

# Hien To

---

University of Southern California  
Computer Science, University of Southern California  
3710 S. McClintock Ave, RTH 323  
Los Angeles, CA 90089

[hto@usc.edu](mailto:hto@usc.edu)  
[www-scf.usc.edu/ hto/](http://www-scf.usc.edu/hto/)  
Skype: ubriela  
Phone: +1 (123) 270-5337

**Research Interests:** *Spatial Crowdsourcing, Data Privacy, Social Media, Multivariate Time-series, Query Optimization, Multimedia.*

## Education

**University of Southern California** *2011 - Current*  
PhD Candidate in Computer Science. Advisor: Professor Cyrus Shahabi  
Thesis: *Task Assignment with Rigorous Privacy Protection in Spatial Crowdsourcing*  
Contributed to projects funded by NSF, Google, Microsoft, NGC: [Spatial Crowdsourcing](#), [MediaQ](#), [iRain](#), [iCampus](#).

**Hanoi University of Science and Technology** *2005 - 2011*  
B.E., Computer Science, 2010.

## Employment

**University of Southern California** *2011 - Present*  
*Research Assistant*  
***Spatial Crowdsourcing (SC):*** Developing new privacy-preserving techniques for efficient task assignment in spatial crowdsourcing/crowdsensing, funded by two NSF grants: : [GeoCrowd](#), [BDD](#).

My SC technology was **deployed at the Center for Hydrometeorology and Remote Sensing (CHRS) at the UC, Irvine**. Check our app [iRain](#) — a novel mobile application for crowdsourcing weather information. More on [iRain Project](#), [UNESCO](#), [US Army Corps of Engineers](#), [LA Times](#), [UCI News](#).

My SC technology was **used by U.S National Geospatial-Intelligence Agency (NGA) and many institutions (e.g., UCI, LMU, HKUST)**. Check our app [MediaQ](#) — a novel media management platform for managing crowdsourced data (pictures, videos), which were provided as services to [GeoQ](#) — a crowdsourcing platform for disaster data developed by NGA. This integration enhances the capability of GeoQ to access real-time geo-tagged images and videos from the affected population of disasters; thus, enable timely analysis of such data for disaster response and recovery. More on [NSF News](#), [MediaQ Project at USC](#), [USC Magazine](#), [PBS Newshour](#), [NATO Summit](#), [Digital Trends](#), [USC Viterbi News](#)

**Time-series Mining:** Developing mining and visualizing high-dimensional and multivariate time-series data (human sensor and medical data). More on [paper](#).

**Tweet Mining:** a) visualizing and classifying disaster-related tweets to enhance situational awareness during disaster response b) estimating social POI boundaries from geo-tagged tweets.

### **Teradata Corporation, Query Optimization Group**

*Software Engineering Intern*

*Summer 2014*

**Large Join Query Optimization:** Developed genetic algorithms for LJQO in shared-nothing architecture, which reduced the estimated cost a few times while the overhead is minimal. More on [patent](#).

### **Teradata Corporation, Query Optimization Group**

*Software Engineering*

*Summer 2012*

**Selectivity Estimation:** Developed one/multidimensional histograms that advance the state-of-the-art to improve accuracy of database statistics. A patent filed, more on [paper](#).

### **VNG Corporation**

*R&D Back-End Engineer*

*Jun 2010 - Jul 2011*

**Key-Value Store:** Developed distributed key-value store from scratch that favors availability with following features: persistent storage, memory caching, high scalability, and fault-tolerance, served as backends for 1) 2M+ online game users and 2) 6M+ monthly active social network users (news feed ranking). More on [report](#).

**Social Faceted Search:** Developed instant search engine that helps users to explore their networks, such as their friends, friends of friends, nearby friends, school friends, and friends with common interests.

### **Bkav Corporation**

*Software Developer*

*Feb 2008 - May 2010*

## Publications

### Book Chapters

**Hien To**, Cyrus Shahabi. Location Privacy in Spatial Crowdsourcing. In *Handbook on Mobile Data Privacy* (under review)

### Journal Articles

**Hien To**, Sasan Tavakkol, Seon Ho Kim, Minh Nguyen, Cyrus Shahabi. On Acquisition and Analysis of Visual Data for Crowdsourcing Disaster Response. (revision)

**Hien To**, Dung D. Vu, Won-Yong Shin, Cyrus Shahabi. SocialPOI: An Efficient Algorithm for Estimating Social POI Boundaries Using Spatio-Textual Information. (under review)

Luan Tran\*, **Hien To\***, Liyue Fan, Cyrus Shahabi. A Real-Time Framework for Task Assignment in Hyperlocal Spatial Crowdsourcing. *ACM Transactions on Intelligent Systems and Technology (TIST)*. March 2017.

**Hien To**, Liyue Fan, Gabriel Ghinita, Cyrus Shahabi. [Differentially Private Location Protection for Worker Datasets in Spatial Crowdsourcing](#). *The 2016 IEEE Transactions on Mobile Computing (TMC 2016)*

**Hien To**, Leyla Kazemi, Cyrus Shahabi. [A Server-Assigned Crowdsourcing Framework](#). *ACM Transactions on Spatial Algorithms and Systems (TSAS 2015)*. Acceptance rate 11%.

### Conference Articles

**Hien To**, Sumeet Agrawal, Seon Ho Kim, Cyrus Shahabi. [On Identifying Disaster-Related Tweets: Matching-based or Learning-based?](#) *The 2017 IEEE International Conference on Multimedia Big Data (IEEE BigMM 2017)*. April 19-21, 2017. Laguna Hills, California, USA. Acceptance rate 23.9%.

**Hien To\***, Kien Nguyen\*, Cyrus Shahabi. [Differentially Private Publication of Location Entropy](#). *The 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. Acceptance rate 18%.

Ying Lu, **Hien To**, Abdullah Alfarrarjeh, Seon Ho Kim, Yifang Yin, Roger Zimmermann and Cyrus Shahabi. [GeoUGV: User-Generated Mobile Video Dataset with Fine Granularity Spatial Metadata](#). *ACM Multimedia Systems 2016 Conference (MMSys 2016)*.

**Hien To**, Hyerim Park, Seon Ho Kim, Cyrus Shahabi. [Incorporating Geo-Tagged Mobile Videos Into Context-Aware Augmented Reality Applications](#). *The 2016 IEEE International Conference on Multimedia Big Data (IEEE BigMM 2016)*.

**Hien To**, Liyue Fan, Luan Tran, Cyrus Shahabi. [Real-Time Task Assignment in Hyper-Local Spatial Crowdsourcing under Budget Constraints](#). *The 2016 IEEE International Conference on Pervasive Computing and Communications (PerCom 2016)*. Acceptance rate 15%.

**Hien To**, Seon Ho Kim, Cyrus Shahabi. [Effectively Crowdsourcing the Acquisition and Analysis of Visual Data for Disaster Response](#). *The 2015 IEEE International Conference on Big Data (IEEE Big Data 2015)*. Acceptance rate 18%.

**Hien To**, Gabriel Ghinita, Cyrus Shahabi. [A Framework for Protecting Worker Location Privacy in Spatial Crowdsourcing](#). *In proceedings of the 40th International Conference on Very Large Data Bases (VLDB 2014)*. Acceptance rate 20%.

**Hien To**, Kuorong Chiang, Cyrus Shahabi. [Entropy-based Histograms for Selectivity Estimation](#). *ACM International Conference on Information and Knowledge Management (CIKM 2013)*. Acceptance rate 17%.

Hung Dang, **Hien To** and Tuan Nguyen. [Maximum Complex Task Assignment: Enabling Crowdsourcing Spatial Complex Task](#). *The 15th International Conference on Information Integration and Web-based Applications & Services (iiWAS2013)*. Acceptance rate 30%.

### Workshop Articles

Sasan Tavakkol, **Hien To**, Seon Ho Kim, Patrick Lynett and Cyrus Shahabi. [An Entropy-based Framework for Efficient Post-disaster Assessment based on Crowdsourced Data](#). *The 2nd Workshop on Emergency Management using GIS (EM-GIS 2016)*. San Francisco, USA, October 2016.

Minh Nguyen, Sanjay Purushotham, **Hien To**, Cyrus Shahabi. [m-TSNE: A Framework for Visualizing High-Dimensional Multivariate Time Series](#). *Proceedings of the 2016 Workshop on Visual Analytics in Healthcare (VAHC 2016)*. November 12th, 2016, Chicago, IL.

**Hien To**, Ruben Geraldès, Cyrus Shahabi, Seon Ho Kim, Helmut Prendinger. [An Empirical Study of Workers' Behavior in Spatial Crowdsourcing](#). *Third International ACM SIGMOD Workshop on Managing and Mining Enriched Geo-Spatial Data (GeoRich 2016)*.

Dung Vu, **Hien To**, Cyrus Shahabi, Won-Yong Shin. [GeoSocialBound: An Efficient Framework for Estimating Social POI Boundaries Using Spatio-Textual Information](#). *Third International ACM SIGMOD Workshop on Managing and Mining Enriched Geo-Spatial Data (GeoRich 2016)*.

**Hien To**, Mohammad Asghari, Dingxiong Deng, and Cyrus Shahabi. [SCAWG: A Toolbox for Generating Synthetic Workload for Spatial Crowdsourcing](#). In *Proceeding of International Workshop on Benchmarks for Ubiquitous Crowdsourcing: Metrics, Methodologies, and Datasets (CROWDBENCH 2016)*.

**Hien To**, Liyue Fan, Cyrus Shahabi. [Differentially Private H-Tree](#). *2nd Workshop on Privacy in Geographic Information Collection and Analysis (GeoPrivacy 2015)*. In conjunction with ACM SIGSPATIAL 2015.

Afsin Akdogan, **Hien To**, Seon Ho Kim, and Cyrus Shahabi. [A Benchmark to Evaluate Mobile Video Upload to Cloud Infrastructures](#). In *Proceedings of the fifth workshop on Big data benchmarks, Performance optimization, and Emerging hardware (BPOE 2014)*.

### Demo Papers

**Hien To**, Gabriel Ghinita, Cyrus Shahabi. [PrivGeoCrowd: A Toolbox for Studying Private Spatial Crowdsourcing](#). In *Proceedings of 32st International Conference on Data Engineering (ICDE 2015)*.

### Symposium Papers

**Hien To**. [Task Assignment in Spatial Crowdsourcing: Challenges and Approaches](#). In *3rd ACM SIGSPATIAL 2016 PhD Symposium*. San Francisco, CA, USA, October 31 - November 3, 2016.

### Posters

**Hien To**, Hyerim Park, Seon Ho Kim, and Cyrus Shahabi. [Selecting User-Generated Videos for Augmented Reality Applications](#). *IMSC Retreat*. June 1, 2016.

**Hien To**, Seon Ho Kim, and Cyrus Shahabi. [Crowdsourcing the Acquisition and Analysis of Mobile Videos for Disaster Response](#). *IMSC Retreat*. June 1, 2016.

**Hien To**, Seon Ho Kim, and Cyrus Shahabi. [A Crowdsourcing Tool for Analyzing User-Generated Videos](#). *IMSC Retreat*. June 1, 2016.

**Hien To**, Seon Ho Kim, and Cyrus Shahabi. [A Benchmark to Evaluate Mobile Video Upload to Cloud Infrastructures](#). *IMSC Retreat*. April 2, 2015.

**Hien To**, Liyue Fan, Luan Tran, Cyrus Shahabi. [Real-time Spatial Task Assignment for Weather Crowdsourcing](#). *IMSC Retreat*. April 2, 2015.

**Hien To**, Giorgos Constantinou, Cyrus Shahabi. [A Generic Framework for Spatial Crowdsourcing](#). *Qualcomm Innovation Fellowship*. Mar 22, 2015.

**Hien To**, Phu Nguyen, Cyrus Shahabi, Soroosh Sorooshian. [iRain: A Spatial Crowdsourcing App for Real-time Rainfall Observation](#). *USC's Environmental Sustainability Research Network*. May 2014.

**Hien To**, Gabriel Ghinita, Cyrus Shahabi. [A Framework for Protecting Worker Location Privacy in Spatial Crowdsourcing](#). *IMSC Retreat*. March 2014.

**Hien To**, Jeong-min Lee, EungJin Lee, Seon Ho Kim. [iCampus Mobile](#). *IMSC Retreat*. April 2012.

**Hien To**, Kuorong Chiang, Cyrus Shahabi. [Entropy-based Histograms](#). *VEF Annual Conference*. January 2013.

## Patents

Kuorong Chiang, **Hien To**, and Wei Tang. [Methods and a system for hybrid large join query optimization](#). Patent No. 2016/0179890. *United States*, June 23, 2016.

Michael Gibas, **Hien To**. Histogram Generation on Multiple Dimensions. *United States*, Application Number 14/588,109. Filed January 2015.

## Media

Spatial Crowdsourcing research on [NSF Award](#) (September 1, 2013)

MediaQ on [NSF Award](#) (April 1, 2015), [NSF News](#) (2015) and [NSF Press Release](#) (March 30, 2015)

iRain on UNESCO: [iRain: new mobile App to promote citizen-science and support water management](#). November 08, 2016.

iRain on US Army Corps of Engineers: [iRAIN: New Mobile App Supports Water Management Around the World](#). November 28, 2016.

iRain on Los Angeles Times: [How much rain did we get? Ask the iRain app created at UCI](#). January 6, 2017.

iRain on UCI News: [UCI introduces iRain smartphone app](#). January 5, 2017.

MediaQ on Digital Trends: [The citizen Journalist: How Ordinary People are Taking Control of the News](#). June 19, 2015.

MediaQ on USC Viterbi News: [The Quest for Information](#). June 19, 2015.

MediaQ for [PBS Newshour Inauguration Coverage](#) (Jan 22, 2013) and [NATO Summit Coverage](#) (Dec 2, 2012).

## Professional Services

**Publicity Chair** of *Fourth International ACM SIGMOD Workshop on Managing and Mining Enriched Geo-Spatial Data*, May 2017.

**PC Member** of *The 4th IEEE PerCom International Workshop on Crowd Assisted Sensing, Pervasive Systems and Communications*, Mar 2017.

**PC Member** of *International Workshop on Benchmarks for Ubiquitous Crowdsourcing: Metrics, Methodologies, and Datasets (CROWDBENCH)*, Mar 2016.

**Reviewer** of *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, Jan 2017.

**Reviewer** of *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, May 2016.

**Reviewer** of *IEEE Transactions on Dependable and Secure Computing (TDSC)*, April 2016.

**Reviewer** of *Information Processing Letters (IPL)*, June 2015.

**External Reviewer** of *The 44th International Conference on Very Large Data Bases (VLDB)*, Oct 2016.

**External Reviewer** of *The 43rd International Conference on Very Large Data Bases (VLDB)*, Jun 2016.

**External Reviewer** of *The 42nd International Conference on Very Large Data Bases (VLDB)*, Mar 2016.

**External Reviewer** of *The 36th IEEE International Conference on Distributed Computing Systems (ICDCS)*, Dec 2015.

**External Reviewer** of *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, Nov 2015.

**External Reviewer** of *The 2015 ACM Special Interest Group on Management of Data (SIGMOD)*, Sep 2014.

**External Reviewer** of *The Journal of Systems and Software*, Dec 2014.

**External Reviewer** of *IEEE ICDE 31st Conference Data Engineering (ICDE)*, Sep 2014.

**External Reviewer** of *The 23rd ACM International Conference on Information and Knowledge Management (CIKM)*, June 2014.

**Student Volunteer** at *CIKM'13, DASFAA'15, IEEE Big Data'15 Conferences*.

## Teaching Experience

### University of Southern California

CSCI 587 - Geospatial Information Management *Fall 2014*

CSCI 101 - Fundamentals of Computer Programming *Fall 2013*

CSCI 587 - Geospatial Information Management *Fall 2012*

## Invited Talks

USC ACM Research Fair. Talk on *USC Infolab Research*. November 16, 2016.

Qualcomm Innovation Fellowship. Talk on *A Generic Framework for Spatial Crowdsourcing*. Mar 22 2015.

USC 6th Annual Graduate Research Symposium. Talk on *Protecting Worker Location Privacy in Spatial Crowdsourcing*. April 1, 2014.

USC for CSCI 587 Class – Geospatial Information Management. Guest lecture on *Privacy in Spatial Crowdsourcing*. Nov 12, 2014.



USC PhD Student Colloquium. Talk on *Entropy-based Histograms for Selectivity Estimation*. Oct 22, 2013.

## Awards

Student Travel Award to SIGSPATIAL 2016  
Student Travel Award to IEEE Big Data 2015  
Prof. Ram Kumar Travel Fellowships for ICDE 2015  
USC GSG Travel Grant for ICDE 2015  
Finalist, Qualcomm Innovation Fellowship 2015 (with George Constantinou)  
Microsoft Azure Research Grant 2014  
Travel Grant for CIKM 2013 Conference  
Vietnam Education Foundation Fellowship 2011  
Ministry of Education and Training Scholarship 2005-2010

## Technical Skills

Programming Languages: Python, Java, C/C++ STL, Android, iOS, PHP, HTML, CSS, JavaScript

Databases: Hadoop, Spark, PostgreSQL, Oracle and NoSQLs, e.g., Redis, Mongo, Neo4j

Data Tools: NumPy, Matplotlib, NLTK, Lucene, Solr, ElasticSearch, EC2, Azure, CodeIgniter

Science Tools: Scipy, Mahout, Word2Vec, Doc2Vec, t-SNE, LDA, LSI, k-means, DBSCAN, SVM, Logit

## Others

Lead a team of 3 students to developed a weather crowdsourcing app, named iRain (iOS & Android) *2014-Current*  
Work in a team of 10+ students to developed a MediaQ app, named MediaQ (iOS & Android) *2013-Current*  
Work in a team of 5 students to developed a bill sharing app, BillShare (iOS & Android) *May, 2014*  
Publish Android app (MiddleboxTester) to detect cellular networks' policies *Dec. 2014*  
Conquer Fansipan – the highest mountain in Indochina *Mar. 2011*  
Volunteer at Boys and Girls Clubs, El Segundo, CA *Summers 2012, 2014*  
Hobbies: Surfing, swimming, hiking, biking, travel to remote places, e.g., national parks

## References

Available upon request.