

HANCHENG CAO

hanchcao@stanford.edu | <http://hanchengcao.me> | (+1) 6503348835

EDUCATION

Stanford University

Stanford, CA

Ph.D. in Computer Science

Sep 2018 – Jun 2023 (Expected)

- Research Interest: Data Science, Ubiquitous Computing, Computational Social Science
- Rotation Advisor for Winter 2019: [Prof. Michael Bernstein](#)
- Rotation Advisor for Fall 2018: [Prof. Jure Leskovec](#)
- Awarded the James D. Plummer Graduate Fellowship – a Stanford University School of Engineering (SoE) Fellowship

Tsinghua University

Beijing, China

B.Eng. in Electronic Engineering (with honors)

Aug 2014 – Jun 2018

- GPA: 92.2/100; Rank: 12/262
- Received Outstanding Graduate Award, the China National Scholarship, the Comprehensive Scholarship, the Research Excellence Award, and the Academic Excellence Award
- Selected to Spark Scientific and Technological Innovation Fellowship (top 1.5% of 3560 Tsinghua students for outstanding research performance)

University of Maryland, College Park

College Park, MD, USA

Exchange Student

Aug 2016 – Dec 2016

- GPA: 3.83/4.0
- Contributed to collaboration between UMD Distinguished [Professor Hanan Samet's lab](#) and [Tsinghua Future Communications & Internet Lab](#)

Massachusetts Institute of Technology

Cambridge, MA, USA

Visiting Student and Research Assistant in the Human Dynamics Group, MIT Media Lab

Jun 2017 – Sep 2017

- Fostered collaboration between [MIT Human Dynamics Group](#), BNU-MIT Intellectual Innovations City Lab and [Tsinghua Future Communications & Internet Lab](#).

PUBLICATIONS

1. **H. Cao**, Z. Chen, F. Xu, Y. Li, V. Kostakos. Revisitation in Urban Space vs. Online: A Comparison across POIs, websites, and Smartphone Apps. In 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp 2019).
2. **H. Cao**, J. Feng, Y. Li, V. Kostakos. Uniqueness in the City: Urban Morphology and Location Privacy. In 2018 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp 2018).
3. **H. Cao**, F. Xu, J. Sankaranarayanan, Y. Li, H. Samet. Habit2vec: Trajectory Semantic Embedding for Living Pattern Recognition in Population. In IEEE Transactions on Mobile Computing (TMC).
4. **H. Cao**, J. Sankaranarayanan, J. Feng, Y. Li, H. Samet. Understanding Metropolitan Crowd Mobility via Mobile Cellular Accessing Data. Submitted to IEEE Transactions on Visualization and Computer Graphics (TVCG). Under Review.
5. **H. Cao***, Z. Chen*, F. Xu, Y. Li, T. Wang. When E-Commerce Meets Intimacy: An Empirical Study of Social Commerce Site Beidian. Submitted to the 13th International AAAI Conference on Web and Social Media (ICWSM 2019). Under Review.
6. M. Zeng, **H. Cao**, M. Chen, Y. Li. User Behavior Modeling, Recommendations, and Purchase Prediction During Online Shopping Festivals. In Springer Electronic Markets (EM).
7. H. Shi, **H. Cao**, X. Zhou, Y. Li, V. Kostakos, F. Sun, F. Meng, C. Zhang. Semantics-Aware Hidden Markov Model for Human Mobility. In 2019 SIAM International Conference on Data Mining (SDM 2019).
8. F. Xu, T. Xia, **H. Cao**, Y. Li, F. Sun, F. Meng. Detecting Popular Temporal Modes in Population-scale Unlabelled Trajectory Data. In 2018 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp 2018).

* Indicates equal contribution.

RESEARCH EXPERIENCE

Stanford University (Department of Computer Science)

Stanford, CA, USA

Research Assistant to [Prof. Jure Leskovec](#),

Data Science for Health

- Ongoing research work aiming at understanding the fundamental characteristics of people's daily eating habits through large-scale (~TB) log data of a health tracking app.

Massachusetts Institute of Technology (Media Lab)

Cambridge, MA, USA

Research Assistant to [Prof. Alex 'Sandy' Pentland](#), Member of National Academies, World Economic Forum Councils, Advisory Board of United Nations, Director of MIT Media Lab Human Dynamics Group, and [Prof. Xiaowen Dong](#)

Jun 2017 – Sep 2017

Purchasing Pattern Recognition in Metropolis

- Independent research work recognizing typical purchasing patterns in population from large-scale credit card transaction data via representation learning based method and Monte Carlo Simulation
- Proposed algorithms embedding people's purchasing behavior to numeric vectors that better preserved original semantics; studied demographic factors (gender, age, etc.) contributing to people's purchasing patterns and the role of social learning in shaping those purchasing patterns
- Described the study and presented results in a paper being prepared for submission as first author

University of Maryland (Department of Computer Science)

College Park, MD, USA

Research Assistant to [Prof. Hanan Samet](#), University Distinguished Professor
IEEE/ACM/AAAS/ICPR/UCGIS Fellow

Sep 2016 – Jun 2017

Project 1 - Habit2vec: Trajectory Semantic Embedding for Living Pattern Recognition in Population

- Proposed a novel method to recognize and cluster metropolitan human living patterns through semantic-rich spatial temporal data breaking through geographic constraints; introduced neural network based representation learning to represent user living habits embedded in individual trajectories through numeric vectors; evaluated the effectiveness of the proposed framework on a large-scale real-world trajectory dataset in Beijing
- Paper accepted at IEEE Transactions on Mobile Computing (TMC, IF: 4.098) as first author

Project 2 - Understanding Metropolitan Crowd Mobility via Mobile Cellular Accessing Data

- Took advantage of a large-scale fine-grained cellular tower access trace and analyzed crowd mobility on city block level recognized through road network data in Shanghai; proposed algorithms to recognize homes, workplaces and stay regions of users; validated the methodology via ground truth data collected from volunteers; developed a visual analytics procedure to discover hidden block-level correlation rules and neighborhood structures formed by crowd mobility via network analysis method
- Submitted a paper to IEEE Transactions on Visualization and Computer Graphics (TVCG, IF: 2.840) as first author

University of Melbourne (School of Computing and Information Systems)

Melbourne, Australia

Research Assistant to [Prof. Vassilis Kostakos](#) and [Prof. Yong Li](#) (Tsinghua)

Sep 2017 – Aug 2018

Project 1 - Revisitation in Urban Space vs. Online: a Comparison Across POIs, websites, and Smartphone Apps

- Studied revisitation patterns in the urban space via large-scale check-in data and localization data
- Compare online and offline revisitation behavior, which contribute to fundamental understanding of human periodic behavior
- Paper accepted in 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp 2019) as first author

Project 2 - Uniqueness in the City: Urban Morphology and Location Privacy

- Studied location uniqueness in cities through Open Street Map (OSM) data to better understand city structure and location privacy
- Proposed an efficient algorithm to re-identify user geolocations supplied by provision of surrounding Point of Interest (POI); conducted experiments in New York, Melbourne, Vancouver, and Zurich to analyze factors including reporting radius, POI density, POI composition, and distance to city center, in shaping location privacy
- Paper accepted in 2018 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp 2018) as first author

Tsinghua University (Department of Electronic Engineering)

Beijing, China

Research Assistant to [Prof. Yong Li](#), Future Communications & Internet Lab

Sep 2015 – Aug 2018

Project 1 - User Behavior Modeling, Recommendations and Purchase Prediction

- Case study on user online browsing and purchasing behaviors during a large shopping festival in China via clickstream data and shopping logs collected from China's leading e-commerce site

- Conducted detailed analysis on user browsing and shopping patterns; proposed collaborative filtering based method to recommend items for different customers; constructed customer model to predict user shopping behavior
- Paper accepted at Springer Electronic Commerce (IF: 3.818) as second author

Project 2 - Semantic-Aware HMM for Human Mobility Modeling

- Proposed a novel human mobility model, which jointly takes into account spatial and temporal activity, as well as user motivation in human mobility
- Introduced graph embedding in mobility model to capture complex semantics in mobility; proposed a von Mises-Fisher mixture clustering for grouping users of similar mobility patterns to tackle data sparsity; trained an ensemble of Hidden Markov Model in embedding space to represent group-level mobility patterns
- After conducting extensive experiments on two large-scale datasets, we found that our model outperformed baselines by a statistically significant margin in the task of next location/activity prediction; the resulting paper was accepted in 2019 SIAM International Conference on Data Mining (SDM 2019) with me as second author

Project 3 - Detecting Popular Temporal Modes in Population-scale Unlabelled Trajectory Data

- Presented the innovative idea of understanding human daily routines by detecting different popular temporal modes (i.e., how different people allocate their time)
- Proposed a novel distance metric to compare the similarity between temporal modes for clustering; using two large-scale spatial temporal datasets in Beijing and Shanghai, we successfully detected distinct and meaningful temporal modes
- Paper accepted in 2018 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp 2018) as third author

Tsinghua University (Department of Electronic Engineering)

*Student Research Training Project with [Prof. Depeng Jing](#),
Future Communications & Internet Lab*

Beijing, China
Oct 2015 – Aug 2016

Linking Physical and Cyber Space via Big Data

- Independent research work exploring the correlation between webpage browsing behavior and the location of web connection using Wi-Fi big data in Shanghai, China through co-clustering

Tsinghua University (Department of Automation)

*Student Research Training Project advised by [Prof. Lihui Peng](#),
Director of Institute of Measurement and Electronic Technology*

Beijing, China
Mar 2015 – Feb 2016

Music Performance Analysis—A Perspective from Signal Processing

- Collected music performance datasets; used signal processing and machine learning method to recognize chord usage, analyzed music genre, and performance style
- Research Group Leader

WORK EXPERIENCE

Tencent Inc.

Research Intern, Tencent Map Service, Mobile Internet Group

Beijing, China
Aug 2018 – Sep 2018

User Check-in Behavior Analysis

- Analyzed user in town and out of town check-in behavior patterns
- Proposed representation learning based algorithms to embed user and POI for location recommendation.
- Results leveraged in Tencent product.

SELECTED AWARDS AND HONORS

- The James D. Plummer Graduate Fellowship – a School of Engineering (SoE) Fellowship, Stanford University, 2018
- UbiComp 2018 Student Travel Grant, 2018
- Beijing Outstanding Graduate Award, 2018 (Highest honor for graduate set by the government of Beijing)
- Outstanding Graduate Award, Tsinghua University, 2018
- China National Scholarship, 2017 (Highest level of scholarship set by the government of China)
- Qualcomm Scholarship, 2017 (Awarded to top 33 of 2562 applicants with excellent scientific potential)
- The China Scholarship Council (CSC) Scholarship, 2016
- Zhang Mingwei Scholarship, 2016 (Awarded to students for outstanding academic performance)

- Changhong Scholarship, 2015 (Awarded to students for outstanding academic performance)
- Philobiblion Scholarship, 2016 (0.5% of 1000 applicants)
- Tsinghua Comprehensive Excellence Award, 2015–17 (Top 5% of 262 students)
- Tsinghua Research Excellence Award, 2015–17 (Top 5% of 262 students)
- Tsinghua Academic Excellence Award, 2015–17 (Top 5% of 262 students)
- 1st Prize for the 32rd National Undergraduate Physics Olympic, 2015 (Top 1%)

ADDITIONAL INFORMATION

- **Academic Service:** Student volunteer, UbiComp 2018
- **Extracurricular activities:** Clavier Team of Tsinghua Student Art Troupe, (Member: 2014 – 2018; Vice Captain 2015 – 2016), Tsinghua Science and Technology Association (Member: 2015 – 2016)
- **Computer skills and proficiencies:** C/C++, MATLAB, Python, R, SQL, D3.js, Data Structure and algorithms, Data Scraping, Machine Learning, LATEX
- **Language skills and proficiencies:** Mandarin Chinese (Native); English (Proficient: TOEFL 117/120; GRE Verbal 161, Quantitative 170, Analytical Writing 3.5); German (Elementary)