 2021
29. Zheng Wang\*, Cheng Long, **Gao Cong**. Trajectory Simplification with Reinforcement Learning, ICDE 2021
30. Yile Chen\*, Xiucheng Li\*, **Gao Cong**, Cheng Long, Zhifeng Bao. Robust Road Network Representation Learning: When Traffic Patterns Meet Traveling Semantics, CIKM 2021.

—2020—

31. Junghoon Kim\*, Tao Guo\*, Kaiyu Feng\*\*, **Gao Cong**, Arijit Khan, and Farhana Choudhury, "Densely Connected User Community and Location Cluster Search in Location-Based Social Networks", Proc. of ACM International Conference on Management of Data (SIGMOD) 2020
32. Zhida Chen\*, **Gao Cong**, Walid G. Aref: STAR: A Distributed Stream Warehouse System for Spatial Data. SIGMOD Conference 2020: 2761-2764 (Demo)
33. Shang Liu\*, Wanli Gu, **Gao Cong**, Fuzheng Zhang: Structural Relationship Representation Learning with Graph Embedding for Personalized Product Search. CIKM 2020: 915-924
34. Yile Chen\*, Cheng Long, **Gao Cong**, Chenliang Li. Context-aware Deep Model for Joint Mobility and Time Prediction, WSDM 2020
35. Lucas Vinh Tran\*, Yi Tay, Shuai Zhang, **Gao Cong**, Xiaoli Li. HyperML: A Boosting Metric Learning Approach in Hyperbolic Space for Recommender Systems. WSDM 2020
36. Yiding Liu\*, Kaiqi Zhao\*, **Gao Cong**, Zhifeng Bao, Online Anomalous Trajectory Detection with Deep Generative Sequence Modeling, ICDE2020
37. Xiucheng Li\*, **Gao Cong**, Yun Cheng, Spatial Transition Learning on Road Networks with Deep Probabilistic Models, ICDE 2020
38. Shanshan Feng\*; Lucas Vinh Tran\*; **Gao Cong**; Lisi Chen; Jing Li; Fan Li: HME: A Hyperbolic Metric Embedding Approach for Next-POI Recommendation, SIGIR 2020
39. Ziyang Wang; Wei Wei; Gao Cong; Xiaoli Li; Xianling Mao; Minghui Qiu: Global Context Enhanced Graph Neural Networks for Session-based Recommendation, SIGIR 2020

—2019—

40. Di Yao\*, Gao Cong, Chao Zhang, Jingping Bi. Computing Trajectory Similarity in Linear Time: A Generic Seed-Guided Neural Metric Learning Approach. Proceedings of the 35th IEEE International Conference on Data Engineering (ICDE) 2019#
41. Xiucheng Li\*, Gao Cong, Aixin Sun, Learning Travel Time Distributions with Deep Generative Model. The Web Conference 2019 proceedings (WWW conference) # (Acceptance: 16%)
42. Zhen Wang\*, Cheng Long, Gao Cong, and Ce Ju, "Effective and Efficient Sports Play Retrieval with Deep Representation Learning", ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2019 (Acceptance: about 170/1200 = 14.2%)
43. Lucas Vinh Tran\*, Tuan-Anh Nguyen Pham\*, Yi Tay, Yiding Liu\*, Gao Cong, Xiaoli Li. Interact and Decide: Medley of Sub-Attention Networks for Effective Group Recommendation. ACM SIGIR 2019

—2018—

44. Tao Guo\*, Kaiyu Feng\*, **Gao Cong**, Zhifeng Bao. Efficient Selection of Geospatial Data on maps for Interactive Visualized Exploration. Proc. of the 2018 ACM SIGMOD International Conference on Management of Data (SIGMOD) [acceptance rate 22%] #
45. Xiangyu Ke, Arijit Khan, and Gao Cong, "Finding Seeds and Relevant Tags Jointly: For Targeted Influence Maximization in Social Networks", in Proc. of ACM International Conference on Management of Data 2018(SIGMOD). [acceptance rate 22%]#
46. Tao Guo, Mingzhao Li, Peishan Li, Zhifeng Bao, Gao Cong. POIsam: a System for Efficient Selection of Large-scale Geospatial Data on Maps. Proc. of the 2018 ACM SIGMOD International Conference on Management of Data (SIGMOD) (Demonstration)
47. Jin Yao Chin\*, Kaiqi Zhao\*, Shafiq Joty, Gao Cong. ANR: Aspect-based Neural Recommender Proceedings of the CIKM 2018: 147-156#
48. Yiding Liu\*, Kaiqi Zhao\*, Gao Cong. Efficient Similar Region Search with Deep Metric Learning. ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2018. #

49. Dezhong Yao\*\*, Peilin Zhao, Gao Cong, Tuan-Anh Nguyen Pham. High-dimensional Similarity Learning via Dual-sparse Random Projection. Proceedings of the 27th International Joint Conference on Artificial Intelligence and the 23rd European Conference on Artificial Intelligence (IJCAI-ECAI-2018) #
50. Ali Zonoozi\*, Kim Jung Jae, Xiaoli Li, Gao Cong. Periodic-CRN: A Convolutional Recurrent Model for Crowd Density Prediction with Recurring Periodic Patterns. Proceedings of the 27th International Joint Conference on Artificial Intelligence and the 23rd European Conference on Artificial Intelligence (IJCAI-ECAI-2018) #
51. Xiucheng Li\*, Kaiqi Zhao\*, **Gao Cong**, Christian S. Jensen, Wei Wei. Deep Representation Learning for Trajectory Similarity Computation. Proceedings of the 34rd IEEE International Conference on Data Engineering (ICDE) 2018 [acceptance rate 25%] #
52. Shanshan Feng\*, **Gao Cong**, Arijit Khan, Yong Liu, Chee Yeow Meng. Inf2vec: Latent Representation Model for Social Influence Embedding. Proceedings of the 34rd IEEE International Conference on Data Engineering (ICDE) 2018 [acceptance rate 25%] #
53. Yixiang Fang\*, Reynold Cheng, **Gao Cong**, Nikos Mamoulis, Yun Li. On Spatial Pattern Matching. Proceedings of the 34rd IEEE International Conference on Data Engineering (ICDE) 2018 [acceptance rate 25%] #
54. Kaiyu Feng\*, Tao Guo\*, **Gao Cong**, Sourav S. Bhowmick, Shuai Ma. SURGE: Continuous Detection of Bursty Regions over a Stream of Spatial Objects. Proceedings of the 34rd IEEE International Conference on Data Engineering (ICDE) 2018 (short paper)

—2017—

55. Tao Guo\*, Xin Cao\*, **Gao Cong**, Jiaheng Lu, Xuemin Lin. Distributed Algorithms on Exact Personalized PageRank. Proc. of the 2017 ACM SIGMOD International Conference on Management of Data (SIGMOD): 479-494 [acceptance rate 25%] #
56. Zhida Chen\*, **Gao Cong**, Zhenjie Zhang, Tom Fu, Lisi Chen, "Distributed Publish/Subscribe Query Processing on the Spatio-Textual Data Stream", Proceedings of the 33rd IEEE International Conference on Data Engineering (ICDE) 2017: 1095-1106 [acceptance rate 16%] #
57. Xiucheng Li\*, Yun Cheng, Gao Cong, Lisi Chen. Discovering Pollution Sources and Propagation Patterns in Urban Area. In ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD): 1863-1872 [acceptance rate 21%] #
58. Tuan-Anh Nguyen Pham\*, Xutao Li\*\*, **Gao Cong**: A General Model for Out-of-town Region Recommendation. The 26th International World Wide Web Conference (WWW): 401-410 [acceptance rate 16%] #
59. Quan Yuan, Wei Zhang, Chao Zhang, Xinhe Geng, Gao Cong and Jiawei Han. PRED: Periodic Region Detection for Mobility Modeling of Social Media Users. ACM International Conference on Web Search and Data Mining (WSDM): 263-272 [acceptance rate 16%] #
60. Shanshan Feng\*, **Gao Cong**, Bo An, Yeow Meng Chee. POI2Vec: Geographical Latent Representation for Predicting Future Visitors. Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17): 102-108 [acceptance rate 25%] #
61. Daniel Rugeles\*, Kaiqi Zhao\*, **Gao Cong**, Manoranjan Dash, Shonali Krishnaswamy. Biclustering: An application of Dual Topic Models. In the SIAM International Conference on Data Mining (SDM): 453-461 [acceptance rate 25%] #
62. Zhuoxuan Jiang\*, Shanshan Feng\*, **Gao Cong**, Chunyan Miao, Xiaoming Li: A Novel Cascade Model for Learning Latent Similarity from Heterogeneous Sequential Data of MOOC. The 2017 Conference on Empirical Methods on Natural Language Processing (EMNLP), 2017: 2768-2773 [acceptance rate 26%] #

—2016—

63. Kaiyu Feng\*, **Gao Cong**, Sourav Bhowmick, Wen-Chih Peng, Chunyan Miao. Towards Best Region Search for Data Exploration. Proc. of the 2016 ACM SIGMOD International Conference on Management of Data (SIGMOD): 1055-1070 [acceptance rate 26%] #

64. Kaiqi Zhao\*, Lisi Chen\*, **Gao Cong**. Topic Exploration in Spatio-Temporal Document Collections. To appear in Proc. of the 2016 ACM SIGMOD International Conference on Management of Data (SIGMOD): 985-998 [acceptance rate 26%] #
65. Zongcheng Ji\*\*, Aixin Sun, Gao Cong and Jialong Han. Joint Recognition and Linking Fine-Grained Locations from Tweets. To appear in the 25th International World Wide Web Conference (WWW): 1271-1281 [Acceptance rate: 16%] #
66. Kaiqi Zhao\*, **Gao Cong**, Aixin Sun. Annotating Points of Interest with Geo-tagged Tweets. Proc. 25th ACM International Conference on Information and Knowledge Management (CIKM): 417-426. [Acceptance rate: 17.6%] #
67. Miao Li\*, Lisi Chen\*, **Gao Cong**, Yu Gu, Ge Yu. Efficient Processing of Location-Aware Group Preference Queries. Proc. 25th ACM International Conference on Information and Knowledge Management (CIKM): 559-568. [Acceptance rate: 17.6%] #
68. Ali Zonoozi\*, Qirong Ho, Shonali Krishnaswamy, and Gao Cong. ConTrack: A Scalable Method For Tracking Multiple Concepts In Large Scale Multidimensional Data. IEEE International Conference on Data Mining (ICDM): 759-768 [Acceptance rate: 10%] #

—2015—

69. Xutao Li\*\*, **Gao Cong**, Xiaoli Li. Rank-GeoFM: A Ranking based Geographical Factorization Method for Point of Interest Recommendation. To appear in Proc. of the 38th International ACM SIGIR Conference on Research & Development in Information Retrieval (SIGIR), 2015: 433-442 [acceptance rate 20%] #
70. Shanshan Feng\*, Xutao Li\*\*, Yifeng Zeng, Gao Cong, Yeow Meng Chee, Quan Yuan: Personalized Ranking Metric Embedding for Next New POI Recommendation. IJCAI 2015: 2069-2075 [acceptance rate 26%] #
71. Xutao Li\*\*, Tuan-Anh Nguyen Pham\*, **Gao Cong**, Quan Yuan, Xiao-Li Li, Shonali Krishnaswamy. Where You instagram? Associating your instagram photos with points of interest). Proc. ACM International Conference on Information and Knowledge Management (CIKM) 2015: 1231-1240 [acceptance rate 17%]#
72. Lisi Chen\*, **Gao Cong**. Diversity-Aware Top-k Publish/Subscribe for Text Stream. Proc. of the 2015 ACM SIGMOD International Conference on Management of Data (SIGMOD): 347-362 [acceptance rate 25%] #
73. Tao Guo\*, Xin Cao\*, **Gao Cong**. Efficient Algorithms for Answering the m-Closest Keywords Query. Proc. of the 2015 ACM SIGMOD International Conference on Management of Data (SIGMOD): 405-418 [acceptance rate 25%] #
74. Lisi Chen\*, **Gao Cong**, Xin Cao\*, Kian-Lee Tan Temporal Spatial-Keyword Top-k Publish/Subscribe. Proceedings of the 31st IEEE International Conference on Data Engineering (ICDE) : 255-266 [Acceptance rate: 25%] #
75. Tuan-Anh Nguyen Pham\*, Xutao Li\*\*, **Gao Cong**, Zhenjie Zhang. A General Graph-based Model for Recommendation in Event-based Social Networks. Proceedings of the 31st IEEE International Conference on Data Engineering (ICDE) : 567-578 [Acceptance rate: 25%] #
76. Kaiqi Zhao\*, **Gao Cong**, Quan Yuan, Kenny Zhu. SAR: A Sentiment-Aspect-Region Model for User Preference Analysis in Geo-tagged Reviews. Proceedings of the 31st IEEE International Conference on Data Engineering (ICDE) : 675-686 [Acceptance rate: 25%] #
77. Miao Xie\*, Qiusong Yang, Qing Wang, Gao Cong, Gerard de Melo. DynaDiffuse: A Dynamic Diffusion Model for Continuous Time Constrained Influence Maximization. In Twenty-ninth AAAI Conference on Artificial Intelligence (AAAI) [Acceptance rate: 25%] #
78. Xuefeng Chen, Yifeng Zeng, Gao Cong, Shengchao Qin, Yanping Xiang, Yuanshun Dai. On Information Coverage for Location Category Based Point-of-Interest Recommendation. In Twenty-ninth AAAI Conference on Artificial Intelligence (AAAI) [Acceptance rate: 25%] #
79. Zongyang Ma, Aixin Sun, Quan Yuan\*, Gao Cong. A Tri-Role Topic Model for Domain-Specific Question Answering. In Twenty-ninth AAAI Conference on Artificial Intelligence (AAAI) [Acceptance rate: 25%] #
80. Jinpeng Wang\*, **Gao Cong**, Wayne Xin Zhao, Xiaoming Li. Mining User Intents in Twitter: A Semi-Supervised Approach to Inferring Intent Categories for Tweets. In Twenty-ninth AAAI Conference on Artificial Intelligence (AAAI) [Acceptance rate: 25%] #

81. Ying Lu\*, **Gao Cong**, Jiaheng Lu, Cyrus Shahabi: Efficient algorithms for answering reverse spatial-keyword nearest neighbor queries. SIGSPATIAL/GIS 2015: 82:1-82:4 (short paper)

—2014—

82. Quan Yuan\*, **Gao Cong**, Aixin Sun. Graph-based Point-of-interest Recommendation with Geographical and Temporal Influences. Proc. 23rd ACM International Conference on Information and Knowledge Management (CIKM). (10 pages) 2014. [Acceptance Rate: 20%] #
83. Zongyang Ma, Aixin Sun, Quan Yuan\*, **Gao Cong**. Tagging Your Tweets: A Probabilistic Modeling of Hashtag Annotation in Twitter. Proc. 23rd ACM International Conference on Information and Knowledge Management (CIKM). (10 pages) 2014. [Acceptance Rate: 20%] #
84. Quan Yuan\*, **Gao Cong**, Chin-Yew Lin. COM: A Generative Model for Group Recommendation. The 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD), 163-172, 2014. [Acceptance Rate: 14%] #
85. Kaiyu Feng\*, **Gao Cong**, Sourav S. Bhowmick, Shuai Ma. In Search of Influential Event Organizers in Online Social Networks. Proc. of the 2014 ACM SIGMOD International Conference on Management of Data (SIGMOD), pp.63-74, 2014. [Acceptance Rate: 25%] #
86. Zhen Hai\*, **Gao Cong**, Kuiyu Chang, Wenting Liu, Peng Cheng. Coarse-to-Fine Review Selection via Supervised Joint Aspect and Sentiment Model. Proc. of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval (SIGIR), pp.617-626, 2014. [Acceptance Rate: 20%] #
87. Siqiang Luo, Yifeng Luo, Shuigeng Zhou, **Gao Cong**, Jihong Guan. Distributed Spatial Keyword Querying on Road Networks. In 17th International Conference on Extending Database Technology (EDBT), pp.235-246, 2014. [Acceptance Rate: 20%] #
88. Shanshan Feng\*, Xuefeng Chen, **Gao Cong**, Yifeng Zeng, Yeow Meng Chee and Yanping Xiang: Influence Maximization with Novelty Decay in Social Networks. In Twenty-Eighth AAAI Conference on Artificial Intelligence (AAAI), pp.37-43, 2014. [Acceptance Rate: 28%] #

—2013—

89. Quan Yuan\*, **Gao Cong**, Zongyang Ma, Aixin Sun, Nadia Magnenat-Thalmann Who, Where, When and What: Discover Spatio-Temporal Topics for Twitter Users. In Proc. of 19th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD), pp.605-613, 2013. [Acceptance Rate: 17%] #
90. Lisi Chen\*, **Gao Cong**, Xin Cao. An Efficient Query Indexing Mechanism for Filtering Geo-Textual Data. In Proc. of the 2013 ACM SIGMOD International Conference on Management of Data (SIGMOD), pp.749-760 2013. [Acceptance Rate: 20%] #
91. Quan Yuan\*, **Gao Cong**, Zongyang Ma, Aixin Sun, Nadia Magnenat-Thalmann. Time-Aware Point-of-interest Recommendation. Proc. of the 36th International ACM SIGIR Conference on Research & Development in Information Retrieval (SIGIR), pp.363-372 2013. [Acceptance Rate: 20%] #
92. Yaqiong Liu\*, Hock Soon Seah, **Gao Cong**, Efficient Proximity Detection among Mobile Objects in Road Networks with Self-Adjustment Methods. The ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (ACM GIS), pp.124-133 2013. [Acceptance Rate: 16%] #
93. Xin Cao\*, Lisi Chen\*, **Gao Cong**, Jihong Guan, Nhan-Tue Phan, Xiaokui Xiao. KORS: Keyword-aware Optimal Route Search System. Proc. of the 29th IEEE International Conference on Data Engineering (ICDE), pp.1340-1343, 2013. (Demonstration)

—2012—

94. Bo Liu\*\*, **Gao Cong**, Dong Xu, Yifeng Zeng. Time Constrained Influence Maximization in Social Networks. Proceedings of the 12<sup>th</sup> IEEE International Conference on Data Mining (ICDM), pp.439-448, 2012. [Acceptance Rate 13%] #
95. Xin Cao\*, Lisi Chen\*, **Gao Cong**, Christian S. Jensen, Qiang Qu, Anders Skovsgaard, Dingming Wu, Man Lung Yiu. Spatial Keyword Querying. 31st International Conference on Conceptual Modeling (ER), pp.16-29, 2012. [Invited paper]

96. Shaohua Li, Gao Cong, Chunyan Miao: Author Name Disambiguation Using a New Categorical Distribution Similarity. The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD), pp. 569-584, 2012 [Acceptance Rate 23%] #
97. Quan Yuan\*, **Gao Cong**, Aixin Sun, Chin-Yew Lin, Nadia M. Thalmann. Category Hierarchy Maintenance: A Data-Driven Approach. Proc. of the 35th International ACM SIGIR Conference on Research & Development in Information Retrieval (SIGIR), pp.791-800, 2012. [Acceptance rate 20%] #
98. Zhen Hai\*, Kuiyu Chang, Gao Cong. One seed to find them all: Mining Opinion Features via Association. The 21st ACM Conference on Information and Knowledge Management (CIKM), pp.255-264, 2012. [Acceptance rate 13.4%] #
99. Zongyang Ma, Aixin Sun, Quan Yuan\*, Gao Cong. Topic-Driven Reader Comments Summarization. The 21st ACM Conference on Information and Knowledge Management (CIKM), pp.265-274, 2012. [Acceptance Rate: 13.4%] #
100. Lei Han, Guojie Song, Gao Cong, Kunqing Xie. Overlapping Decomposition for Causal Graphical Modeling. In ACM International Conference on Knowledge Discovery and Data Mining (SIGKDD), pp.114-122, 2012. [Acceptance Rate: 17.6%] #
101. Peng Tang, Siu Cheung Hui, and Gao Cong. Functional Feature Extraction and Chemical Retrieval. 24th International Conference on Scientific and Statistical Database Management (SSDBM), pp.507-525, 2012

—2011—

102. Xin Cao\*, Gao Cong, Christian Jensen, Beng Chin Ooi, Collective Spatial Keyword Querying. In ACM SIGMOD International Conference on Management of Data (SIGMOD) pp.373-384, 2011. [Acceptance Rate: 23%] #
103. Jiaheng Lu, Ying Lu\*, Gao Cong. Reverse Spatial and Textual k Nearest Neighbor Search. In ACM SIGMOD International Conference on Management of Data (SIGMOD), pp.349-360, 2011. [Acceptance Rate: 23%] #
104. Qingye Jiang, Guojie Song, Gao Cong, Kunqing Xie, Yu Wang, Wenjun Si. Simulated Annealing based Influence Maximization in Social Networks. The 25th AAAI Conference on Artificial Intelligence (AAAI), 2011. [Acceptance Rate: 25%] #
105. Wei Wei\*, **Gao Cong**, Xiaoli Li, See-Kiong Ng, Guohui Li. Integrating Community Question and Answer Archives. The 25th AAAI Conference on Artificial Intelligence (AAAI), 2011. [Acceptance Rate: 25%] #
106. Dingming Wu, Man Lung Yiu, Christian S. Jensen, Gao Cong. Efficient Continuously Moving Top-K Spatial Keyword Query Processing. Proceedings of the 27th IEEE International Conference on Data Engineering (ICDE), pp.541-552, 2011. [Acceptance Rate: 20%] #

—2010—

107. Zhao-Yan Ming, Tat-Seng Chua, Gao Cong. Exploring Domain-specific Term Weight in Archived Question Search. Proceedings of the 19th ACM Conference on Information and Knowledge Management (CIKM), 1605-1608 2010 [Acceptance rate 30%, short paper ]
108. Lars K. Schunk\*, **Gao Cong**. Using Transactional Data from ERP Systems for Expert Finding. 21st International Conference on Database and Expert Systems Applications (DEXA), pp. 267-276, 2010
109. Bin Cui, Junjie Yao, Gao Cong, and Yuxin Huang. Evolutionary Taxonomy Construction from Dynamic Tag Space. The 11th International Conference on Web Information System Engineering (WISE), pp. 105-119, 2010
110. Yu Wang, Gao Cong, Guojie Song, Kunqing Xie. Community-based Greedy Algorithm for Mining Top-K Influential Nodes in Mobile Social Networks, In ACM International Conference on Knowledge Discovery and Data Mining (SIGKDD), pp.1039-1048, July 2010. [Acceptance rate 17%] #
111. Bin Cui, Ce Zhang, Gao Cong. Content-enriched Classifier for Web Video Classification. Accepted by the 33rd Annual ACM SIGIR Conference on Research & Development in Information Retrieval (SIGIR), pp. 619-626, June 2010. [Acceptance Rate: 16.7%] #
112. Xin Cao\*, **Gao Cong**, Bin Cui, Christian S. Jensen. A Generalized Framework of Exploring Category Information for Question Retrieval in Community Question Answer Archives. The 19th International World Wide Web Conference (WWW), pp. 201-210, April 2010. [Acceptance Rate: 14%] #

113. Bin Cui, Jiakui Zhao, Gao Cong. ISIS: A New Approach for Efficient Similarity Search in Sparse Databases. The 15th International Conference on Database Systems for Advanced Applications (DASFAA), pp.231-245 April 2010. [Acceptance Rate: 23.2%]

—2009—

114. Yuanzhe Cai, Gao Cong, Xu Jia, Hongyan Liu, Jun He, Jiaheng Lu, Xiaoyong Du. Efficient Algorithm for Computing Link-based Similarity in Real World Networks. Proceedings of the 9<sup>th</sup> IEEE International Conference on Data Mining (ICDM), pp.734-739, 2009. [Acceptance Rate: 17.8%] #
115. Xin Cao\*, **Gao Cong**, Bin Cui, Christian S. Jensen and Ce Zhang. The use of categorization information in language models for question retrieval. The 18th ACM Conference on Information and Knowledge Management (CIKM), pp.265-274, 2009. [Acceptance Rate: 14.5%] #
116. Yanhong Zhou, Gao Cong, Bin Cui, Christian S. Jensen, Junjie Yao. Routing Questions to the Right Users in Online Communities. The 25th IEEE International Conference on Data Engineering (ICDE), pp.700-711, 2009. [Acceptance rate: 16.8%] #
117. Ce Zhang, Bin Cui, Gao Cong, YuJing Wang. A Revisit of Query Expansion with Different Semantic Levels, The 14th International Conference on Database Systems for Advanced Applications (DASFAA), pp.662-676, 2009. [Acceptance Rate: 20.9%]

— before 2008—

118. **Gao Cong**, Long Wang\*, Chin-Yew Lin, Y.I. Song. Finding Question-Answer Pairs from Online Forums. In the 31<sup>st</sup> Annual International ACM SIGIR Conference on Research & Development in Information Retrieval (SIGIR), pp.467-474, 2008. [Acceptance Rate: 17%] #
119. Shilin Ding\*, **Gao Cong**, Chin-Yew Lin and Xiaoyan Zhu. Using Conditional Random Fields to Extract Contexts and Answers of Questions from Online Forums. In 46th Annual Meeting of the Association for Computational Linguistics (ACL), pp.710—718, June, 2008 [Oral Presentation, Acceptance Rate: 20%] #
120. Guihua Sun\*, **Gao Cong**, Xiaohua Liu, Chin-Yew Lin, Ming Zhou. Mining sequential patterns and tree patterns to detect erroneous sentences. *Proceedings of the 21st AAAI Conference on Artificial Intelligence (AAAI)*, pp.925-930, 2007. [Acceptance Rate: 27%] #
121. Guihua Sun\*, Xiaohua Liu, Gao Cong, Ming Zhou, Zhongyang Xiong, Chin-Yew Lin. Detecting Erroneous Sentences using Automatically Mined Sequential Patterns. In the 45th Annual Meeting of the Association for Computational Linguistics (ACL), pp.81—88, June, 2007. [Acceptance Rate: 22%] #
122. **Gao Cong**, Wenfei Fan, Floris Geerts, Xibei Jia, Shuai Ma. *Improving Data Quality: Consistency and Accuracy*. In the 33rd International Conference on Very Large Data Bases (VLDB), pp.315-326 2007. [Acceptance Rate: 16.9%] #
123. **Gao Cong**, Wenfei Fan, Anastasios Kementsietsidis. Distributed Query Evaluation with Performance Guarantees. In ACM SIGMOD International Conference on Management of Data (SIGMOD), pp.509-520 2007. [Acceptance Rate: 14.5%] #
124. Wenfei Fan, Gao Cong, Philip Bohannon. Querying XML with update syntax. In ACM SIGMOD International Conference on Management of Data (SIGMOD), pp.293-304, 2007. [Acceptance Rate: 14.5%] #
125. Byron Choi, Gao Cong, Wenfei Fan, Stratis D. Viglas. Updating recursive XML views of relations. The 23rd International Conference on Database Engineering (ICDE), pp.766-775, 2007. [Acceptance rate: 18.5%] #
126. **Gao Cong**, Wenfei Fan, Floris Geerts. Annotation Propagation Revisited for Key Preserving Views. The 15th ACM Conference on Information and Knowledge Management (CIKM), pp.632-641, 2006. [Acceptance Rate: 15%] #
127. Peter Buneman, Gao Cong, Wenfei Fan, Anastasios Kementsietsidis. Using Partial Evaluation in Distributed Query Evaluation. The 32nd International Conference on Very Large Data Bases (VLDB), pp.211-222, 2006. [Acceptance Rate: 13.9%, Google Scholar: 53] #
128. Hanyu Li, Mong Li Lee, Wynne Hsu, Gao Cong. An Estimation System for for XPath Expressions. In 22<sup>th</sup> IEEE International Conference on Data Engineering (ICDE), pp.648-659, April 2006 [Acceptance Rate: 19.5%] #
129. Jialie Shen, Bin Cui, Gao Cong, Heng Tao Shen, Cui Yu. Exploring composite acoustic features for efficient

- music similarity query. In Proceedings of 14<sup>th</sup> ACM International Conference on Multimedia (ACM MM), pp.412-420, 2006. [Acceptance Rate: 16%]
130. **Gao Cong**, Bin Cui, Yingxin Li, Zonghong Zhang. Summarizing frequent patterns using profiles. In 11th International Conference on Database Systems for Advanced Applications (DASFAA), pp.171-186, 2006. [Acceptance Rate: 24.5%]
  131. **Gao Cong**, Kian-Lee Tan, Anthony K.H. Tung, Xin Xu. Mining Top-k Covering Rule Groups for Gene Expression Data. In 24<sup>th</sup> ACM SIGMOD International Conference on Management of Data (SIGMOD), pp. 670-681, 2005. [Acceptance Rate: 15.1%] #
  132. Bin Cui, Anirban Mondal, Jialie Shen, Gao Cong, Kian-lee Tan. On effective E-mail Classification via Neural Networks. In 16th International Conference on Database and Expert Systems Applications (DEXA), pp. 85-94, 2005
  133. **Gao Cong**, Kian-Lee Tan, Anthony K.H. Tung, Feng Pan. Mining Frequent Closed Patterns in Microarray Data. In 4<sup>th</sup> IEEE International Conference on Data Mining (ICDM). pp.363-366, 2004 #
  134. **Gao Cong**, Anthony K.H. Tung, Xin Xu, Feng Pan, J. Yang. FARMER: Finding interesting rule groups in microarray datasets. In 23rd ACM SIGMOD International Conference on Management of Data (SIGMOD), pp.143-154, 2004. [Acceptance Rate:16%] #
  135. Xin Xu, Gao Cong, Beng Chin Ooi, Kian-lee Tan, Anthony K.H. Tung. Semantic Mining and Analysis of Gene Expression Data. Proceedings of the 30th International Conference on Very Large Data Bases (VLDB), pp.1261-1264, 2004 (Demonstration)
  136. Feng Pan, Anthony K.H. Tung, Gao Cong, Xin Xu. COBBLER: Combining Column and Row Enumeration for Closed Pattern Discovery. In 16th International Conference on Scientific and Statistical Database Management (SSDBM), pp. 21-30, 2004
  137. **Gao Cong**, Beng Chin Ooi, Kian-Lee Tan, Anthony K.H. Tung. Go Green: Recycle and Reuse Frequent Patterns. In 20<sup>th</sup> IEEE International Conference on Data Engineering (ICDE), pp. 128-139, 2004 [Acceptance Rate: 14.3%] #
  138. Cuiping Li, Gao Cong, Anthony K.H. Tung, Shan Wang. Incremental Maintenance of Quotient Cube for Median. Proceedings of the 10<sup>th</sup> ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD), pp.226-235, 2004. [Acceptance Rate: 11.9%] #
  139. **Gao Cong**, Weesun Lee, Haoran Wu, Bing Liu. Semi-Supervised Text Classification Using Partitioned EM. In 9th International Conference on Database Systems for Advanced Applications (DASFAA), pp. 482-493, 2004 [Acceptance rate: 22%]
  140. Feng Pan, Gao Cong, Anthony K.H. Tung, Jiong Yang, Mohammed J. Zaki. CARPENTER: Finding Closed Patterns in Long Biological Datasets. In 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD), pp. 637-642, 2003. [Acceptance rate: 27.1%] #
  141. **Gao Cong**, Bing Liu. Speed-up Iterative Frequent Itemset Mining with Constraint Changes. In IEEE International Conference on Data Mining (ICDM), pp. 107-114, 2002. [Acceptance rate: 19.5%] #
  142. **Gao Cong**, Lan Yi, Bing Liu, Ke Wang. Discovering frequent substructures from hierarchical semi-structured data. In the SIAM International Conference on Data Mining (SDM), pp.175-192, 2002. [Acceptance rate: 23.3%] #

### Workshop & poster papers

1. Gao Cong, Kaiyu Feng, Kaiqi Zhao: Querying and mining geo-textual data for exploration: Challenges and opportunities. ICDE Workshops 2016: 165-168
2. Guangtao Wang, Jiayu Zhou, Jingjie Ni, Tingjin Luo, Wei Long, Hai Zhen, Gao Cong, Jieping Ye: Robust Self-Tuning Sparse Subspace Clustering. ICDM Workshops 2017: 858-865
3. **Gao Cong**, Wenfei Fan, Xibei Jia, Shuai Ma. PRATA: A system for XML publishing, integration and view maintenance, Proceedings of the UK e-Science All Hands Meeting 2006, pp: 432-435, 2006
4. **Gao Cong**. Query and Update through XML Views. Databases in Networked Information Systems-International Work (DNIS), pp.81-95, 2007 [Invited paper]
5. Shuai Ma, Kaiyu Feng, Jianxin Li, Haixun Wang, Gao Cong, Jinpeng Huai: Proxies for Shortest Path and Distance Queries. ICDE 2017: 61-62
6. Long Guo, Dongxiang Zhang, Gao Cong, Wei Wu, Kian-Lee Tan: Influence Maximization in Trajectory Databases. ICDE 2017: 27-28
7. Zongyang Ma, Aixin Sun, Gao Cong. Will This #Hashtag Be Popular Tomorrow? Proc. of the 35th



International ACM SIGIR Conference on Research & Development in Information Retrieval (SIGIR), pp.1173-1174, 2012. [Acceptance rate 36%, Poster Paper]

8. Quan Yuan\*, **Gao Cong**, Nadia Magnenat-Thalmann. Enhancing naive bayes with various smoothing methods for short text classification. In International World Wide Web Conference (WWW), pp. 645-646, 2012 [Acceptance Rate: 28%, Poster paper]
9. Ce Zhang, Yujing Wang, Bin Cui, Gao Cong. Semantic Similarity Based on Compact Concept Ontology In International World Wide Web Conference (WWW), pp.1125-1126, 2008 [Poster Paper]

### Tutorial papers

1. Gao Cong, Jingyi Yang, Yue Zhao: Machine Learning for Databases: Foundations, Paradigms, and Open Problems
2. Gao Cong, Christian S. Jensen: Querying Geo-Textual Data: Spatial Keyword Queries and Beyond. SIGMOD Conference 2016: 2207-2212

### Encyclopedia entry

1. Gao Cong, Christian S. Jensen: Spatio-textual data. Encyclopedia of Big Data Techniques 2019
2. Gao Cong, Kaiqi Zhao, Quan Yuan: Spatiotemporal topic detection from social media. Encyclopedia of Social Network Analysis and Mining 2018

### Edited Proceedings/books

1. Junhu Wang, Gao Cong, Jinjun Chen, Jianzhong Qi: Databases Theory and Applications - 29th Australasian Database Conference, ADC 2018, Gold Coast, QLD, Australia, May 24-27, 2018, Proceedings. Lecture Notes in Computer Science 10837, Springer 2018, ISBN 978-3-319-92012-2
2. Xiuzhen Zhang, Shuliang Wang, Gao Cong, Alfredo Cuzzocrea: Social Big Data: Mining, Applications, and Beyond. Complexity 2019: 2059075:1-2059075:2 (2019)
3. Gao Cong, Wen-Chih Peng, Wei Zhang, Chenliang Li, Aixin Sun. Advanced Data Mining and Applications:13<sup>th</sup> International Conference, ADMA 2017, Singapore, November 5-6, 2017, Proceedings

### Innovation

#### Technology transfer to industry

Title of License Issued	Name of company issued to	Date of issue
TECH/039/14 Who where when and what: discover spatio temporal topics for twitter users	Wizdom Pte Ltd	May 23, 2014
TD/144/13 One Seed To Find Them All: Mining Opinion Features Via Association (USPRV: 61877519)	Techbiz Xccelerator Pte Ltd	May 20, 2013

### Patents

1. NTU Ref: 2019-076. Title: Spatial Transition Learning With Deep Probabilistic Model. Inventors: 1) LI Xiucheng; 2) CONG Gao Singapore provisional patent application.
2. Yuan Quan, Gao Cong, Who, Where, When and What: A Non-Parametric Bayesian Approach To Context-Aware Recommendation And Search (Provisional application number 10201600644T, filed in 2016)
3. Gao Cong, Chin-Yew Lin. Discovering question and answer pairs (US20100063797 A1 , filed through Microsoft Research, U.S. 2008 )
4. Gao Cong, Chin-Yew Lin, Shilin Ding. Summarizing online forums into question-context-answer triples

(US20100076978 A1, filed through Microsoft Research, U.S. 2008 )

5. Gao Cong, Wenfei Fan, Anastasios Kementsietsidis.  
Methods and apparatus for evaluating XPathfilters on fragmented and distributed XMLdocuments  
(US8745082 B2, filed through Bell Labs, U.S. 2007)

## Technology Disclosure

1. 4 TD for 2023
2. Robust Road Network Representation Learning Framework Inventors: 1) CHEN Yile; 2) CONG Gao Our Ref: 2020-306
3. A Unified Deep Model Of Learning From Both Data And Queries For Cardinality EstimationInventors: 1) WU Peizhi; 2) CONG Gao Our Ref: 2020-218
4. NTU Ref: 2019-077. Title: Learning Travel Time Distributions With Deep Generative Model. Inventors: 1) LI Xiucheng; 2) CONG Gao
5. NTU Ref: 2019-076. Title: Spatial Transition Learning With Deep Probabilistic Model. Inventors: 1) LI Xiucheng; 2) CONG Gao
6. NTU Ref: 2019-280. Title: Context-Aware Deep Model For Joint Mobility And Time Prediction. Inventors: 1) CHEN Yile; 2) CONG Gao
7. TD 2018-240 Title: A System For Region Search And Exploration 2018
8. TD 2018-241 Title: Towards Personalized Maps: Mining User Preferences From Geo-Textual Data 2018
9. TD 2018-065 Geographical Latent Representation for Predicting Future Visitors
10. TD 2018-066 Efficient Selection of Geospatial Data on Maps for Interactive and Visualized Exploration
11. Title: A Non-Parametric Bayesian Approach To Context-Aware Recommendation And Search. 2016
12. Title: Two Step semi-Supervised Approach To Identifying Intent Tweets (NTU Ref: TD/262/14, Gao Cong, Wang Jinpeng ) 2015
13. Time Aware Point-of-Interest Recommendation, (Gao Cong, Quan Yuan, Aixin Sun, 2014)
14. Who Where When and What: Discover Spatio-temporal Topics for Twitter Users (Quan Yuan, Gao Cong, Aixin Sun, 2014)
15. Predicting the Popularity of Newly Emerging Hashtags in Twitter (Aixin Sun, Zongyang Ma, Gao Cong, 2014)
16. One Seed to Find Them All: Mining Opinion Features via Association (USPRV: 61877519) (Zhen Hai, Kuiyu Chang, Gao Cong), 2013

---

## TEACHING SUMMARY

---

### Courses Taught

#### Development of Teaching Materials

- Preparation of course materials
  - Development of slides, syllabus for MiniMaster course: Data Processing Ethics 2022
  - Development of slides, tutorial questions, and projects for course CZ4032 Data Mining and Analysis, 2021
  - Development of slides, tutorial questions, and projects for new course CZ4031 Database System Principles, Semester 1, 2014
  - Development of slides, tutorial questions, and projects for new course CSC421 Natural Language Processing, Semester 1, 2010
- Preparation of syllabus for several courses

- CZ4027 Natural Language Processing
- CZ4031 Database System Principles
- CZ4033 Advanced Data Management

### **Subject Instructor**

- In NTU, he has lectured at both undergraduate and postgraduate levels. He has gained experience on teaching both small and large classes ranging from tens of students to hundreds of students.
  - o CZ4032 Data Mining and Analysis. 2021
  - o CZ 2007 Introduction to databases (400 students). 2018--2022
  - o MH6201 Information Retrieval and Analysis (Master Course). 2017—2022
  - o Mini Master course: Data Processing Ethics. 2021--2022
  - o CZ4031 Database System Principles (100+ students). 2013--2017
  - o CI6205 Database Systems (Postgraduate Course, 50--100 students). 2013—2017
  - o CSC206/CPE303 Database Systems (300+ students). 2012--2013
  - o CSC421 Natural Language Processing (50+ students). 2010--2011
  - o CSC415 Information Retrieval (100+ students). 2011
- In Aalborg University, Denmark, he lectured three subjects.
  - o Introduction to Internet Technologies (Undergraduate Students). 2008, 2009
  - o Database Technology (Master Students). 2009
  - o Data Warehouse and Data Mining (Continuing Education Students). 2009

**Tutor and Lab Supervisor:** Tutor and lab supervisor for several subjects, such as CSC206/CPE303 Database Systems, CSC202 Algorithms, CZ2002 Object Oriented Design and Programming, and CZ2007 Introduction to Database.

**Student Feedback on Lectures and Tutorials.** His teaching feedback index scores are mostly above 80 for lectures and tutorials. For example,

- His feedback index for lecture CSC421 (Natural Language Processing) is 89.23, CZ4031 (Database System Principles) is 83.26, CI6205 (Database Systems) is 87.37.
- His feedback index for tutorial CPE303 is 95 (AY 10/11), and CZ2007 is 93.97.

### **Academic Supervision and Mentoring**

#### **PhD and Master Student Supervision**

- He has successfully graduated the following PhD students
  - o Liu Shang
  - o Meng Zizhong (Placement: SCALE lab, NTU)
  - o Gu Tu (Placement: Huawei)
  - o Chen Yile (Placement: SCALE lab, NTU)
  - o Zheng Wang (Placement: Huawei)
  - o Chin Jinyao (Placement: Huawei)
  - o KIM JUNGHOON (Placement: assistant professor in UNIST, Korea)
  - o Chen Yue (Placement: Alibaba)
  - o Lucas VINH Tran (Placement: Apple)
  - o Li Xiucheng (Assistant professor, HIT, China)
  - o Liu Yiding (Placement: Baidu)
  - o Daniel Rugeles
  - o Chen Zhida (Placement: Alibaba)
  - o Seyed Ali Majid Zonoozi (Placement: Grab)
  - o Feng Kaiyu 2018 (Placement: Professor, Beijing BIT)
  - o Zhao Kaiqi 2018 (Placement: Assistant professor, University of Auckland)
  - o Liu Yaqiong (co-supervised)
  - o Guo Tao 2017 (Placement: Google)
  - o Pham Nguyen Tuan-Anh 2017 (Research Fellow, NTU-Roll Royce Lab)
  - o Feng Shanshan 2017 (Research Scientist, A star)
  - o Chen Lisi, 2016 (Assistant Professor, Hong Kong Baptist University )
  - o Cao Xin, 2015 (Lecturer at Queen Belfast → Lecturer at New South Wales University, University )

- Yuan Quan, 2015 (postdoc at UIUC → Facebook Research)
  - Hai Zhen, 2015 (Research Scientist, A star)
  - Kevin Sim, 2012 (co-supervised).
- He is supervising the following PhD students as the main supervisor
    - Zhao Yue (IPP)
    - SHI Jiachen
    - Pasquale BALSEBRE
    - LI Yi
    - Liu Shuai
    - Yang Jingyi (IPP)
    - Jiang Yue (IPP)
    - Mo Songsong
    - Yin Ziqi
    - Li Jiangneng (co-supervisor)
  - He supervised a number of visiting postgraduate students from NTNU(Norway), University of Hong Kong, Peking University, Renmin University, North Eastern University, Huazhong University of Science and Technology in the last several years. Most of them have papers published in top-tier venues under his supervision.
  - He graduated 2 Master students by research (Koh Wee Boon, and Lu You)

#### **Postdoc Supervision**

- Several postdocs: Placement with Kyoto University, Facebook, HIT (China) as an associate Professor, UTHealth (US).

#### **Undergraduate Student Supervision**

- He has supervised 50+ FYP students, and 20+ URECA students.

#### **Students' Achievements and Awards**

- His PhD student Liu Yiding won the honorable mention award for best PhD thesis of SCSE NTU 2019
- His PhD student Chen Lisi won the runner-up award for best PhD thesis of SCSE NTU 2017
- His PhD students won a number of travel awards from top-tier conferences (ACM SIGIR, ACM SIGMOD, VLDB, ACM SIGKDD) in the last several years. For example, Zhao Kaiqi was awarded USD1500 for his paper published in SIGMOD 2016.
- His FYP students and URECA students were actively involved in research under his supervision. For example:
  - His FYP student Cui Yan was actively involved in a demonstration system on spatial-keyword publish/subscribe, published in VLDB14, a prestigious data management conference.
  - FYP student Ng Junjie was involved in a demonstration system on spatial-keyword querying, which is published in VLDB12, a prestigious data management conference.
  - His URECA student Phan Nhan-Tue was actively involved in a demonstration system on keyword-aware route planning system, which is published in ICDE12.
  - His FYP student Chen Long's FYP report is selected as one of three best SCE FYP reports, 2013
  - His URECA student Jonathan Samraj won the second prize, URECA poster Exhibition and Competition, NTU, 2012

#### **Teaching Awards**

- Koh Boon Hwee Scholar Tutor Award, 2012
- DSO-URECA award, DSO National Laboratories, 2012
- The second prize, URECA (Undergraduate Research Experience on CAmpus) poster Exhibition and Competition, NTU, 2012 (together with my student Jonathan Samraj)

---

#### **SERVICE SUMMARY**

## **Services to NTU**

- 2018 – Now: co-director for SCALE@NTU Corp Lab
- 2016 – Now: Assistant Chair (Alumni), School of Computer Science and Computer Engineering
- 2021—Now: School Faculty Search Committee Chair.
- 2022--2024: NTU teaching council member
- 2024: NTU Tier-2 reviewing committee
- 2022: members of School Tenure and Promotion Committee
- 2020: SCSE Faculty Search Committee Chair
- 2021: Chair of School Tenure and Promotion Committee 26
- 2020, 2022: Member of Business School's Faculty Search Committee
- 2016 -- 2017: Faculty Search Committee Chair, SCSE
- 2020: Member of Tenure and Promotion Committee 24
- 2014 – 2016: Exam Committee member, SCSE
- 2012 — 2013: Year 2 Course Coordinator
- 2010 –Now: serve for PhD Thesis Examination, PhD Oral Defense, Confirmation, TAC, etc.

## **Services to Academic Community**

### **Editorial board for journals**

- Associate Editor, ACM transactions on Database Systems (TODS ), 2016—now
  - Associate Editor, Proceeding of Very Large Database (PVLDB ), 2013—2014
  - Editorial board, Proceeding of Very Large Database (PVLDB ), 2011—2012, 2018-2023
  - Editorial board, Journal of In Journal of Intelligent Information Systems, and
  - Editorial board, Data-Centric Engineering (DCE)
  - Associate Editor: Data Science and Engineering Journal. 2024 – present
  - International Journal of Social Network Mining (IJSNM), 2012—present
  - Member of The Scientific Committee Conference: Modelling, Data Analytics and AI in Engineering
- Co-editor for second version of Encyclopedia on Social network analysis and mining
    - I am invited to co-edit the Section "Spatio-Temporal Aspects in Social Networks and Social Media."

### **Conference PC members**

- Tutorial track co-chair, WWW 2024
- PC co-chair, ICDE 2022
- Associate General Chair, KDD 2021
- PC co-chair, Workshop track of CIKM'21
- PC co-chair, PAKDD 2020 Tutorial Track
- PC Vice Chair, ICDE 2018
- PC co-chair, ADC 2018
- PC co-chair, PhD forum of ASONAM 2018
- PC co-chair, ADMA 2017
- PC co-chair for E&A track of VLDB 2014
- Senior PC for KDD'22, KDD'2021, KDD'2020, PAKDD'2021, PAKDD'2020, AAI'2021, CIKM'22, CIKM'17, CIKM'18, IEEE Big Data'17
- 2023: SIGMOD, VLDB, etc
- 2022: SIGMOD, WWW, etc
- 2021: VLDB, ICDE, WWW, etc
- 2020: VLDB, SIGMOD, KDD, ICDE, DASFAA, etc.
- 2019: VLDB, KDD, ICDE, ICDM, CIKM, DASFAA, etc.
- 2018: VLDB, SIGMOD, WWW, KDD, DASFAA, etc.
- 2017: VLDB, CIKM, WWW, WSDM, AAI, ICDM, IEEE Big Data, SSTD, SIGSpatial, DASFAA, etc.
- 2016: KDD, ICDM, VLDB, CIKM, WWW, AAI, IEEE Big Data, etc.
- 2015: ICDE, KDD, CIKM, WWW, ACL, DASFAA, WISE, WAIM, IEEE Big Data, etc.
- 2014: VLDB, ICDE, KDD, SDM, CIKM, PAKDD, WWW, ACL, EMNLP, COLING, DASFAA, WAIM
- 2013: VLDB, KDD, WWW, ICDE, ICDM, SDM, CIKM, DASFAA, IJCAI, PAKDD, WAIM, etc.

- 2012: VLDB, ICDE, NAAACL-HLT, SIGKDD, ICDM, SIGIR, CIKM, ASONAM, DASFAA, PAKDD, etc
- 2011: ICDE; EMNLP, DASFAA; ASONAM; ADC; DEXA; CIKM; SIGIR poster; AINA; PAKDD-DSDM
- 2010: KDD; ACL; ICDM ; WWW; ASONAM; DASFAA; DEXA; DNIS; ICDKE; IEEE HPCC;
- 2009: VLDB; ICDE; ICDM; ASONAM; PAKDD; DEXA; APWeb;
- 2008: SIGMOD; VLDB; KDD; ICDM; DEXA; XANTEC; WWW; Apweb; DASFAA; AReS; WAIM;

### **Journal Reviewer**

- ACM Transactions on Database Systems (TODS)
- ACM Transactions on Information Systems (TOIS)
- IEEE Transactions on Knowledge and Data Engineering (TKDE)
- ACM Transactions on Knowledge Discovery from Data (TKDD)
- Data Mining and Knowledge Discovery (DMKD)
- The VLDB journal (VLDBJ)
- Information Systems
- Knowledge and Information Systems (KAIS)
- WWW journal
- Bioinformatics
- IEEE Intelligent Systems

### **Services to Public**

- Expert Review Panel as well as Expert Approval Panel for the China-Singapore Tianjin Eco-city, 2019
- Interviewer for SINGA Joint Panel Interviews 2014
- Judges for MOE Singapore Science & Engineering Fair competition (SSEF) 2014, and 2015 Feb--Sept, 2015
- Expert mentor for MOE Innovative program 2015
- Jan 2014 -- Dec 2015 Assessor for H3 Science Research and NRP 2014 (OAFa)

### **Services to Industry**

- Consultancy services to company Meme Analytics Pte. Ltd. 2018
- Consultancy services to Data Analytics Master Program of NTU 2017, 2018
- He has licensed two TDs to two local companies

### **Services to Other Universities and Funding Agency**

- Reviewer for a grant proposal of Israel Science Foundation, 2024
- Peking University, Nankai University, UNSW, Chinese Univ of HK, University of Macau, Academic reviewing for Promotion & Tenure 2023--2024
- City University of HK, Academic reviewing for Promotion & Tenure. 2022
- Tsinghua University, Academic reviewing for recruitment. 2021
- Renmin University, Academic reviewing for Promotion & Tenure. 2020
- National Taiwan University, Academic reviewing for promotion of professor. 2020
- Hong Kong Polytechnic University, Academic reviewing for promotion of professor. 2019
- Singapore Management University, Academic reviewing for promotion of associate professor. 2018
- University of Macau, Academic reviewing for promotion of associate professor, Feb—March 2016.
- National Taiwan University, Academic reviewing for promotion of associate professor. March --July 2015
- External Examiner for several PhD students from Universities in Australia, including University of Queensland, University of Adelaide, University of New South Wales, UTS, 2015-2020
- Dissertation Committee for a PhD student from Singapore Management University, June 2016--June 2017
- Dissertation Committee for a PhD student, National University of Singapore, 2015
- Dissertation Committee for a PhD student (Gong Wei) from Singapore Management University, June 2015 --August 2016
- Dissertation Committee for a PhD student (Wang Yazhe) from Singapore Management University, Jan - August 2014

- External Reviewer, Research Grants Council (RGC) of Hong Kong, 2016, 2017, 2018, 2020--2024
- Reviewer for a grant proposal of Swiss National Science Foundation, 2009

### **Services to Societies**

- Vice Chair ACM SIGKDD Singapore Chapter