

Education

University of Southern California (USC)
3rd Ph.D in Computer Science.

Aug 2011 - Present

Hanoi University of Science and Technology (HUST)

Sep 2005 - Jul 2010

B.S in Computer Science. Excellent. Rank: 4/39

Thesis: GOODAS: DataGrid-Based Interscholastic System for Document Plagiarism Detection

Technical Skills

Operating Systems	Linux (Ubuntu, Fedora, RedHat), Windows, MacOS
Programming Languages	Java, C/C++ STL, Python, Ruby, Android, PHP, JSP, JavaScript, jQuery
Database Systems	MySQL, MSSQL, Teradata DB, Oracle, DB2, Redit, BerkeleyDB, CouchDB, Neo4j
Technologies	Distributed Systems, NoSQL, Amazon EC2, Lucene, Solr, ElasticSearch, Struts, Hibernate, IBM InfoSphere, Google-Facebook-Twitter APIs, Globus Toolkit

Honors and Awards

Travel grant & volunteer at CIKM'13 Conference

Nov 2013

VEF fellowship to pursue Ph.D in the U.S

Aug 2010

Certificate of Merit for excellent graduation grade

Jul 2010

Ministry of Education and Training Scholarship for excellent students

2005 - 2010

Publications

[1] **Hien To** et. al. *Differential Privacy Preserving in Spatial Crowdsourcing*. (ongoing)

[2] **Hien To**, Leyla Kazemi, Cyrus Shahabi. *GeoCrowd: A Server-Assigned Crowdsourcing Framework*, ACM Transactions on Spatial Algorithms and Systems (TSAS). (in submission)

[3] 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%)

[4] **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%)

[5] Le Duc Tung, Nguyen Viet Phuong, Nguyen Hong Thanh, **Hien To**. *Data grid solution for management of electronic documents stored in distributed virtual organizations*, 4th National Symposium of Fundamental and Applied Information Technology, (FAIR 2009). (in Vietnamese)

Industry Experience

Software Engineer Architect (part time), Mobivi

Sep 2012 - Sep 2013

NextGiving: a charitable social marketing platform to provide efficient way for charities organizations to receive donations from anyone via social network (30K active users in Jan 2013)

Software Engineer Intern, Teradata

Summer 2012

Selectivity Estimation: developed one/multidimensional histograms for Teradata DB Optimizer

R&D Engineer, VNG

Jun 2010 - Jul 2011

ZStore: developed distributed key-value storage from scratch that assures great availability with the following features: persistent storage, memory caching, high scalability, and fault-tolerance (1M+ active game users)

Feed Ranking (FR): developed backend that provides social network users (6M+ active users in Jan 2011) with personalized ranked feeds to overcome information overload

Social Faceted Search (SFS): developed instant search engine that helps users to search/explore their network, such as their friends, friends of friends, nearby friends, school friends, and friends with common interests

Software Engineering Intern, VMG

Feb 2008 - May 2009

Mobile Apps: developed song finder, news reader, and games on limited devices (CLDC)

Research Experience

Research Assistant, InfoLab, USC

Aug 2011 - Present

GeoCrowd: research on scalability, trust and privacy issues in spatial crowdsourcing

MediaQ: develop a novel online media management framework to collect, organize, share, search, and trade user-generated mobile images and public videos

Janus: developed a multisource event detection system for effective surveillance of criminal activity, in which data streams come from multiple sources, e.g. realtime videos, user-generated media, realtime text report

iCampusMobile: developed an Android app that provides location-based services for students and integrates multiple sources of data, including public data, user-generated data and USC's private data

Coursework, InfoLab, USC

Aug 2011 - Present

Parallel Clustering and Analysis of Twitter Data: retrieved twitter's (trending) topics and conducted parallel k-means clustering analysis of the data using Apache Mahout on Amazon EMR.

Research Intern, High Performance Computing Center, HUST

May 2009 - Jun 2010

Plagiarism Detection: developed grid-based document management system that addresses the key issues of security at multiple levels, high scalability, and mirroring of data for reliability

Teaching Experience

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

Fall 2013

Teaching Assistant, CSCI 587 - Geospatial Information Management, USC

Fall 2012