

Education

University of Southern California (USC)

Aug 2011 - Present

PhD Candidate, Computer Science. Expected to graduate in May 2017.

Contributed to various projects funded by NSF, Google, Microsoft, Oracle and Northrop Grumman in developing new **privacy-preserving** techniques for **efficient task assignment**, **task pricing** in **mobile crowdsourcing/crowdsensing**.

Hanoi University of Science and Technology (HUST)

Sep 2005 - Jul 2010

Awards

IEEE Big Data 2015 Student Travel Award

Oct 2015

Prof. Ram Kumar Travel Fellowships for ICDE'15 Conference

Apr 2015

USC Graduate Student Government Travel Grant for ICDE'15 Conference

Mar 2015

Finalist, Qualcomm Innovation Fellowship 2015

Mar 2015

Microsoft Azure Research Grant (\$40K worth of grant)

May 2014

Travel Grant for CIKM'13 Conference

Nov 2013

Vietnam Education Foundation Fellowship (\$52K in two years)

Aug 2011

Certificate of Merit for Excellent Graduation Grade

Jul 2010

Ministry of Education and Training Scholarship for Excellent Students

2005 - 2010

Refereed Journal Articles

Hien To, Liyue Fan, Gabriel Ghinita, Cyrus Shahabi. *Differentially Private Location Protection for Worker Datasets in Spatial Crowdsourcing*. 2016 IEEE Transactions on Mobile Computing. (under 2nd revision).

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

Refereed Conferences Articles

Ying Lu, **Hien To**, Abdullah Alfarrarjeh, Seon Ho Kim, Yifang Yin, Roger Zimmermann and CyrusShahabi. *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%.

Refereed Workshop Papers

Hien To, Ruben Gerald, 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. *SocialGeoFence: Detection of 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).

Refereed 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).

Posters

Afsin Akdogan, **Hien To**, Seon Ho Kim, 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.

Invited Talks

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

University of Southern California for CSCI 587: Geospatial Information Management. Guest lecture on *Privacy in Spatial Crowdsourcing*. Nov 12, 2014.

Professional Services

Technical Program Committee, CROWDBENCH 2016

Reviewer of IEEE Transactions on Dependable and Secure Computing, Information Processing Letters

External Reviewer of: VLDB 2016, ICDCS 2016, TKDE 2015, SIGMOD 2015, ICDE 2015, CIKM 2014, etc.

Student volunteer @ CIKM'13, DASFAA'15, IEEE Big Data'15 Conferences

Teaching Experience

Teaching Assistant, CSCI 587 - Geospatial Information Management, USC

Fall 2014

Teaching Assistant, CSCI 101 - Fundamentals of Computer Programming, USC

Fall 2013

Teaching Assistant, CSCI 587 - Geospatial Information Management, USC

Fall 2012

Research Experience

Spatial Crowdsourcing

USC

Aug 2011 - Present

Online task assignment in spatial crowdsourcing with emphasis on location privacy issues.

Data Grid

High Performance Computing Center @ HUST

May 2009 - Jun 2010

Developed datagrid-based interscholastic system for document plagiarism detection using LDA technique

Industry Experience

Multimedia Data Management @ IMSC

USC

Aug 2011 - Present

Developing a novel media management platform, namely MediaQ, for collecting, organizing, sharing, searching, and trading user-generated mobile pictures and videos using advanced geospatial and video analytics technologies. Link: <http://mediaq.usc.edu/>

Networking & Data Science Projects

USC

Aug 2011 - Present

Analyzed crowdsourced measurement data to identify proxy policies (e.g., middlebox existence, caching, delay handshaking, redirection and content rewriting) of large cellular networks, including 4 large U.S carriers.

Retrieved twitter's (trending) topics and conducted parallel k-means clustering analysis of the data using Apache Mahout on Amazon EMR.

Software Engineer Intern, Query Optimization

Teradata Corporation

Summer 2014

Developed genetic algorithms for large join query optimization in shared-nothing architecture. The proposed technique reduced the estimated cost a few times while the overhead is minimal.

Software Engineer Intern, Query Optimization

Teradata Corporation

Summer 2012

Developed one/multidimensional histograms to improve accuracy of database statistics. The two proposed histograms are the best for equality and range queries, respectively.

R&D Back End Engineer

VNG Corporation

Jun 2010 - Jul 2011

Developed distributed key-value storage that serves as a backend for 2 millions+ active game online users and 6 millions+ active social network users. The system assures high availability with the following features: persistent storage, memory caching, high scalability, and fault-tolerance.

Developed instant social search engine that serves 6 millions+ active users. The system enables users to search/explore their network, such as their friends, friends of friends, nearby friends, school friends, and friends with common interests.

Technical Skills

Programming Languages

Java, C/C++ STL, Python, Android, iOS, Ruby, PHP, JSP, HTML, CSS, JavaScript

Databases

Hadoop, Spark, PostgreSQL, MySQL, MSSQL, Oracle, Redis, Mongo, Neo4j, BerkeleyDB, DB2

Technologies

Distributed Systems, Web Services, NoSQL, Amazon EC2, Microsoft Azure, Lucene, Solr, ElasticSearch, Struts, Hibernate, IBM InfoSphere, Google-Facebook-Twitter APIs

Others

Lead a team of 3 students to developed a weather crowdsourcing app, **iRain** (iOS & Android)

2014-Current

Work in a team of 10 students to developed a MediaQ app, **MediaQ** (iOS & Android)

2013-Current

Work in a team of 5 students to developed a bill sharing app, **BillShare** (iOS & Android)

May. 2014

Publish **MiddleboxTester** (Android app) to detect cellular networks' policies

Dec. 2014

Hobbies: Surfing, swimming, hiking, biking, travel to remote places, e.g., national parks

Conquer Fansipan the highest mountain in Indochina

Mar. 2011

Teradata Cares @ Boys and Girls Clubs, El Segundo, CA