

[Home](#) | 
 [Projects](#) | 
 [People](#) | 
 [Research](#) | 
 [Publications](#) | 
 [Datasets](#) | 
 [News](#) | 
 [More Infolab!](#) | 
 [Related Links](#)

## Research

### [Traffic Prediction](#)

### [Human Performance Optimization](#)

### [Outlier detection in data streams](#)

### [Scenic trip planning on road networks](#)

### [Mobile Video Data Management](#)

### [Spatial Crowdsourcing](#)

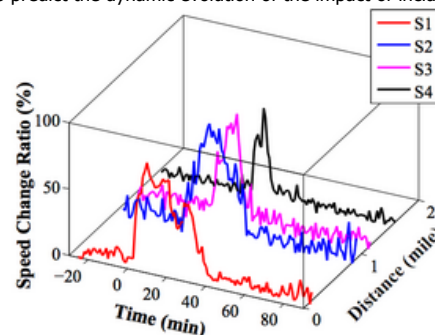
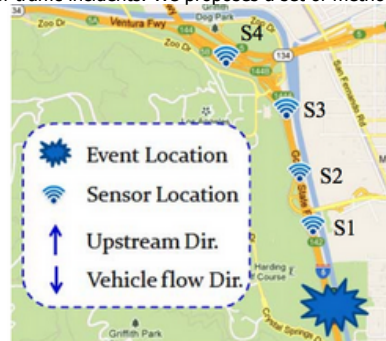
### [GeoSocial Networks](#)

### [Past Research](#)

- [WOLAP: Wavelet-based Online Analytical Processing](#)
- [Privacy in Location-Based Services](#)
- [Visibility Computation for Trajectories](#)
- [Image Acquisition Planning](#)
- [K-Nearest Neighbor Queries Over Moving Objects on Road Networks](#)
- [Surface KNN Queries](#)
- [Database Outsourcing](#)
- [Private Participatory Urban Sensing](#)
- [Road Network Traffic Modeling](#)
- [Location-based Web Search](#)
- [Subspace Method](#)
- [Texture Generation](#)
- [Nearest Neighbor Search](#)
- [Geospatial Data Fusion](#)
- [Spatial Data Stream](#)
- [Temporal Abstraction of Immersive Data Stream](#)
- [Multivariate Time Series Analysis](#)
- [Peer-to-Peer Search](#)
- [Web Services Discovery](#)
- [Query Customization](#)
- [Continuous Media Servers](#)
- [Shape Retrieval](#)

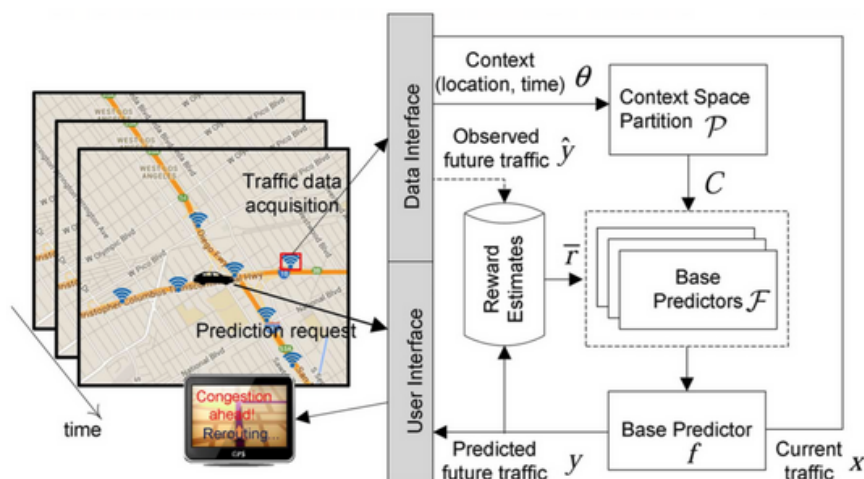
## Traffic Prediction

In this research effort, we focused on traffic prediction problem via utilizing the traffic sensor dataset. Due to thorough sensor instrumentations of road network in Los Angeles as well as the vast availability of auxiliary commodity sensors from which traffic information can be derived (e.g., CCTV cameras, GPS devices), a large volume of real-time and historical traffic data at very high spatial and temporal resolutions have become available. Therefore, how to mine valuable information from these data has become a key topic in spatiotemporal data mining. We have studied the traffic prediction problem for road segments. We utilized the spatiotemporal behaviors of rush hours and events to perform a more accurate prediction of both short-term and long-term average speed on road-segments, even in the presence of infrequent events (e.g., accidents). By utilizing both the topology of road network and sensor dataset, we are able to overcome the sparsity of our sensor dataset and extend the prediction task to the whole road network. We have also addressed the problems related to the impact of traffic incidents. We propose a set of methods to predict the dynamic evolution of the impact of incidents.



We then study the online traffic prediction problem. One key challenge in traffic prediction is how much to rely on

prediction models that are constructed using historical data in real-time traffic situations, which may differ from that of the historical data and change over time. To overcome this, we propose a novel online framework that could learn from the current traffic situation (context) in real-time and predict the future traffic by matching the current situation to the most effective prediction model trained using historical data. As real-time traffic arrives, the traffic context space is adaptively partitioned in order to efficiently estimate the effectiveness of each base predictor in different situation.



### People Involved

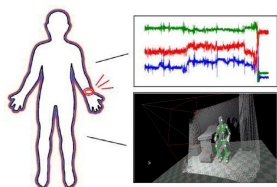
- Bei Pan (Penny), Dingxiong Deng

### Publications

- Jie Xu, Dingxiong Deng, Ugur Demiryurek, Cyrus Shahabi and Mihaela van der Schaar. Mining the Situation: Spatiotemporal Traffic Prediction with Big Data. Journal of Selected Topics Signal Processing, 2015
- Jie Xu, Dingxiong Deng, Ugur Demiryurek, Cyrus Shahabi, and Mihaela van der Schaar. [Context-Aware Online Spatiotemporal Traffic Prediction](#), ICDM Workshop on Spatial and Spatial-Temporal Data Mining (SSTD), 2014
- Bei Pan, Ugur Demiryurek, Chetan Gupta, and Cyrus Shahabi, [Forecasting Spatiotemporal Impact of Traffic Incidents on Road Networks](#), ICDM'13, Dallas, Texas, USA, Dec 7-10, 2013
- Bei Pan, Yu Zheng, David Wilkie, and Cyrus Shahabi, [Crowd Sensing of Traffic Anomalies based on Human Mobility and Social Media](#), ACM SIGSPATIAL GIS'13, Orlando, Florida, USA, November 5-8, 2013
- Bei Pan, Ugur Demiryurek, and Cyrus Shahabi, [Utilizing Real-World Transportation Data for Accurate Traffic Prediction](#), IEEE International Conference on Data Mining (ICDM), Brussels, Belgium, December 2012

[Top](#)

## Human Performance Optimization



Recent development of wearable technology and specialized camera has opened up great opportunities for human performance evaluation applications in various domains. In this line of research, we aim to analyze human motion data extracted from various sensors and cameras to determine human performance level. To monitor physical activities and measure the performance of an individual, wrist-worn sensors embedded in smartwatches, fitness bands, and clip-on devices can be used to collect various types of data, as the subject performs regular daily activities. Furthermore, specialized cameras are proposed to capture human body motion for further examination. The motivation of this research comes from the inquiry how to set the status of an object, especially in the field of cancer diagnosis and treatment. Data collected from accelerometers and skeleton-captured cameras are processed and analyzed using data mining and machine learning algorithms.

### People Involved

- Minh Nguyen

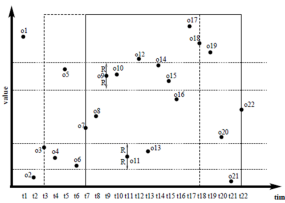
### Publications

- Minh Nguyen, Liyue Fan, and Cyrus Shahabi, [Activity Recognition Using Wrist-Worn Sensors for Human Performance Evaluation](#), The Sixth Workshop on Biological Data Mining and its Applications in Healthcare in conjunction with the 14th IEEE International Conference on Data Mining (ICDM 2015), Atlantic City, New Jersey, USA, November 14-17, 2015.

[Top](#)

## Outlier Detection In Streams of Data

The discovery of distance-based outliers from huge volumes of streaming data is critical for modern applications ranging from credit card fraud detection to moving object monitoring. In this research topic, we aim to find data points in windows that have fewer than  $k$  neighbors in distance  $R$ . The number of dimensions of data can be from 1 to thousands. Most algorithms utilize a data structure that support range query to find neighbors of data points.



But for high dimensional data, the time for processing is large. We use some transformations to reduce the dimensions of data and compute upper and lower bound distance in transformed metric to prune data points. We propose exact algorithms to detect distance-based outliers.

#### People Involved

- Luan Tran

#### Publications

[Top](#)

### Scenic Trip Planning on Road Networks



Traditional route planning problems mainly focus on finding the shortest path considering the travel distance or time. In this research topic, we aim to find the most scenic path that offers the most beautiful sceneries on the arcs of a path while the total travel cost (distance or time) is within a user-specified budget. This is a challenging problem as the optimization objective is to maximize the value of the path (i.e., its scenic value) instead of minimizing its cost (distance or time). The problem can be formulated as a variant of the Arc Orienteering Problem (AOP), which is a well-known NP-hard combinatorial optimization problem. Due to the fast response-time requirements of interactive mobile and online applications (e.g., within 300 milliseconds) and the large scale of real-world road networks, existing heuristic algorithms for AOP in operation research community fail to solve the most scenic road problem. To this end, we utilize the techniques from the field of spatial database, e.g., spatial indexing and pruning, to find the scenic path on a large-scale road network with high efficiency and accuracy.

#### People Involved

- Ying Lu

#### Publications

[Top](#)

### Mobile Video Data Management



An unprecedented number of user-generated videos (UGVs) are currently being collected by mobile devices (e.g., smartphones, Google Glass), however, such unstructured data are very hard to index and search. Due to the advanced sensors technologies of the mobile devices, UGVs can be geo-tagged, e.g., GPS locations and compass directions, at the acquisition time at a very fine spatial granularity. Ideally, each video frame can be tagged by the spatial extent of its coverage area, termed Field-Of-View (FOV). Consequently, the problems of mobile video indexing and querying are converted to the problems of FOV indexing and querying on FOVs. However, the challenges of indexing and querying FOVs are: 1) an FOV is shaped similar to a piece of pie and contains both location and orientation information, conventional spatial indexes, such as R-trees, cannot index them efficiently; 2) since UGVs are usually captured in a casual way with various shooting and moving directions, moving trajectories, zoom levels and camera lenses, no pattern can be made to condense FOVs in an index structure. To overcome these two challenges, we proposed a class of new indexes called Orientation-R-trees (OR-trees) combining the orientation and viewable distance of the FOVs into the indexes for filtering and optimization. Moreover, UGVs are in large-scale and are updated frequently. According to the recent YouTube statistics, there are more than 3 hours mobile videos uploading to YouTube. These additional challenges open new research directions for FOV indexing and query processing.

#### People Involved

- Ying Lu, Abdullah Alfarrarjeh, Hien To

#### Publications

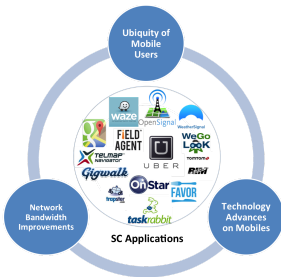
- Ying Lu, Hien To, Abdullah Alfarrarjeh, Seon Ho Kim, Yifang Yin, Roger Zimmermann, and Cyrus Shahabi, [GeoUGV: User-Generated Mobile Video Dataset with Fine Granularity Spatial Metadata](#), In the 7th ACM Multimedia Systems Conference (MMSys), Klagenfurt am Worthersee, Austria, May 10-13, 2016
- Hien To, Hyerim Park, Seon Ho Kim, and Cyrus Shahabi, [Incorporating Geo-Tagged Mobile Videos Into Context-Aware Augmented Reality Applications](#), [The Second IEEE International Conference on Multimedia Big Data \(IEEE BigMM 2016\)](#), DOI: 10.1109/BigMM.2016.64, Taipei, Taiwan, April 20-22, 2016
- Hien To, Seon Ho Kim, and Cyrus Shahabi, [Effectively Crowdsourcing the Acquisition and Analysis of Visual Data for Disaster Response](#), In proceeding of 2015 IEEE International Conference on Big Data (IEEE Big Data 2015), Pages 697-706, (Acceptance rate ~ 18%), Santa Clara, CA, USA, October 29-November 1, 2015
- Yinghao Cai, Ying Lu, Seon Ho Kim, Luciano Nocera, and Cyrus Shahabi, [GIFT: A Geospatial Image and Video Filtering Tool for Computer Vision Applications with Geo-tagged Mobile Videos](#), IEEE International Workshop on Mobile Multimedia Computing, pages: 1 - 6 (Best Paper Award), Torino, Italy, June 29 - July 3, 2015, (Best Paper Award, one out of 57 submissions)
- Ying Lu, Cyrus Shahabi, and Seon Ho Kim, [An Efficient Index Structure for Large-scale Geo-tagged Video Databases](#), 23rd ACM SIGSPATIAL International Conference on Advances in GIS, pages: 465 - 468, Dallas, Texas, USA, 2014 (short paper, acceptance rate: 22.8%)

- Guanfeng Wang, Ying Lu, Luming Zhang, Abdullah Alfarrarjeh, Roger Zimmermann, Seon Ho Kim, and Cyrus Shahabi, [Active Key Frame Selection for 3D Model Reconstruction From Crowdsourced Geo-tagged Videos](#), In the Proceedings of International Conference of Multimedia & Exp (ICME), pages: 1-6, Chengdu, China, July 14 - 18, 2014
- Seon Ho Kim, Ying Lu, Giorgos Constantinou, Cyrus Shahabi, Guanfeng Wang, and Roger Zimmermann, [MediaQ: Mobile Multimedia Management System. In the conference of Multimedia Systems \(MMSys\)](#), pages: 224 - 235, Singapore, March 19 -21, 2014, (acceptance rate: 25%)
- Seon Ho Kim, Ying Lu, Junyuan Shi, Abdullah Alfarrarjeh, Cyrus Shahabi, Guanfeng Wang, and Roger Zimmermann, Frame Selection Algorithms for Automatic Generation of Panoramic Images from Crowdsourced Geo-tagged Videos, In the conference of Web and Wireless Geographic Information System (W2GIS), Seoul, South Korea, 2014

[Top](#)

## Spatial Crowdsourcing

### Spatial Crowdsourcing



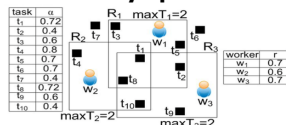
We are experiencing an astonishing growth in the number of smartphone users, the phone's hardware, and the broadband bandwidth. Exploiting this large volume of potential users and their movability, a transformative area of research is to utilize this new platform for various tasks, among which the most promising is spatial crowdsourcing. Spatial crowdsourcing (SC) engages individuals, groups, and communities in the act of collecting, analyzing, and disseminating high-fidelity urban, social, and other spatiotemporal information (e.g., picture, video, audio and their metadata such as location, time, speed, direction, and acceleration). Spatial crowdsourcing requires the workers to perform a set of tasks by physically traveling to certain locations at particular times. This new paradigm for data collection has various applications in environmental sensing, disaster response, and transportation decision-making. Our studies aim to focus on addressing three major challenges with SC in real-world applications, scalability, trust and privacy issues. First, to continuously match thousands of spatial crowdsourcing campaigns, where each campaign consists of many spatiotemporal tasks, with millions of workers, a SC-Server must be able to run efficient task-assignment strategies that can scale. Second, while in small campaigns, the workers may be known and trusted, with spatial crowdsourcing, the workers cannot always be trusted. Third, individuals with mobile devices need to physically travel to the specified locations of interest. An adversary with access to these individuals' whereabouts can infer sensitive details about a person (e.g., health status and political views); thus, protecting worker location privacy is an important concern in spatial crowdsourcing. Finally, we developed system prototypes of spatial crowdsourcing, including MediaQ and iRain, served as testbeds for research in video data/metadata management, indexing and retrieval of such data and crowd sensing, respectively.

### Scalable Spatial Crowdsourcing



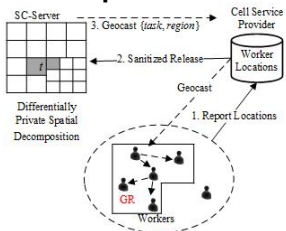
Given that the task assignment is the main bottleneck of the system, the spatial and temporal aspects of the tasks are exploited to reduce the complexity of assignment. We proposed the Maximum Task Assignment (MTA) problem for task assignment from the workers' perspective. Specifically, given a worker and a set of spatial tasks, MTA aims to maximize the number of tasks the worker could complete considering the physical location and the expiration time of each task. Next, we studied task assignment at scale from the server's perspective. In particular, we extended our server-assigned SC approach (first proposed in ACMGIS'12) from a single centralized system to a distributed environment.

### Trustworthy Spatial Crowdsourcing



We tackle the issue of trust by having tasks performed redundantly by multiple workers. We would then require maintaining a reputation scheme for workers and requesters and then utilize the aggregate reputation of the workers to see if the pending tasks would meet the confidence levels set by their requesters. Thus, the problem is to maximize the number of spatial tasks that are assigned to a set of workers while satisfying the confidence levels of those tasks. The "spatial" aspect of problem to make it unique is that every task should be assigned to enough number of workers such that their aggregated reputation satisfies the confidence of the task.

### Private Spatial Crowdsourcing



In this study, our goal is to develop an efficient framework for worker-selected spatial crowdsourcing which enable the participation of the crowdsourcing workers without compromising their privacy. There have been several solutions proposed to protect location-based queries, i.e., given an individual's location find point of interests in proximity without disclosing the actual coordinates. However, we argue that existing location privacy techniques are not sufficient for SC (i.e., a worker's location is no longer part of the query but rather the result of a spatial query around the task), and we propose a mechanism based on differential privacy and geocasting that achieves effective SC services while offering privacy guarantees to workers. Toward this end, we investigate analytical models and task assignment strategies that balance multiple crucial aspects of SC functionality, including task completion rate, worker travel distance and system overhead. We focus on both releasing one time snapshot of the location data and protecting worker location across multiple timestamps.

### System Prototypes

We develop [MediaQ](#), a novel online media management system to collect, organize, share, and search mobile multimedia contents using automatically tagged geospatial metadata. User-generated videos can be uploaded to the MediaQ from users' smartphones, iPhone and Android, and displayed accurately on a map interface according to their automatically sensed geospatial and other metadata. A basic spatial crowdsourcing mechanism has been implemented in MediaQ for on-demand user-generated video collection. In collaboration with environmental scientists at UC Irvine, we utilize spatial crowdsourcing technology to enable human workers to report precipitation condition, particularly rain/norain observation to improve realtime global satellite precipitation estimation (i.e. [G-WADI](#) system). We develop a spatial crowdsourcing system namely iRain, which consists of a server and mobile apps for smartphones and tablets using [iPhone](#) or [Android](#). SC tasks (e.g., collecting rain information at a particular place during time period) can be posted to iRain from users' phones or via a [web portal](#). Thereafter the

tasks will be crowdsourced among available nearby workers. Workers report 'rain' with spatiotemporal information to the server when they see rain occurring around their locations.

### People Involved

- Leyla Kazemi, Hien To, Mohammad Asghari, Dingxiong Deng, Ying Lu, Abdullah Alfarrarjeh, Giorgos Constantinou

### Publications

- Hien To, [Task Assignment in Spatial Crowdsourcing: Challenges and Approaches](#), [The 3rd ACM SIGSPATIAL 2016 PhD Symposium](#), San Francisco, CA, USA, October 31 - November 3, 2016
- Hien To\*, Kien Nguyen\*, and Cyrus Shahabi, [Differentially Private Publication of Location Entropy](#), [In Proceeding of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems \(SIGSPATIAL 2016\)](#), (Acceptance rate ~18%), San Francisco, CA, USA, October 31 - November 3, 2016
- Hien To, Gabriel Ghinita, Liyue Fan, and Cyrus Shahabi, [Differentially Private Location Protection for Worker Datasets in Spatial Crowdsourcing](#), [IEEE Transactions on Mobile Computing \(TMC 2016\)](#), DOI <http://dx.doi.org/10.1109/TMC.2016.2586058>, June 29, 2016
- Hien To, Ruben Gerald, Cyrus Shahabi, Seon Ho Kim, and Helmut Prendinger, [An Empirical Study of Workers' Behavior in Spatial Crowdsourcing](#), [Third International ACM SIGMOD Workshop on Managing and Mining Enriched Geo-Spatial Data \(GeoRich 2016\)](#), DOI:<http://dx.doi.org/10.1145/2948649.2948657>, San Francisco, CA, USA, June 26 - July 1, 2016
- Dingxiong Deng, Cyrus Shahabi, Ugur Demiryurek, and Linhong Zhu, [Task Selection in Spatial Crowdsourcing from Worker's Perspective](#), [GeoInformatica](#), DOI 10.1007/s10707-016-0251-4, March 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 \(CROWDBENCH 2016\)](#), Sydney, Australia, March 14-18, 2016
- Hien To, Liyue Fan, Luan Tran, and Cyrus Shahabi, [Real-Time Task Assignment in Hyperlocal Spatial Crowdsourcing under Budget Constraints](#), [In Proceeding of IEEE International Conference on Pervasive Computing and Communications \(PerCom 2016\)](#), Sydney, Australia, March 14-18, 2016 (Acceptance rate <15%)
- Hien To, Seon Ho Kim, and Cyrus Shahabi, [Effectively Crowdsourcing the Acquisition and Analysis of Visual Data for Disaster Response](#), [In proceeding of 2015 IEEE International Conference on Big Data \(IEEE Big Data 2015\)](#), Pages 697-706, (Acceptance rate ~ 18%), Santa Clara, CA, USA, October 29-November 1, 2015
- Hien To, Liyue Fan, and Cyrus Shahabi, [Differentially Private H-Tree](#), [In proceeding of 2nd Workshop on Privacy in Geographic Information Collection and Analysis \(GeoPrivacy 2015\)](#). In conjunction with ACM SIGSPATIAL, Seattle, Washington, USA, November 3-6, 2015
- Dingxiong Deng, Cyrus Shahabi, and Linhong Zhu, [Task Matching and Scheduling for Multiple Workers in Spatial Crowdsourcing](#), [23rd ACM SIGSPATIAL International Conference on Advances in GIS](#), Seattle, Washington, USA, November 3-6, 2015, (acceptance rate: 17.5%)
- Abdullah Alfarrarjeh, Tobias Emrich, and Cyrus Shahabi, [Scalable Spatial Crowdsourcing: A study of distributed algorithms](#), [16th IEEE International Conference on Mobile Data Management \(MDM\)](#), Pittsburgh, Pennsylvania, USA, June 15-18, 2015, (acceptance rate: <25%)
- Hien To, Leyla Kazemi, and Cyrus Shahabi, [A Server-Assigned Spatial Crowdsourcing Framework](#), [Journal ACM Transactions on Spatial Algorithms and Systems \(TSAS\)](#), Volume 1 Issue 1, Article No. 2. DOI: <http://doi.acm.org/10.1145/2729713> (Acceptance Rate ~ 11%), August 2015
- Hien To, Gabriel Ghinita, and Cyrus Shahabi, [PrivGeoCrowd: A Toolbox for Studying Private Spatial Crowdsourcing](#), [2015 IEEE 31st International Conference on Data Engineering \(ICDE 2015\)](#), Page 1404 - 1407, DOI: <http://dx.doi.org/10.1109/ICDE.2015.7113387>, (demonstration), Seoul, Korea, April 13-17, 2015
- Afsin Akdogan, Hien To, Seon Ho Kim, and Cyrus Shahabi, [A Benchmark to Evaluate Mobile Video Upload to Cloud Infrastructures](#), [BPOE workshop at VLDB](#), Volume 8807, 2014, pp 57-70. DOI: [http://dx.doi.org/10.1007/978-3-319-13021-7\\_5](http://dx.doi.org/10.1007/978-3-319-13021-7_5), Hangzhou, China, September 2014
- Hien To, Gabriel Ghinita, and Cyrus Shahabi, [A Framework for Protecting Worker Location Privacy in Spatial Crowdsourcing](#), [In Proceedings of the 40th International Conference on Very Large Data Bases \(VLDB 2014\)](#), Volume 7, No 10, Pages 919-930. (Acceptance rate 20%), Hangzhou, China, September 2014
- Hung Dang, Tuan Nguyen, and Hien To, [Maximum Complex Task Assignment: Towards Tasks Correlation in Spatial Crowdsourcing](#), [In Proceedings of International Conference on Information Integration and Web-based Applications and Services \(iiWAS 2013\)](#), No 77, Pages 77-81, DOI: <http://doi.acm.org/10.1145/2539150.2539243>, Vienna, Austria, 2-4 December 2013
- Leyla Kazemi, Cyrus Shahabi, and Lei Chen, [GeoTruCrowd: Trustworthy Query Answering with Spatial Crowdsourcing](#), [International Conference on Advances in Geographic Information Systems \(ACM SIGSPATIAL GIS 2013\)](#), Orlando, Florida, November 5-8, 2013



- Dingxiong Deng, Cyrus Shahabi, and Ugur Demiryurek, [Maximizing the Number of Worker's Self-Selected Tasks in Spatial Crowdsourcing](#), [ACM SIGSPATIAL GIS '13](#), Orlando, Florida, USA, November 5-8, 2013
- Leyla Kazemi and Cyrus Shahabi, [GeoCrowd: Enabling Query Answering with Spatial Crowdsourcing](#), [ACM SIGSPATIAL GIS](#), Redondo Beach, CA, November 2012
- Ali Khodaei and Cyrus Shahabi, [Social-Textual Search and Ranking](#), [CrowdSearch workshop at WWW 2012](#), Lyon, France, April 2012

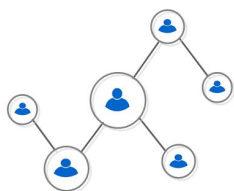
[Top](#)

## GeoSocial Networks

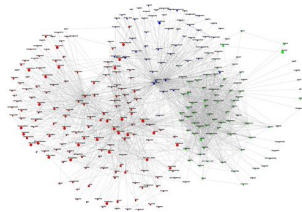
### Inferring Real-World Social Strength and Influence from Spatiotemporal Data



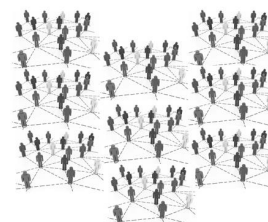
Social networks have been studied by social scientists since the pre-Internet era. Beginning in the 1970s, social studies were confined to extremely small social graphs collected from surveys. In the 1990s, the Internet revolution started to fuel online services, and resulted in giant online social networks in the 2000s – millions of users and their social connections (Facebook, Twitter, LinkedIn), which, in turn, fueled social studies and brought them to a new extent. Social scientists started to look at quantitative social connections and social influence by utilizing the huge online social graphs. However, these social relationships are only specific to online social media (aka the virtual world), and may not necessarily correspond to those in the real lives (aka the real world).



During 1970s, 1980s  
Surveys



Started in 1990s, popular in mid/late 2000s  
Virtual connections



Started in 2008/2009  
Real-world connections

In this line of research, we aim to bridge the gap between the virtual world and the real world by studying social behaviors that take place in the real world. The motivation of this research comes from the pervasiveness of GPS-enabled mobile devices, because of which the giant social networks have all gone mobile and produced massive data that represents the movements of people in the real world – aka spatiotemporal data. Such collections of spatiotemporal data constitute a rich source of information for studying and inferring various social behaviors, including **social connections** and **spatial influence**. For example, for social connections, the intuition is that if two people have been to the same places at the same time (aka *co-occurrences*), there is a good chance that they are socially related. Another example is in inferring spatial influence, whose intuition is based on the concept of *followship*, which captures the phenomenon of an individual visiting a real-world location (e.g., restaurant) due to the influence of another individual who has visited that same location in the past. We infer and quantify such social connections and influence by considering various geo-social aspects, including the impact of the frequency of co-occurrences/followship, the popularity of locations and the inherent coincidences in individuals' geospatial behaviors.

#### People Involved

- Huy Pham

#### Publications

- Cyrus Shahabi and Huy Pham, [Inferring Real-World Relationships from Spatiotemporal Data](#), Data Engineering Bulletin Issues, IEEE Computer Society, June 2015
- Nikos Armenatzoglou, Huy Pham, Vasilis Ntranos, Dimitris Papadias, and Cyrus Shahabi, [Real-Time Multi-Criteria Social Graph Partitioning: A Game Theoretic Approach](#), [SIGMOD](#), Melbourne, AU, June 2015
- Huy Pham, Cyrus Shahabi, and Yan Liu, [EBM - An Entropy-Based Model to Infer Social Strength from Spatiotemporal Data](#), [SIGMOD](#), New York, USA, June 2013
- Huy Pham, Ling Hu, and Cyrus Shahabi, [GEOSO - A Geo-Social Model: From Real-World Co-occurrences to Social Connections](#), [DNIS 2011](#), University of Aizu, Japan, December 2011
- Huy Pham, Ling Hu, and Cyrus Shahabi, [Towards Integrating Real World Spatiotemporal Data with Social Networks](#), [ACM SIG Spatial 2011](#), Chicago, Illinois, November 2011

[Top](#)

## Multidimensional Data Analysis

### WOLAP: Wavelet-based Online Analytical Processing

OLAP is an elegant approach to efficiently support analytical queries on massive multidimensional datasets. Several fundamental classes of OLAP queries, such as aggregate queries, slice-and-dice queries, or roll-up queries, have



been addressed by numerous researchers for the last couple of years. Most of these methods share the disadvantage of either providing only approximate answers by compressing the data or sacrificing the update cost for better query processing performance. However, we propose to employ wavelet transform to efficiently process such queries with exact results and without significant increases on update performance. Leveraging from multi-resolution property of wavelets, we also incorporate approximate query processing in case of space or time limitation as our approach is fundamentally progressive. Furthermore, we prepare and maintain wavelet data by providing a framework to efficiently create, insert, and update data. By developing a real system with its deployment in practice, we address all the essential steps toward realization of our proposed work. We are currently extending our framework by supporting a new class of analytical queries, plot queries, as the most frequently performed queries in scientific applications for extraction of useful information.

### People Involved

- Seyyed Jalal Kazemitabar

### Publications

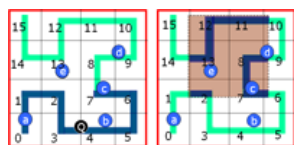
- Cyrus Shahabi and Rolfe Schmidt, [Methods for Fast Progressive Evaluation of Polynomial Range-Sum Queries on Real-time Datacubes](#), Utility Patent No. 8,090,730, United States, January 2012
- Ugur Demiryurek, Farnoush Banaei-Kashani, Cyrus Shahabi, and Anand Ranganathan, [Online Computation of Fastest Path in Time-Dependent Spatial Networks](#), [12th International Symposium on Spatial and Temporal Databases \(SSTD11\)](#), Minneapolis, MN, USA, August 2011
- Hyunjin Yoon and Cyrus Shahabi, [Robust Time-Referenced Segmentation of Moving Object Trajectories](#), [Eighth IEEE International Conference on Data Mining \(ICDM\)](#), Pisa, Italy, December 2008
- Ali Khodaei, Hyokyeong Lee, Farnoush Banaei-Kashani, Cyrus Shahabi, and Iraj Ershaghi, [A Mutual Information-Based Metric for Identification of Nonlinear Injector](#), [SPE Western Regional](#), San Jose, CA, March 2009
- Mehrdad Jahangiri and Cyrus Shahabi, [Plot Query Processing with Wavelets](#), [20th International Conference on Scientific and Statistical Database Management \(SSDBM\)](#), Hong Kong, China, July 2008
- Cyrus Shahabi, Mehrdad Jahangiri, and Farnoush Banaei-Kashani, [ProDA: An End-to-End Wavelet-Based OLAP System for Massive Datasets](#), [IEEE Computer Magazine](#), Vol. 41, No. 4, Pages 69-77, ISSN: 0018-9162, April 2008
- Ali Khoshgozaran, Ali Khodaei, Mehdi Sharifzadeh, and Cyrus Shahabi, [A hybrid aggregation and compression technique for road network databases](#), [Knowledge and Information Systems](#), ISSN 0219-1377 (Print) 0219-3116 (Online),
- Ali Khoshgozaran, Ali Khodaei, Mehdi Sharifzadeh, and Cyrus Shahabi, [A Multi-Resolution Compression Scheme for Efficient Window Queries over Road Network Databases](#), [Spatial and Spatio-temporal Data Mining \(SSTDm\)](#), Hong Kong, December 2006
- Mehrdad Jahangiri and Cyrus Shahabi, [Essentials for Modern Data Analysis Systems](#), [Second NASA Data Mining Workshop](#), Pasadena, CA, May 2006
- Cyrus Shahabi, [Wavelets and Web-Services \(W2\) for Online Scientific Data Analysis](#), Lawrence Livermore National Laboratory, Livermore, CA, August 2005
- Mehdi Sharifzadeh, Farnaz Azmoodeh, and Cyrus Shahabi, [Change Detection in Time Series Data Using Wavelet Footprints](#), [9th International Symposium on Spatial and Temporal Databases \(SSTD05\)](#), Angra dos Reis, Brazil, August 2005
- Mehrdad Jahangiri, Dimitris Sacharidis, and Cyrus Shahabi, [SHIFTSPLOT: I/O Efficient Maintenance of Wavelet-Transformed Multidimensional Data](#), [24th ACM SIGMOD International Conference on Management of Data](#), Baltimore, Maryland, USA, June 2005
- Mehrdad Jahangiri and Cyrus Shahabi, [ProDA: A Suite of WebServices for Progressive Data Analysis](#), [24th ACM SIGMOD International Conference on Management of Data](#), (demonstration), Baltimore, Maryland, USA, June 2005
- Cyrus Shahabi, Mehrdad Jahangiri, and Dimitris Sacharidis, [Hybrid Query and Data Ordering for Fast and Progressive Range-Aggregate Query Answering](#), [International Journal of Data Warehousing and Mining](#), Vol. 1, No. 2, Pages 49-69, Idea Group Publishing, ISSN: 1548-3924, April 2005
- Cyrus Shahabi, Dimitris Sacharidis, and Mehrdad Jahangiri, [Wavelets for Querying Multidimensional Datasets](#), [Encyclopedia of Data Warehousing and Mining](#), ISBN 1-59140-557-2, Vol. 1, Idea Group Reference, April 2005
- Cyrus Shahabi, Xiaoming Tian, and Wugang Zhao, [TSA-tree: A Wavelet-Based Approach to Improve the Efficiency of Multi-Level Surprise and Trend Queries on Time-Series Data](#), [The 12th International Conference on Scientific and Statistical Database Management \(SSDBM 2000\)](#), Berlin, Germany, July, 2000
- Rolfe R. Schmidt and Cyrus Shahabi, [ProPolyne: A Fast Wavelet-based Algorithm for Progressive Evaluation of Polynomial Range-Sum Queries \(extended version\)](#), [VIII. Conference on Extending Database Technology](#), Prague, March, 2002
- Cyrus Shahabi, Seokkyung Chung, and Maytham Safar, [A Wavelet-Based Approach to Improve the Efficiency of Multi-Level Surprise Mining](#), [PAKDD International Workshop on Mining Spatial and Temporal data](#), Hong Kong, April, 2001

- Cyrus Shahabi, Seokkyung Chung, Maytham Safar, and George Hajj, [2D TSA-tree: A Wavelet-Based Approach to Improve the Efficiency of Multi-Level Spatial Data Mining](#), [The 13th International Conference on Scientific and Statistical Database Management \(SSDBM 2001\)](#), Fairfax, Virginia, July, 2001
- Jose Luis ambite, Cyrus Shahabi, Rolfe R. Schmidt, and Andrew Philpot, [Fast Approximate Evaluation of OLAP Queries for Integrated Statistical Data](#), [National Conference for Digital Government Research](#), Los Angeles, California, May, 2001
- Rolfe R. Schmidt and Cyrus Shahab, [Wavelet Based Density Estimators for Modeling Multidimensional Data Sets](#), [SIAM Workshop on Mining Scientific Data Sets](#), Chicago, IL, April, 2001
- Rolfe R. Schmidt and Cyrus Shahabi, [How to Evaluate Multiple Range-Sum Queries Progressively](#), [21st ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems \(PODS\)](#), Madison, Wisconsin, June, 2002

[Top](#)

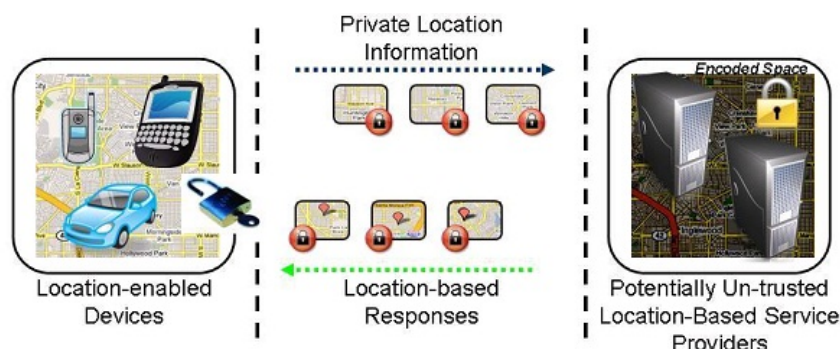
## Geospatial Information Management

### Privacy in Location-Based Services

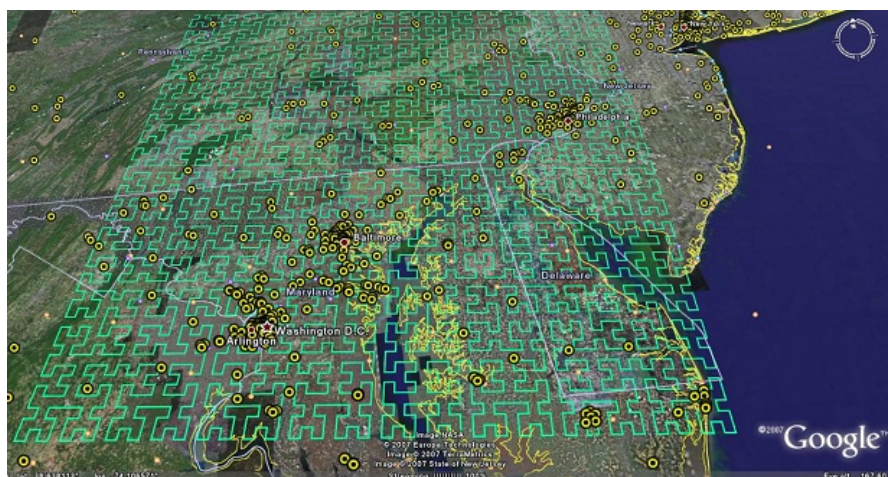


The past few years have witnessed a dramatic growth in capabilities of cellphone devices as well as the same growth in public demand for using such services and features. Such growth and trend in technology is often not seamless. Recent concerns over how such services can jeopardize user's private information have coined a term, until recently unknown, location privacy. With the advent of GPS devices embedded in cellphones, and their growing use, several breaching of subscriber's privacy by stalking their locations have been reported and many researchers and organizations have raised the need to explore the threats associated with location-based services and misuse of users' private location information.

The concerns over protecting user's location while using location-based services have led to the discussion that the user has to compromise his privacy for the service because more accurate responses from location servers demand more accurate information about where the user is located. However, we believe it is possible to allow users to enjoy the same quality of service while not being worried about their private location information.



In April 2007 we embarked on the new research topic of location privacy. We are studying different approaches to satisfy significantly more stringent privacy guarantees as compared to the first generation approaches based on location cloaking or anonymity. Our approaches are inspired by the work in the area of encryption and Private Information Retrieval (PIR). Both of these research fields have been around for a long time enabling privacy protection but not yet fully exploited for location privacy. You may refer to the research topics section for more information on these approaches.

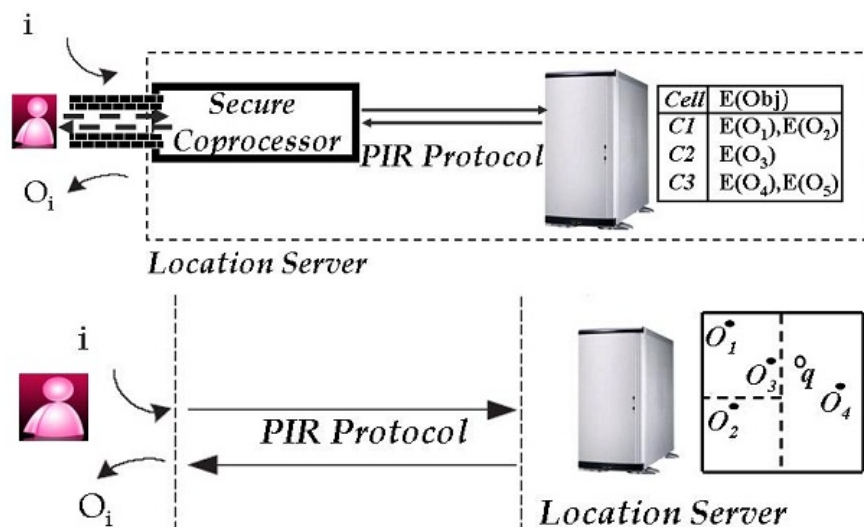


Space Encoding of Objects for Location Privacy

Hardware-Based PIR Schemes for Location Privacy

Computational-PIR-Based Schemes for Location Privacy





### People Involved

- Jeff Khoshgozaran

### Publications

- Ying Lu, Jiaheng Lu, Gao Cong, Wei Wu, and Cyrus Shahabi, [Efficient Algorithms and Cost Models for Reverse Spatial-Keyword k-Nearest Neighbor Search](#), In ACM Transactions On Database Systems (TODS) 39 (2): 13, 2014
- Seon Ho Kim, Ying Lu, Giorgos Constantinou, Cyrus Shahabi, Guanfeng Wang, and Roger Zimmermann, [MediaQ: Mobile Multimedia Management System. In the conference of Multimedia Systems \(MMSys\)](#), In the conference of Multimedia Systems (MMSys), pages: 224 - 235, Singapore, March 19 -21, 2014, (acceptance rate: 25%)
- Leyla Kazemi and Cyrus Shahabi, [TAPAS: Trustworthy privacy-aware participatory sensing](#), Knowledge and Information Systems, DOI: 10.1007/s10115-012-0573-y, 2012
- Leyla Kazemi and Cyrus Shahabi, [GeoCrowd: Enabling Query Answering with Spatial Crowdsourcing](#), ACM SIGSPATIAL GIS, Redondo Beach, CA, November 2012
- Ugur Demiryurek and Cyrus Shahabi, [Indexing Network Voronoi Diagrams](#), The 17th International Conference on Database Systems for Advanced Applications, Busan, South Korea , April 2012
- Ali Khoshgozaran and Cyrus Shahabi, [Towards Private Navigation of Tree Structured Spatial Indexes](#), The Third International Conference on Emerging Databases (EDB 2011), Incheon, Korea, August 2011
- Huy Pham, Ling Hu, and Cyrus Shahabi, [GEOSO - A Geo-Social Model: From Real-World Co-occurrences to Social Connections](#), DNIS 2011, University of Aizu, Japan, December 2011
- Huy Pham, Ling Hu, and Cyrus Shahabi, [Towards Integrating Real World Spatiotemporal Data with Social Networks](#), ACM SIG Spatial 2011, Chicago, Illinois, November 2011
- Ali Khodaei, Cyrus Shahabi, and Chen Li, [SKIF-P: a point-based indexing and ranking of web documents for spatial-keyword search](#), Geoinformatica, Publisher: Springer Netherlands, Issn: 1384-6175, Doi: 10.1007/s10707-011-0142-7, October 2011
- Leyla Kazemi and Cyrus Shahabi, [Towards Preserving Privacy in Participatory Sensing](#), PerCom, (short paper), Seattle, WA, March 2011
- Ugur Demiryurek, Farnoush Banaei-Kashani, and Cyrus Shahabi, [A Case for Time-Dependent Shortest Path Computation in Spatial Networks](#), ACM SIGSPATIAL, San Jose, CA, November 2010
- Bei Pan, Ugur Demiryurek, Farnoush Banaei-Kashani, and Cyrus Shahabi, [Spatiotemporal Summarization of Traffic Data Streams](#), ACM SIGSPATIAL, IWGS, San Jose, CA, November 2010
- Songhua Xing and Cyrus Shahabi, [Scalable Shortest Paths Browsing on Land Surface](#), ACM GIS, San Jose, CA, November 2010
- Ugur Demiryurek, Farnoush Banaei-Kashani, and Cyrus Shahabi, [Efficient K-Nearest Neighbor Search in Time-Dependent Spatial Networks](#), 21st International Conference on Database and Expert Systems Applications (DEXA10), Bilbao, Spain, August 2010
- Ali Khodaei, Cyrus Shahabi, and Chen Li, [Hybrid Indexing and Seamless Ranking of Spatial and Textual Features of Web Documents](#), 21st International Conference on Database and Expert Systems Applications (DEXA10), Bilbao, Spain , August 2010
- Ugur Demiryurek, Farnoush Banaei Kashani , and Cyrus Shahabi, [Towards K-Nearest Neighbor Search in Time-Dependent Spatial Network Databases](#), 6th International Workshop on Databases in

[Networked Systems\(DNIS 2010\)](#), Aizu, Japan, March 2010

- Cyrus Shahabi, Farnoush Banaei-Kashani, Ali Khoshgozaran, Luciano Nocera, and Songhua Xing, [GeoDec: A Framework to Effectively Visualize and Query Geospatial Data for Decision-Making](#), [IEEE Multimedia](#), ISSN: 1070-986X,
- Ching-Chien Chen, Craig A. Knoblock, Cyrus Shahabi, [System and Method for Fusing Geospatial Data](#), Patent No. 7,660,441, United States, February 2010
- Leyla Kazemi, Farnoush Banaei-Kashani, Cyrus Shahabi, and Ramesh Jain, [Efficient Approximate Visibility Query in Large Dynamic Environments](#), [DASFAA](#), Tsukuba, Japan, April 2010
- Ugur Demiryurek, Farnoush Banaei-Kashani, and Cyrus Shahabi, [TransDec: A Spatiotemporal Query Processing Framework for Transportation Systems \(Demo\)](#), [26th IEEE International Conference on Data Engineering \(ICDE 2010\)](#), Long Beach, California, USA, March 2010
- Cyrus Shahabi, [Participatory Urban Data Collection: Planning and Optimization](#), Tsinghua University, Beijing, China, October 2009
- Ali Khodaei, Hyokyeong Lee, Farnoush Banaei-Kashani, Cyrus Shahabi, and Iraj Ershaghi, [A Mutual Information-Based Metric for Identification of Nonlinear Injector](#), [SPE Western Regional](#), San Jose, CA, March 2009
- Ugur Demiryurek, Bei Pan, Farnoush Banaei-Kashani, and Cyrus Shahabi, [Towards Modeling the Traffic Data on Road Networks](#), [ACMGIS 2009 Workshop on Computational Transportation Science](#), Seattle, Washington, November 2009
- Houtan Shirani-Mehr, Farnoush Banaei-Kashani, and Cyrus Shahabi, [Efficient Viewpoint Assignment for Urban Texture Documentation](#), [ACM SIGSPATIAL GIS 2009](#), Seattle, Washington, November 2009
- Ali Khoshgozaran and Cyrus Shahabi, [Private Buddy Search: Enabling Private Spatial Queries in Social Networks](#), [Symposium on Social Intelligence and Networking \(SIN09\)](#), In conjunction with [IEEE International Conference on Social Computing \(SocialCom09\)](#), Vancouver, Canada, August 2009
- Houtan Shirani-Mehr, Farnoush Banaei-Kashani, and Cyrus Shahabi, [Efficient Viewpoint Selection for Urban Texture Documentation](#), [Third International Conference on Geosensor Networks](#), Oxford, United Kingdom, July 2009
- Ugur Demiryurek, Farnoush Banaei-Kashani, and Cyrus Shahabi, [Efficient Continuous Nearest Neighbor Query in Spatial Networks using Euclidean Restriction](#), [11th International Symposium on Spatial and Temporal Databases \(SSTD09\)](#), Aalborg, Denmark, July 2009
- Yao-Yi Chiang, Craig A. Knoblock, Cyrus Shahabi, and Ching-Chien Chen, [Automatic and Accurate Extraction of Road Intersections from Raster Maps](#), [GeoInformatica](#), ISSN:1384-6175 (Print) 1573-7624 (Online), Issue:Volume 13, Number 2, June 2009
- Cyrus Shahabi, [GeoRealism: Expanding the Human Ability to Comprehend a Larger Geo-Space](#), NRO, Washington, D.C., February 2009
- Gabriel Ghinita, Panos Kalnis, Ali Khoshgozaran, Cyrus Shahabi, and Kian-Lee Tan, [Private Queries in Location Based Services: Anonymizers are not Necessary](#), [In Proceedings of ACM SIGMOD Conference 2008](#), Vancouver, Canada, June 2008
- Mehdi Sharifzadeh, Cyrus Shahabi, and Leyla Kazemi, [Processing Spatial Skyline Queries in both Vector Spaces and Spatial Network Databases](#), [accepted for publication in ACM TODS](#), June, 2009
- Ali Khoshgozaran, Houtan Shirani-Mehr, and Cyrus Shahabi, [SPIRAL: A Scalable Private Information Retrieval Approach to Location Privacy](#), [The 2nd International Workshop on Privacy-Aware Location-based Mobile Services \(PALMS\) In conjunction with the 9th International Conference on Mobile Data Management \(MDM08\)](#), Beijing, China, April 2008
- Ali Khoshgozaran, Ali Khodaei, Mehdi Sharifzadeh, and Cyrus Shahabi, [A hybrid aggregation and compression technique for road network databases](#), [Knowledge and Information Systems](#), ISSN 0219-1377 (Print) 0219-3116 (Online) ,
- Ali Khoshgozaran and Cyrus Shahabi, [Blind Evaluation of Nearest Neighbor Queries Using Space Transformation to Preserve Location Privacy](#), [10th International Symposium on Spatial and Temporal Databases](#), Boston, MA, July 2007

[Top](#)

## Visibility Computation for Trajectories



Visibility query is fundamental to many analysis and decision-making tasks in virtual environments. Visibility computation is time complex and the complexity escalates in large and dynamic environments, where the visibility set (i.e., the set of visible objects) of any viewpoint is prone to change at any time. However, exact visibility query is rarely necessary. Besides, it is inefficient, if not infeasible, to obtain the exact result in a dynamic environment. In this research, we formally define an Approximate Visibility Query (AVQ) as follows: given a viewpoint  $v$ , a distance  $e$  and a probability  $p$ , the answer to an AVQ for the viewpoint  $v$  is an approximate visibility set such that its difference with the exact visibility set is guaranteed to be less than  $e$  with confidence  $p$ . We propose an approach to correctly and efficiently answer AVQ in large and dynamic environments.

## People Involved

- Leyla Kazemi

## Image Acquisition Planning



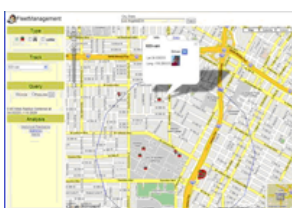
We envision GeoSIM (Geo Social Image Mapping) as a system with which a group of individuals with camera equipped mobile phones participate in collaborative collection of urban texture information. An important issue in this system is to generate participation plans for the users for fast and comprehensive social image mapping. The participation plans should satisfy various user constraints (e.g., limited available time) while maximizing the exposure of uncovered sight to users. For each user, the participation plan consists of a traversal path with multiple points along at which the user takes photos and collects texture. We model the problem of finding users' participation plans as a two step process. At the first step, termed viewpoint selection, a minimum number of points in the urban environment are selected from which the texture of the entire urban environment (the part visible to cameras) can be collected. At the second step, called viewpoint assignment, the selected viewpoints are assigned to the participating users such that given a limited number of users with various constraints, users can collectively capture the maximum amount of texture information within a limited time interval.

### People Involved

- Farnoush Banaei-Kashani
- Leyla Kazemi
- Houtan Shirani-Mehr
- Songhua Xing

[Top](#)

## K-Nearest Neighbor Queries Over Moving Objects on Road Networks



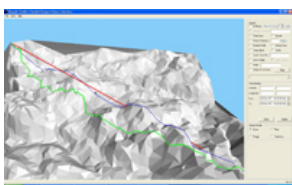
The wide spread usage of GPS devices and latest developments in wireless technologies has led to extensive usage of location-based services as well as emerging global positioning services. With rapidly growing interest in GPS enabled handheld devices, end users are able to report their positions hence becoming a point of interests and utilize various location-based services from the providers. One main class of such services is k-nearest neighbor ((k-NN) queries over moving objects on road networks. A k-nearest neighbor query finds the k point of interests (ie: restaurant, taxis) in a dataset that lie closest to a given query point. In the past, many efficient algorithms have been proposed to address k-NN query processing in Euclidian space. In real-world scenarios, however, the query points move in road networks, and the distance between it's point of interest is defined as the length of the shortest path connecting them. When processing k-NN queries over moving objects, the main challenges arise from a) the mobility of query points and/or point of interests b) fast road network distance calculation c) management of location updates and indexes. To cope with challenges, we propose a novel approach which initially retrieves the k-NN from a grid structure and then utilizes threshold-based algorithms to refine the result set. In addition, our approach takes into account changes in the network edges because of varying road conditions such as traffic, road construction, accidents. All these algorithms are designed for update streams, where the central server manages the location updates from a particular set of queries and objects.

### People Involved

- Ugur Demiryurek

[Top](#)

## Surface KNN Queries



A very important class of queries in Geo-spatial database is the k Nearest Neighbor (kNN) query. Various studies have focused on this problem and its variant in Euclidean space and road networks. Moreover, real world applications often result in the formulation of kNN queries on land surface. Surface kNN query (skNN) is a kind of constraint based query. In such queries, the objects can only move on the terrain. A skNN query actually involves two tasks, it needs to select out the k nearest neighbors to q, just as a traditional kNN query. What is more, the users are not satisfied by only telling which one is nearest, they also want to know how to get there. Thus the system has to point out the path from q to such points on the surface. We propose a novel approach to process skNN query with two spatial indexes, namely Loose Surface Index (LSI) and Tight Surface Index (TSI). With those two indexes, we bring down the computation cost (both I/O and CPU time) for surface distance computation. Our algorithm finds out skNN in an "expanding strategy": firstly we locate the query point, and search in the around area for the first nearest neighbor, then we expand the search area with certain manner to report k nearest neighbors incrementally.

A variant of the skNN query, known as CskNN query takes the time dimension into consideration: Given a fixed query point q, a CskNN query continuously monitor the k nearest moving neighbors of q. We propose two methods to tackle this problem. The first method, inspired by the existing techniques in monitoring kNN in road networks maintains an analogous counterpart of the Dijkstra Expansion Tree on land surface, called Surface Expansion Tree (SE-Tree). However, we show the concept of expansion tree for land surface does not work as SE-tree suffers from intrinsic defects: it is fat and short, and hence does not improve the query efficiency. Hence, we propose a superior approach that partitions SE-Tree into hierarchical chunks of pre-computed surface distances, called Angular Surface Index Tree (ASI-Tree). Unlike SE-tree, ASI-Tree is a well balanced thin and tall tree. With ASI-Tree, we can continuously monitor skNN queries efficiently with low CPU and I/O overheads by both speeding up the surface shortest path computations and localizing the searches.

### People Involved

- Songhua Xing

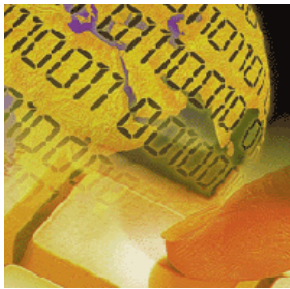
### Publications

- Songhua Xing and Cyrus Shahabi, [Scalable Shortest Paths Browsing on Land Surface](#), [ACM GIS](#), San Jose, CA, November 2010

- Songhua Xing, Cyrus Shahabi, and Bei Pan, [Continuous Monitoring of Nearest Neighbors on Land Surface](#), [35th International Conference on Very Large Data Bases](#), Lyon, France, August 2009
- Cyrus Shahabi, Lu-An Tang, and Songhua Xing, [Indexing Land Surface for Efficient kNN Query](#), [The 34th International Conference on Very Large Data Bases \(VLDB 08\)](#), Auckland, New Zealand, August 2008

[Top](#)

## Database Outsourcing



Outsourcing data to third party data providers is becoming a common practice for data owners to avoid the cost of managing and maintaining databases. Meanwhile, due to the popularity of location-based-services (LBS), the need for spatial data (e.g., gazetteers, vector data) is increasing exponentially. Consequently, we are witnessing a new trend of outsourcing spatial datasets by data collectors.

Two main challenges with outsourcing datasets is to keep the data private (from the data provider) and ensure the integrity of the query result (for the clients). Unfortunately, most of the techniques proposed for privacy and integrity do not extend to spatial data in a straightforward manner. Hence, recent studies proposed various techniques to support either privacy or integrity (but not both) on spatial datasets.

We propose a technique that can ensure both privacy and integrity for outsourced spatial data. In particular, we first use a one-way spatial transformation method based on Hilbert curves, which encrypts the spatial data before outsourcing and hence ensures its privacy. Next, by probabilistically replicating a portion of the data and encrypting it with a different encryption key, we devise a technique for the client to audit the trustworthiness of the query results. We show the applicability of our approach for both k-nearest-neighbor and spatial range queries, the building blocks of any LBS application.

### People Involved

- Ling Hu

### Publications

- Ling Hu, Wei-Shinn Ku, Spiridon Bakiras, and Cyrus Shahabi, [Spatial Query Integrity with Voronoi Neighbors](#), [TKDE](#), in press, 2012
- Ling Hu, Wei-Shinn Ku, Spiridon Bakiras, and Cyrus Shahabi, [Verifying Spatial Queries using Voronoi Neighbors](#), [ACM SIGSPATIAL](#), San Jose, California , November 2010
- Wei-Shinn Ku, Ling Hu, Cyrus Shahabi, and Haixun Wang, [Query Integrity Assurance of Location-based Services Accessing Outsourced Spatial Databases](#), [The 11th International Symposium on Spatial and Temporal Databases \(SSTD09\)](#), Aalborg, Denmark, July 2009

[Top](#)

## Private Participatory Urban Sensing



Mobile devices are becoming the largest sensor network around the world. They could be used to collect a large amount of data with little effort and cost which is leading to a promising future for participatory sensing networks or urban sensing. With these rapid technology advances, privacy concerns of the mobile users are the major inhibitors hindering massive participation.

Consider an application system as the following: suppose we have built a 3D model of an urban environment such as New York city and we want to have a participatory sensing system in which mobile users can upload city images taken with their mobile phones to a server. With these images, the server can texture map urban images onto their target Geo-locations to create a photo-realistic 3D representation. Such systems can recreate the real world with unprecedented details and quality and answer any questions with both temporal and spatial precision.

However, the user may not be willing to disclose his/her association with the picture to any party, including the server. With a user participatory system, mobile users may not want to release the ownership information due to safety considerations or political reasons. In this research work, we address the privacy problems in such participatory sensing systems and propose a private participatory sensing system framework.

### People Involved

- Ling Hu

[Top](#)

## Road Network Traffic Modeling



A spatiotemporal network is a spatial network (e.g., road network) along with the corresponding time-dependent weight (e.g., travel time) for each edge of the network. The design and analysis of policies and plans on spatiotemporal networks require realistic models that accurately represent the temporal behavior of such networks. We propose a traffic modeling framework for road networks that enables 1) generating an accurate temporal model from archived temporal data collected from a spatiotemporal network (so as to be able to publish the temporal model of the spatiotemporal network without having to release the real data), and 2) augmenting any given spatial network model with a corresponding realistic temporal model custom-built for that specific spatial network (in order to be able to generate a spatiotemporal network model from a solely spatial network model). We validate the accuracy of our proposed modeling framework via experiments. We also used the proposed framework to generate the temporal model of the Los Angeles County freeway network and publish it for public use.

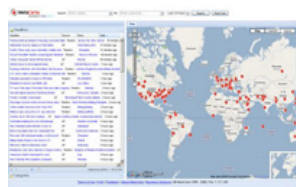
### People Involved



- Ugur Demiryurek
- Bei Pan (Penny)
- Farnoush Banaei-Kashani

[Top](#)

## Location-based Web Search



There is a significant commercial and research interest in location-based web search engines. Given a number of search keywords and one or more locations that user is interested in, a location-based web search finds and ranks the most textually and spatially relevant web pages according to query keywords and locations. With this type of search, both the spatial and textual information should be indexed. This generates a few challenges: space and text are two totally different data types requiring different index structures. For instance, conventional text engines are set-oriented while location indexes are usually based on two-dimensional and Euclidean spaces. This creates new research problems on how to represent the location(s) of each web page and also how to combine these two types of indexes efficiently. Such combined index structure, should also be able to simultaneously handle both the spatial and textual features of the web-pages and also rank them based on a well-defined scoring function. Existing approaches either filter the data based on one feature (space or text) first and then search based on the second feature or use hybrid index structures performing poor when one feature is more selective than the other. Also most of these approaches cannot rank web-pages based on a combination of space and text (they either do not rank, or rank based on one feature) and are not easy to integrate into existing search engines. We propose to address all the above shortcomings by designing a new index structure to expedite location-based web searches. Our indexing technique can easily be integrated into the current search engines.

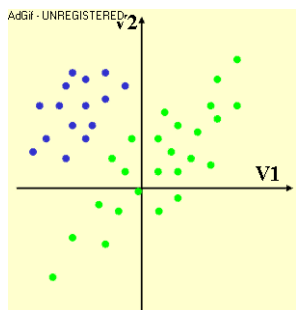
### People Involved

- Ali Khodaei

[Top](#)

## Past Research

### Subspace Method



Subspace Method (SM) is a process to find a lower-dimensional representation of high-dimensional data with particularly useful mathematic properties. It is therefore useful when dealing with multidimensional data sets that are commonly and inevitably suffering from the curse of dimensionality. In Linear Subspace Method (LSM), the lower-dimensional subspace is spanned by a small number of new components (or directions) that are formed as a linear function of the original variables. Therefore, the process is simply to find the linear coefficients of these linear functions by minimizing or maximizing a certain criterion.

LSM have been profitably exploited in many disciplines such as dimension reduction, feature selection/extraction, data compression/reconstruction, data visualization, etc, typically based on or extended from Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Multidimensional Scaling (MS), and more. PCA is to find a subspace that maximally preserves the variability in data, LDA is to find a discriminative subspace with maximum separation between classes, and MS is to find a two-dimensional subspace that maximally preserves the inter-point distances between the data items.

At InfoLab, LSM was exploited for feature subset selection of Multivariate Time Series (MTS) data set. The linear coefficients obtained from LSM are often interpreted as the contributions of the corresponding original variables, which was exploited to find an optimal subset of original variables least redundant each other and potentially effective to the subsequent data mining task as a whole. In specific, a subspace that is common to all MTS data items in the set and most effectively preserves the variability in each MTS item was first extracted, then the original variables were clustered based on their coefficients/contributions on forming this common subspace, and one representative variable from each cluster was finally selected as the subset. More recently, LSM has been exploited for the visualization of high-dimensional data. That is, a new linear subspace method, name Progression-Preserving Projection, was proposed to find a two- or three-dimensional subspace that most effectively preserves the class separation and the within-class progression structure so as to visually discover the most useful and informative structures presented in the original multidimensional data, which is particularly useful in the application of motor rehabilitation after stroke. Now, we are exploiting LSM approaches for the predictive modeling of multidimensional data.

### People Involved

- Hyunjin Yoon

### Publications

- Ugur Demiryurek, Farnoush Banaei-Kashani, Cyrus Shahabi, and Frank Wilkinson, [Neural Network Based Sensitivity Analysis for Injector Producer Relationship Identification](#), *SPE Intelligent Energy Conference*, Amsterdam, Netherlands, February 2008
- Hyunjin Yoon, Kiyoun Yang, and Cyrus Shahabi, [Feature Subset Selection and Feature Ranking for Multivariate Time Series](#), *IEEE Transactions on Knowledge and Data Engineering (TKDE)* - Special Issue on Intelligent Data Preparation, Volume: 17, Issue: 9, pp. 1186- 1198, ISSN: 1041-4347, September 2005
- Kiyoun Yang, Hyunjin Yoon, and Cyrus Shahabi, [CLEVer: A Feature Subset Selection Technique for Multivariate Time Series](#), *The Ninth Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD-05)*, LNAI, Volume 3518, pages 516-522, ISBN: 3-540-26076-5, Hanoi, Vietnam, May 2005
- Kiyoun Yang, Hyunjin Yoon, and Cyrus Shahabi, [CLEVer: a Feature Subset Selection Technique for Multivariate Time Series](#), University of Southern California, Computer Science Department, Technical Report 05-845, March 2005

## Texture Generation



Our research looks at practical and efficient ways to use images taken on a campaign setting using handheld devices for the purpose of creating textures to be used on virtualized locations. The motivation for this work stems by our intent to device a rapid and accurate modeling pipeline able to produce quality models on a large scales. In addition, existing models commonly obtained using LiDar imagery have poor textures and we expect to be able to find a pipeline that allow for their re-use.

Our approach attempts to blend techniques from computer vision and computer graphics and use all available pre-existing information to constrain the solution space. Available constraints include existing geometric models and textures of buildings associated with geospatial data, as well as information readily available from the mobile devices such as their GPS position at the time of acquisition.

Finally we plan to explore how those textures can be linked to information in a geographical information system to facilitate decision making type applications.

### People Involved

- Luciano Nocera
- Balki Ranganathan
- Frederick Zyda

## Nearest Neighbor Search



A very important class of queries in GIS applications is the class of K-Nearest Neighbor queries, where a set of K closest points of interest (e.g., hospitals or restaurants) to a query point is requested. Moreover, new real-world applications often result in the formulation of new variations of the KNN queries each demanding new a solution in presence of certain constraints. The challenge in all of these queries is to process the query as fast as possible when: 1) The dataset is usually large, and hence, conventional techniques usually lead to very long query processing time. 2) The distance function is usually computationally complex, which again usage of conventional approaches (e.g., Dijkstra's algorithm to find the shortest path between two points on a road network) will be impractical. Different approaches to address these challenges have been proposed. These approaches are usually based on approximation of the actual distance function with a simpler one (e.g., using Euclidean distance instead of shortest path) and utilization of spatial index structures. The problem with these approaches is that the simple distance function may not (and often does not) provide an acceptable approximation for the distance function. At USC InfoLab, we focus on nearest neighbor problem for road networks, where the distance function between two points is usually their shortest path. We address this problem from different angles. First, we propose a more accurate (as compare to Euclidean distance) approximation of the shortest path distance function by transforming the original network to a higher dimensional space. This approach is appropriate when some level of inaccuracy can be tolerated. Second, we propose an approach based on pre-calculating and indexing the solution space. This approach, based on Voronoi diagram, provides 100% accuracy and can also be extended to address constrained nearest neighbor queries and to find the distance between two arbitrary points. Finally, we studied an unexploited and novel form of KNN queries named Optimal Sequenced Route (OSR) query in both vector and metric spaces. OSR strives to find a route of minimum length starting from a given source location and passing through a number of typed locations in a specific sequence imposed on the types of the locations. Our experiments verified that our algorithms significantly outperform the naive approach in terms of processing time (up to two orders of magnitude) and required workspace (up to 90% reduction on average).

### People Involved

- Mohammad-Reza Kolahdouzan
- Mehdi Sharifzadeh

### Publications

- Ling Hu, Yinan Jing, Wei-Shinn Ku, and Cyrus Shahabi, [Enforcing k Nearest Neighbor Query Integrity on Road Networks](#), *ACM SIGSPATIAL GIS*, Redondo Beach, CA, November 2012
- Ling Hu, Wei-Shinn Ku, Spiridon Bakiras, and Cyrus Shahabi, [Spatial Query Integrity with Voronoi Neighbors](#), *TKDE*, in press, 2012
- Ugur Demiryurek and Cyrus Shahabi, [Indexing Network Voronoi Diagrams](#), *The 17th International Conference on Database Systems for Advanced Applications*, Busan, South Korea , April 2012
- Ali Khodaei, Cyrus Shahabi, and Chen Li, [SKIF-P: a point-based indexing and ranking of web documents for spatial-keyword search](#), *Geoinformatica*, Publisher: Springer Netherlands, Issn: 1384-6175, Doi: 10.1007/s10707-011-0142-7, October 2011
- Ugur Demiryurek, Farnoush Banaei-Kashani, Cyrus Shahabi , and Anand Ranganathan, [Online Computation of Fastest Path in Time-Dependent Spatial Networks](#), *12th International Symposium on Spatial and Temporal Databases (SSTD11)*, Minneapolis, MN, USA, August 2011
- Afsin Akdogan, Ugur Demiryurek, Farnoush Banaei-Kashani , and Cyrus Shahabi, [Voronoi-based Geospatial Query Processing with MapReduce](#), *IEEE Cloud Computing Technology and Science (CloudCom)*, Best Paper Award, Indianapolis, IN, November 2010
- Ugur Demiryurek, Farnoush Banaei-Kashani, and Cyrus Shahabi, [A Case for Time-Dependent Shortest Path Computation in Spatial Networks](#), *ACM SIGSPATIAL*, San Jose, CA, November 2010
- Ugur Demiryurek, Farnoush Banaei-Kashani, and Cyrus Shahabi, [Efficient K-Nearest Neighbor Search in Time-Dependent Spatial Networks](#), *21st International Conference on Database and Expert Systems*

[Applications \(DEXA10\)](#), Bilbao, Spain, August 2010

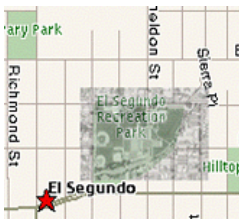
- Ali Khodaei, Cyrus Shahabi, and Chen Li, [Hybrid Indexing and Seamless Ranking of Spatial and Textual Features of Web Documents](#), [21st International Conference on Database and Expert Systems Applications \(DEXA10\)](#), Bilbao, Spain, August 2010
- Ugur Demiryurek, Farnoush Banaei Kashani, and Cyrus Shahabi, [Towards K-Nearest Neighbor Search in Time-Dependent Spatial Network Databases](#), [6th International Workshop on Databases in Networked Systems \(DNIS 2010\)](#), Aizu, Japan, March 2010
- Ugur Demiryurek, Farnoush Banaei-Kashani, and Cyrus Shahabi, [TransDec: A Spatiotemporal Query Processing Framework for Transportation Systems \(Demo\)](#), [26th IEEE International Conference on Data Engineering \(ICDE 2010\)](#), Long Beach, California, USA, March 2010
- Cyrus Shahabi and Mehdi Sharifzadeh, [Voronoi Diagrams for Query Processing](#), [Encyclopedia of GIS](#) 2008, 1235-1240,
- Cyrus Shahabi and Mehdi Sharifzadeh, [Voronoi Diagrams](#), [Encyclopedia of Database Systems](#) 2009, 3438-3440,
- Cyrus Shahabi, [Road Networks](#), [Encyclopedia of Database Systems](#) 2009, 2442-2447,
- Ugur Demiryurek, Bei Pan, Farnoush Banaei-Kashani, and Cyrus Shahabi, [Towards Modeling the Traffic Data on Road Networks](#), [ACMGIS 2009 Workshop on Computational Transportation Science](#), Seattle, Washington, November 2009
- Ali Khoshgozaran and Cyrus Shahabi, [Private Buddy Search: Enabling Private Spatial Queries in Social Networks](#), [Symposium on Social Intelligence and Networking \(SIN09\). In conjunction with IEEE International Conference on Social Computing \(SocialCom09\)](#), Vancouver, Canada, August 2009
- Ali Khoshgozaran and Cyrus Shahabi, [Private Information Retrieval Techniques for Enabling Location Privacy in Location-Based Services](#), [Privacy in Location-based Applications](#), ISBN: 978-3-642-03510-4, Bettini, C.; Jajodia, S.; Samarati, P.; Wang, X.S. Eds., October 2009
- Ugur Demiryurek, Farnoush Banaei-Kashani, and Cyrus Shahabi, [Efficient Continuous Nearest Neighbor Query in Spatial Networks using Euclidean Restriction](#), [11th International Symposium on Spatial and Temporal Databases \(SSTD09\)](#), Aalborg, Denmark, July 2009
- Mehdi Sharifzadeh and Cyrus Shahabi, [Processing Optimal Sequenced Route Queries Using Voronoi Diagrams](#), [GeoInformatica](#), Publisher: Springer Netherlands. Page: 411-433, ISSN: 1384-6175 (Print) 1573-7624 (Online), December 2008
- Wan D. Bae, Shayma Alkobaisi, Seon Ho Kim, Sada Narayanappa, and Cyrus Shahabi, [Web data retrieval: solving spatial range queries using k -nearest neighbor searches](#), [GeoInformatica](#), ISSN:1384-6175 (Print) 1573-7624 (Online), August 2008
- Cyrus Shahabi, [Geospatial Decision Making and Spatial Skyline Queries](#), University of Tokyo, Tokyo, Japan, October 2007
- Wan D. Bae, Shayma Alkobaisi, Seon Ho Kim, Sada Narayanappa, and Cyrus Shahabi, [Supporting Range Queries on Web Data Using k-Nearest Neighbor Search](#), [W2GIS](#), Cardiff, UK, November 2007
- Wan D. Bae, Shayma Alkobaisi, Seon Ho Kim, Sada Narayanappa, and Cyrus Shahabi, [Supporting Range Queries on Web Data Using k -Nearest Neighbor Search](#), [WebDB 2007](#), Beijing, China, June 2007
- Ali Khoshgozaran and Cyrus Shahabi, [Blind Evaluation of Nearest Neighbor Queries Using Space Transformation to Preserve Location Privacy](#), [10th International Symposium on Spatial and Temporal Databases](#), Boston, MA, July 2007
- Mehdi Sharifzadeh, Mohammad Kolahdouzan, and Cyrus Shahabi, [The Optimal Sequenced Route Query](#), [The VLDB Journal: The International Journal on Very Large Data Bases](#), SSN: 1066-8888 (Print) 0949-877X (Online), DOI: 10.1007/s00778-006-0038-6, Issue: Online First, 2007
- Cyrus Shahabi, [GeoDec: Enabling Geospatial Decision Making](#), ETRI, Korea, September 2006
- Ching-Chien Chen, Cyrus Shahabi, Craig A. Knoblock, and Mohammad Kolahdouzan, [Automatically and Efficiently Matching Road Networks with Spatial Attributes in Unknown Geometry Systems](#), [The Third Workshop on Spatio-Temporal Database Management \(STDBM'06 in conjunction with VLDB'06\)](#), ISSN 1613-00174, online CEUR-WS.org/Vol-174/paper1.pdf, Seoul, Korea, September 2006
- Cyrus Shahabi, [GeoDec: Enabling Geospatial Decision Making](#), Air Force Research Lab (AFRL) - Chief Scientist seminar series, Rome, NY, June 2006
- Cyrus Shahabi, [GeoDec: Enabling Geospatial Decision Making](#), Google, Mountain View, CA, May 2006
- Mohammad R. Kolahdouzan and Cyrus Shahabi, [Alternative Solutions for Continuous K Nearest Neighbor Queries in Spatial Network Databases](#), [GeoInformatica](#), 9:4, pp. 321-341, 2005
- Mehdi Sharifzadeh, Cyrus Shahabi, and Craig A. Knoblock, [Learning Approximate Thematic Maps from Labeled Geospatial Data](#), [Next Generation Geospatial Information: From Digital Image Analysis to Spatiotemporal Databases](#), ISBN 0415380499, Peggy Agouris and Arie Croituru Eds., Taylor & Francis Group, July 2005
- Mehdi Sharifzadeh, Mohammad Kolahdouzan, and Cyrus Shahabi, [The Optimal Sequenced Route Query](#), University of Southern California, Technical Report 05-840, Computer Science Department, January

2005

- Mohammad R. Kolahdouzan and Cyrus Shahabi, [Continuous K Nearest Neighbor Queries in Spatial Network Databases](#), *The Second Workshop on Spatio-Temporal Database Management (STDBM04)*, co-located with VLDB04, Toronto, Canada, August 2004
- Mohammad R. Kolahdouzan and Cyrus Shahabi, [Voronoi-Based K Nearest Neighbor Search for Spatial Network Databases](#), *30th International Conference on Very Large Data Bases (VLDB)*, Toronto, Canada, September 2004
- Cyrus Shahabi, Mohammad R. Kolahdouzan, and Mehdi Sharifzadeh, [A Road Network Embedding Technique for K-Nearest Neighbor Search in Moving Object Databases](#), *10th ACM International Symposium on Advances in Geographic Information Systems (ACM-GIS'02)*, McLean, VA, November, 2002
- Cyrus Shahabi, Mohammad R. Kolahdouzan, and Mehdi Sharifzadeh, [A Road Network Embedding Technique for K-Nearest Neighbor Search in Moving Object Databases](#), *Geoinformatica*, Kluwer Academic Publishers, Volume 7, Issue 3, September 2003

[Top](#)

## Geospatial Data Fusion



There is a vast amount of geospatial data available, including image, map, vector, point, and elevation data. The users of these data products often want these different data sources displayed in some integrated fashion. However; accurately and automatically integrating diverse geo-spatial datasets is a difficult task, since there are often certain spatial inconsistencies between geo-spatial datasets. They are multiple reasons why different data products may not align: they may have been collected at different resolutions, they may use different projections of the earth, they may have been corrected in different ways, etc. Conflation is often a term used to describe the integration or alignment of different geospatial data sets. However, the difficulty with current conflation techniques is that they require the manual identification of a set of control points to properly conflate two data sets. We are developing techniques to automatically conflating different sources of geospatial data. The critical challenge in automatic conflation is to automatically identify an appropriate set of control points on the two sources to be integrated. We propose to do this by combining different sources of information from each of the sources to be integrated. For example, in integrating vector data with imagery, we can exploit what is known or can be inferred about each of the data sources. For the imagery, we can exploit points that have been identified in the imagery or perform imager processing on the image. For vector data of roads, we can exploit meta-data about the vectors, such as address ranges, road names, or even the number of lanes and type of road surface. In addition, we can only analyze the road network to determine the location of intersections. For other types of vector data, there are similar types of meta data and inferences that can be exploited. This information can then be augmented with other sources of data. For example, points on the imagery can be determined by looking up the points in an online telephone book. Furthermore, we propose an information integration approach, which utilizes common vector datasets as "glue" to automatically conflate imagery with street maps. We have also demonstrated the utility of our approach using multiple evaluation methods.

### People Involved

- Dr. Ching-Chien (Jason) Chen

### Publications

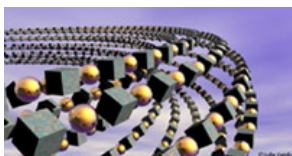
- Ali Khodaei, Cyrus Shahabi, and Chen Li, [SKIF-P: a point-based indexing and ranking of web documents for spatial-keyword search](#), *Geoinformatica*, Publisher: Springer Netherlands, Issn: 1384-6175, Doi: 10.1007/s10707-011-0142-7, October 2011
- Ugur Demiryurek, Farnoush Banaei-Kashani, Cyrus Shahabi, and Anand Ranganathan, [Online Computation of Fastest Path in Time-Dependent Spatial Networks](#), *12th International Symposium on Spatial and Temporal Databases (SSTD11)*, Minneapolis, MN, USA, August 2011
- Ugur Demiryurek, Farnoush Banaei-Kashani, and Cyrus Shahabi, [Efficient K-Nearest Neighbor Search in Time-Dependent Spatial Networks](#), *21st International Conference on Database and Expert Systems Applications (DEXA10)*, Bilbao, Spain, August 2010
- Cyrus Shahabi, Farnoush Banaei-Kashani, Ali Khoshgozaran, Luciano Nocera, and Songhua Xing, [GeoDec: A Framework to Effectively Visualize and Query Geospatial Data for Decision-Making](#), *IEEE Multimedia*, ISSN: 1070-986X,
- Ugur Demiryurek, Farnoush Banaei-Kashani, and Cyrus Shahabi, [TransDec: A Spatiotemporal Query Processing Framework for Transportation Systems \(Demo\)](#), *26th IEEE International Conference on Data Engineering (ICDE 2010)*, Long Beach, California, USA, March 2010
- Cyrus Shahabi, [Participatory Urban Data Collection: Planning and Optimization](#), Tsinghua University, Beijing, China, October 2009
- Luciano Nocera, Arjun Rihan, Songhua Xing, Ali Khodaei, Ali Khoshgozaran, Farnoush Banaei-Kashani, and Cyrus Shahabi, [GeoDec : A Multi-Layered Query Processing Framework for Spatio-Temporal Data](#), *ACM SIGSPATIAL GIS*, Seattle, Washington, November 2009
- Ugur Demiryurek, Bei Pan, Farnoush Banaei-Kashani, and Cyrus Shahabi, [Towards Modeling the Traffic Data on Road Networks](#), *ACMGIS 2009 Workshop on Computational Transportation Science*, Seattle, Washington, November 2009
- Ugur Demiryurek, Farnoush Banaei-Kashani, and Cyrus Shahabi, [TransDec: A Data-Driven Framework for Decision-Making in Transportation Systems](#), *Transportation Research Forum (TRF09)*, Portland, OR, March 2009



- Cyrus Shahabi, **GeoRealism: Expanding the Human Ability to Comprehend a Larger Geo-Space**, NRO, Washington, D.C., February 2009
- Ching-Chien Chen, Craig A. Knoblock, and Cyrus Shahabi, **Automatically and Accurately Conflating Raster Maps with Orthoimagery**, [GeoInformatica](#), Page:377-410, ISSN:1384-6175 (Print) 1573-7624 (Online),, September 2008
- Cyrus Shahabi, **Location Privacy in Geospatial Decision Making**, [5th International Workshop on Databases in Networked Information Systems \(DNIS 2007\)](#), Aizu-Wakamatsu, Japan, October 2007
- Craig A. Knoblock and Cyrus Shahabi, **Geospatial Data Integration**, The Handbook of Geographic Information Science, ISBN: 9781405107969, Pages 185-196, Editors: John P. Wilson , A. Stewart Fotheringham; Publisher: Blackwell, 2008
- Cyrus Shahabi, **Geospatial Decision Making and Spatial Skyline Queries**, Peking University (PKU), Beijing, China, July 2007
- Cyrus Shahabi, **Geospatial Decision Making and Spatial Skyline Queries**, Chinese Academy of Sciences (CAS), Beijing, China, July 2007
- Cyrus Shahabi, **Geospatial Information Extraction and Integration**, Lawrence Livermore National Laboratory, Livermore, CA, February 2007
- Ching-Chien Chen, Craig A. Knoblock, and Cyrus Shahabi, **Automatically Conflating Road Vector Data with Orthoimagery**, [GeoInformatica](#), DOI: 10.1007/s10707-006-0344-6, Vol.10, No.4, Page 495-530, 2006
- Cyrus Shahabi, **GeoDec: Enabling Geospatial Decision Making**, ETRI, Korea, September 2006
- Cyrus Shahabi, **GeoDec: Enabling Geospatial Decision Making**, Air Force Research Lab (AFRL) - Chief Scientist seminar series, Rome, NY, June 2006
- Cyrus Shahabi, **GeoDec: Enabling Geospatial Decision Making**, Google, Mountain View, CA, May 2006
- Cyrus Shahabi, Yao-Yi Chiang, Kelvin Chung, Kai-Chen Huang, Jeff Khoshgozaran-Haghighi, Craig Knoblock, Sung Chun Lee, Ulrich Neumann, Ram Nevatia, Arjun Rihan, Snehal Thakkar, and Suya You, **GeoDec: Enabling Geospatial Decision Making**, [IEEE International Conference on Multimedia & Expo \(ICME\)](#), Toronto, Canada, July 2006
- Cyrus Shahabi, **Enabling Geospatial Decision Making**, Microsoft Research Center, Redmond, WA, December 2005
- Mehdi Sharifzadeh, Cyrus Shahabi, and Craig A. Knoblock, **Learning Approximate Thematic Maps from Labeled Geospatial Data**, [Next Generation Geospatial Information: From Digital Image Analysis to Spatiotemporal Databases](#), ISBN 0415380499, Peggy Agouris and Arie Croituru Eds., Taylor & Francis Group, July 2005
- Ching-Chien Chen, Cyrus Shahabi, and Craig A. Knoblock, **Utilizing Road Network Data for Automatic Identification of Road Intersections from High Resolution Color Orthoimagery**, [The Second Workshop on Spatio-Temporal Database Management \(STDBM04\)](#), co-located with the 30th International Conference on Very Large Data Bases (VLDB'04), Toronto, Canada, August 2004
- Ching-Chien Chen, Craig A. Knoblock, Cyrus Shahabi, Yao-Yi Chiang, and Snehal Thakkar, **Automatically and Accurately Conflating Orthoimagery and Street Maps**, [The 12th ACM International Symposium on Advances in Geographic Information Systems \(ACM-GIS'04\)](#), Washington, D.C., November 2004
- Cyrus Shahabi, Mohammad R. Kolahdouzan, and Maytham Safar, **Alternative Strategies for Performing Spatial Joins on Web Sources**, [Knowledge and Information Systems \(KAIS\) by Springer-Verlag](#), Volume 6, Number 3, pp. 290-314, ISSN: 0219-1377 (Paper) 0219-3116 (Online), May 2004
- Ching-Chien Chen, Craig A. Knoblock, Cyrus Shahabi, and Snehal Thakkar, **Building Finder: A System to Automatically Annotate Buildings in Satellite Imagery**, [International Workshop on Next Generation Geospatial Information \(NG2I'03\)](#), Cambridge (Boston), Massachusetts, USA, October, 2003
- Ching-Chien Chen, Snehal Thakkar, Craig Knoblock, and Cyrus Shahabi, **Automatically Annotating and Integrating Spatial Datasets**, [The 8th International Symposium on Spatial and Temporal Databases \(SSTD'03\)](#), ISBN: 3-540-40535-6, pp. 469-488, Santorini island, Greece, July, 2003
- Ching-Chien Chen, Craig A. Knoblock, Cyrus Shahabi, and Snehal Thakkar, **Automatically and Accurately Conflating Satellite Imagery and Maps**, [International Workshop on Next Generation Geospatial Information \(NG2I'03\)](#), Cambridge (Boston), Massachusetts, USA, October, 2003

[Top](#)

## Spatial Data Stream



A spatial data stream is a data stream consisting of geometric objects such as points, lines, polygons or intervals associated with a set of non-spatial attributes. These data items are streamed to a central processing module from a stream source such as a sensor node within a sensor network, a GPS-equipped truck adventuring in a desert or a tracker device attached to a user's body in an immersive environment. The emerging research area of algorithms for spatial data streams is the development of novel one-pass computational geometry algorithms which use small (sub-linear) amount of memory space and per-item processing time. An example of such algorithms is to find and maintain an approximate convex hull of a stream of 2-d points over a specified time window. The algorithm can be

applied to a network of moving sensors to monitor the extent of the area they cover. Such geometric algorithms are building blocks of spatial queries on spatial data streams. They are also capable of processing non-spatial streams. For example, flows in an IP network may be modeled as intervals ([start time, end time]) in a 1-d space, or user document downloads on a peer-to-peer network can be mapped into a high dimensional space generating a spatial data stream.

At USC infolab, we intend to focus on addressing different query processing issues on spatial data streams. We study the practicality of our algorithms on sensor networks and immersive environments as our running applications.

### People Involved

- Mehdi sharifzadeh

### Publications

- Seyed Jalal Kazemitabar, Ugur Demiryurek, Mohamed Ali, Afsin Akdogan, and Cyrus Shahabi, [Geospatial Stream Query Processing using Microsoft SQL Server StreamInsight, VLDB](#), (demonstration), Singapore, September 2010
- Ugur Demiryurek, Farnoush Banaei-Kashani, and Cyrus Shahabi, [TransDec: A Spatiotemporal Query Processing Framework for Transportation Systems \(Demo\)](#), 26th IEEE International Conference on Data Engineering (ICDE 2010), Long Beach, California, USA, March 2010
- SunHee Yoon and Cyrus Shahabi, [Distributed Spatial Skyline Query Processing in Wireless Sensor Networks](#), 3rd International Workshop on Sensor Webs, Databases and Mining in Networked Sensing Systems (SWDMNSS 2009), Carnegie Mellon University, Pittsburgh, USA, June 2009
- Ugur Demiryurek, Farnoush Banaei-Kashani, and Cyrus Shahabi, [TransDec: A Data-Driven Framework for Decision-Making in Transportation Systems](#), Transportation Research Forum (TRF09), Portland, OR, March 2009
- Cyrus Shahabi, [GeoRealism: Expanding the Human Ability to Comprehend a Larger Geo-Space](#), NRO, Washington, D.C., February 2009
- SunHee Yoon and Cyrus Shahabi, [A Distributed Algorithm to Compute Spatial Skylines in Wireless Sensor Networks](#), IPSN 2009, poster presentation, San Francisco, April 2009
- Mehdi Sharifzadeh and Cyrus Shahabi, [Approximate voronoi cell computation on spatial data streams](#), The VLDB Journal, ISSN: 1066-8888 (Print) 0949-877X (Online), Volume 18, Number 1 / January, 2009,
- Cyrus Shahabi, [Geospatial Decision Making and Spatial Skyline Queries](#), Peking University (PKU), Beijing, China, July 2007
- Cyrus Shahabi, [Geospatial Decision Making and Spatial Skyline Queries](#), Chinese Academy of Sciences (CAS), Beijing, China, July 2007
- SunHee Yoon and Cyrus Shahabi, [The Clustered AGgregation \(CAG\) technique leveraging spatial and temporal correlations in wireless sensor networks](#), ACM Transactions on Sensor Networks (TOSN), ISSN:1550-4859, Volume 3, Issue 1, March 2007
- Cyrus Shahabi, [GeoDec: Enabling Geospatial Decision Making](#), ETRI, Korea, September 2006
- Mehdi Sharifzadeh and Cyrus Shahabi, [Additively Weighted Voronoi Diagrams for Optimal Sequenced Route Queries](#), The Third Workshop on Spatio-Temporal Database Management (STDBM'06 in conjunction with VLDB'06), Seoul, Korea, September 2006
- Mehdi Sharifzadeh and Cyrus Shahabi, [The Spatial Skyline Queries](#), 32nd International Conference on Very Large Data Bases (VLDB'06), Seoul, Korea, September 2006
- Cyrus Shahabi, [GeoDec: Enabling Geospatial Decision Making](#), Air Force Research Lab (AFRL) - Chief Scientist seminar series, Rome, NY, June 2006
- Cyrus Shahabi, [GeoDec: Enabling Geospatial Decision Making](#), Google, Mountain View, CA, May 2006
- Cyrus Shahabi, Yao-Yi Chiang, Kelvin Chung, Kai-Chen Huang, Jeff Khoshgozaran-Haghighi, Craig Knoblock, Sung Chun Lee, Ulrich Neumann, Ram Nevatia, Arjun Rihan, Snehal Thakkar, and Suyu You, [GeoDec: Enabling Geospatial Decision Making](#), IEEE International Conference on Multimedia & Expo (ICME), Toronto, Canada, July 2006
- Mehdi Sharifzadeh and Cyrus Shahabi, [Utilizing Voronoi Cells of Location Data Streams for Accurate Computation of Aggregate Functions in Sensor Networks](#), GeoInformatica, Volume 10, Number 1, pages 9-36, ISSN: 1384-6175 (Paper) 1573-7624 (Online), March 2006
- Mehdi Sharifzadeh, Cyrus Shahabi, and Craig A. Knoblock, [Learning Approximate Thematic Maps from Labeled Geospatial Data](#), Next Generation Geospatial Information: From Digital Image Analysis to Spatiotemporal Databases, ISBN 0415380499, Peggy Agouris and Arie Croituru Eds., Taylor & Francis Group, July 2005
- Mehdi Sharifzadeh and Cyrus Shahabi, [Voronoi Cell Computation on Geometric Data Streams](#), University of Southern California, Technical Report 04-835, Computer Science Department, 2004

- Mehdi Sharifzadeh and Cyrus Shahabi, [Supporting Spatial Aggregation in Sensor Network Databases](#), *The 12th ACM International Symposium on Advances in Geographic Information Systems (ACM-GIS'04)*, pages 166-175, ISBN 1-58113-979-9, Washington, D.C., November 2004
- Mehdi Sharifzadeh, Cyrus Shahabi, Bahareh Navai, Farid Parvini, and Albert A. Rizzo, [AIDA: an Adaptive Immersive Data Analyzer](#), *The 30th International Conference on Very Large Data Bases (VLDB04)*, (demonstration), Toronto, Canada, September, 2004
- Mehdi Sharifzadeh, Cyrus Shahabi, and Craig A. Knoblock, [Learning Approximate Thematic Maps from Labeled Geospatial Data](#), *International Workshop on Next Generation Geospatial Information (NG2I'03)*, Cambridge (Boston), Massachusetts, USA, October, 2003

[Top](#)

## Temporal Abstraction of Immersive Data Stream



Immersive data streams or immersidata consist of multidimensional sensor data streams acquired from a user's interactions with an immersive environment. One example of this type of data is the set of data streams received from the sensors of a Cyber Glove (a virtual-reality glove). The Cyber Glove data streams can then be used for hand-gesture recognition. While traditional gesture recognition techniques often suffer from being highly device-dependent and not extendible for recognizing new signs, we propose a novel multi-layer framework for representing immersive sensor data streams that addresses both shortcomings. Our framework achieves both device independence and extendibility by separating the process from data within its multiple layers.

This framework consumes the streams of data and constructs a multi-level, spatio temporal representation of the data for further manipulation and processing. The lowest layer, termed 'raw data layer' continuously acquires sensor data streams. The next layer contains a set of predicates that describe general posture of the hand and its movement in space. The top layer contains a set of templates that corresponds to different hand gestures. The focus of this research is on the 'temporal' aspect of the data streams to extract hand movement trajectories. In particular we plan to:

- Define a set of temporal predicates that can describe the movements of the hand,
- Detect the trajectory of the hand movement from the sensor data streams,
- Design an algorithm that can extract the temporal predicates based on these trajectories, and
- Recognize the gestures by combining the extracted temporal templates with the spatial templates.

### People Involved

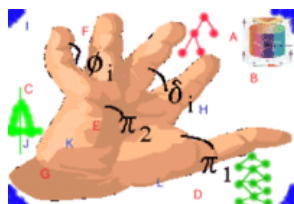
- Farid Parvini

### Publications

- Cyrus Shahabi, [TransDec: A Big-Data Framework for Decision-Making in Transportation Systems](#), Database Seminar Series, UC-San Diego, San Diego, CA, May 2012
- Hyunjin Yoon and Cyrus Shahabi, [Robust Time-Referenced Segmentation of Moving Object Trajectories](#), *Eighth IEEE International Conference on Data Mining (ICDM)*, Pisa, Italy, December 2008
- SunHee Yoon and Cyrus Shahabi, [Distributed Spatial Skyline Query Processing in Wireless Sensor Networks](#), *3rd International Workshop on Sensor Webs, Databases and Mining in Networked Sensing Systems (SWDMNSS 2009)*, Carnegie Mellon University, Pittsburgh, USA, June 2009
- SunHee Yoon and Cyrus Shahabi, [A Distributed Algorithm to Compute Spatial Skylines in Wireless Sensor Networks](#), *IPSN 2009*, poster presentation, San Francisco, April 2009
- SunHee Yoon and Cyrus Shahabi, [The Clustered AGgregation \(CAG\) technique leveraging spatial and temporal correlations in wireless sensor networks](#), *ACM Transactions on Sensor Networks (TOSN)*, ISSN:1550-4859, Volume 3, Issue 1, March 2007
- Farid Parvini and Cyrus Shahabi, [UTGeR: A User-Independent Technique for Gesture Recognition](#), *11th International Conference on Human-Computer Interaction*, (demonstration), Las Vegas, July 2005
- Farid Parvini and Cyrus Shahabi, [Utilizing Bio-Mechanical Characteristics For User-Independent Gesture Recognition](#), *International Workshop on Biomedical Data Engineering (BMDE2005)*, Tokyo, Japan, April 2005

[Top](#)

## Multivariate Time Series Analysis



The Multivariate Time Series (MTS) dataset is a common data type in various scientific, medical and financial fields, where more than one variable (e.g., sensor) is used for monitoring the environment and its status changes over time. An MTS is usually very high dimensional with its main distinguishing characteristic being the inter-correlations and/or interdependencies among its variables. Consequently, MTS may not be easily broken into multiple univariate time series and should be treated as a whole.

There are several research directions for Multivariate Time Series Analysis (MTSA). One is to find an appropriate similarity measure for comparing two MTS's. The similarity measure should take the characteristics of MTS into consideration, and must be efficiently computable. Another is to devise indexing structures for MTS data. The traditional SAM (Spatial Access Method) techniques may not work here due to the high dimensionality of the MTS datasets. Currently, we are extending SVD-based similarity measures for MTS datasets by representing an MTS as a matrix. For more general cases, an n-way SVD/PCA technique should be employed to deal with MTS's represented as tensors. So far, we have focused on k Nearest Neighbor search for MTS data, and we are trying to extend this to more general analysis on MTS streams.

## People Involved

- Kiyoung Yang
- Hyunjin Yoon

## Publications

- Hyunjin Yoon and Cyrus Shahabi, [Robust Time-Referenced Segmentation of Moving Object Trajectories](#), [Eighth IEEE International Conference on Data Mining \(ICDM\)](#), Pisa, Italy, December 2008
- Ali Khodaei, Hyokyeong Lee, Farnoush Banaei-Kashani, Cyrus Shahabi, and Iraj Ershaghi, [A Mutual Information-Based Metric for Identification of Nonlinear Injector](#), [SPE Western Regional](#), San Jose, CA, March 2009
- Hyunjin Yoon and Cyrus Shahabi, [Feature Subset Selection on Multivariate Time Series with Extremely Large Spatial Features](#), [Sixth IEEE International Conference on Data Mining - Workshops \(ICDMW'06\)](#), 18-22 December 2006, Hong Kong, China
- Cyrus Shahabi, Tim Marsh, Kiyoung Yang, Hyunjin Yoon, Albert A. Rizzo, Minyoung Mun, and Margaret McLaughlin, [Immersidata Analysis: Four Case Studies](#), [IEEE Computer volume 40](#), number 7, pp45-52, July 2007
- Tim Marsh, Kiyoung Yang, Cyrus Shahabi, and Seon Ho Kim, [Immersidata Management and Analysis For Game Development and Assessment For Staying There](#), [International Conference on Digital Interactive Media Entertainment & Arts \(DIME\)](#), Bangkok, Thailand, October 2006
- Tim Marsh, Kiyoung Yang, and Cyrus Shahabi, [Game development for experience through staying there](#), [ACM SIGGRAPH Video Game Symposium \(co-located with SIGGRAPH2006\)](#), Boston, MA, USA, July 2006
- Tim Marsh, Shamus P. Smith, Kiyoung Yang, and Cyrus Shahabi, [Continuous and Unobtrusive Capture of User-Player Behavior and Experience to Assess and Inform Game Design and Development](#), [1st World Conference for Fun 'n Games](#), Preston, England, June 2006
- Kiyoung Yang and Cyrus Shahabi, [On the Stationarity of Multivariate Time Series for Correlation-Based Data Analysis](#), [The Fifth IEEE International Conference on Data Mining \(ICDM'05\)](#), pp. 805-808, ISBN 0-7695-2278-5, Houston, Texas, USA, November 2005
- Kiyoung Yang, Tim Marsh, Minyoung Mun, and Cyrus Shahabi, [Continuous Archival and Analysis of User Data in Virtual and Immersive Game Environments](#), [The 2nd ACM Workshop on Capture, Archival and Retrieval of Personal Experiences \(CARPE'05\)](#), pp. 13 - 22, ISBN:1-59593-246-1, Singapore, November 2005
- Kiyoung Yang and Cyrus Shahabi, [A PCA-based Kernel for Kernel PCA on Multivariate Time Series](#), [ICDM 2005 Workshop on Temporal Data Mining: Algorithms, Theory and Applications held in conjunction with The Fifth IEEE International Conference on Data Mining \(ICDM'05\)](#), Houston, Texas, USA, November 2005
- Kiyoung Yang and Cyrus Shahabi, [On the Stationarity of Multivariate Time Series for Correlation-Based Data Analysis](#), University of Southern California, Computer Science Department, Technical Report 05-845, September 2005
- Tim Marsh, Kiyoung Yang, Cyrus Shahabi, Wee Ling Wong, Luciano Nocera, Eduardo Carriazo, Aditya Varma, Hyunjin Yoon, and Chris Kyriakakis, [Automating the Detection of Breaks in Continuous User Experience with Computer Games](#), [Conference on Human Factors in Computing Systems CHI '05 extended abstracts on Human factors in computing systems](#), pp. 1629 - 1632, ISBN:1-59593-002-7, Portland, Oregon, USA, April 2005
- Kiyoung Yang and Cyrus Shahabi, [A Multilevel Distance-based Index Structure for Multivariate Time Series](#), [The 12th IEEE International Symposium on Temporal Representation and Reasoning \(TIME'05\)](#), pp. 65 - 73, ISSN: 1530-1311, Burlington, Vermont, USA, June 2005
- Hyunjin Yoon, Kiyoung Yang, and Cyrus Shahabi, [Feature Subset Selection and Feature Ranking for Multivariate Time Series](#), [IEEE Transactions on Knowledge and Data Engineering \(TKDE\) - Special Issue on Intelligent Data Preparation](#), Volume: 17, Issue: 9, pp. 1186- 1198, ISSN: 1041-4347, September 2005
- Cyrus Shahabi, [Wavelets and Web-Services \(W2\) for Online Scientific Data Analysis](#), , Lawrence Livermore National Laboratory, Livermore, CA, August 2005
- Kiyoung Yang, Hyunjin Yoon, and Cyrus Shahabi, [CLEVer: A Feature Subset Selection Technique for Multivariate Time Series](#), [The Ninth Pacific-Asia Conference on Knowledge Discovery and Data Mining \(PAKDD-05\)](#), LNAI, Volume 3518, pages 516-522, ISBN: 3-540-26076-5, Hanoi, Vietnam, May 2005
- Mehdi Sharifzadeh, Farnaz Azmoodeh, and Cyrus Shahabi, [Change Detection in Time Series Data Using Wavelet Footprints](#), [9th International Symposium on Spatial and Temporal Databases \(SSTD05\)](#), Angra dos Reis, Brazil, August 2005
- Kiyoung Yang, Hyunjin Yoon, and Cyrus Shahabi, [A Supervised Feature Subset Selection Technique for Multivariate Time Series](#), [International Workshop on Feature Selection for Data Mining: Interfacing Machine Learning with Statistics \(FSDM\) in conjunction with 2005 SIAM International Conference on Data Mining \(SDM'05\)](#), Newport Beach, CA, April 2005
- Tim Marsh, Wee Ling Wong, Eduardo Carriazo, Luciano Nocera, Kiyoung Yang, Aditya Varma, Hyunjin Yoon, Yih-lun Huang, Chris Kyriakakis, and Cyrus Shahabi, [User Experiences and Lessons Learned from](#)

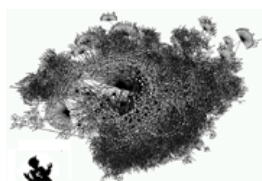


[Developing and Implementing an Immersive Game for the Science Classroom, The 11th International Conference on Human-Computer Interaction](#), Las Vegas, Nevada, USA, July 2005

- Kiyoun Yang, Hyunjin Yoon, and Cyrus Shahabi, [CLEVer: a Feature Subset Selection Technique for Multivariate Time Series](#), University of Southern California, Computer Science Department, Technical Report 05-845, March 2005
- Kiyoun Yang and Cyrus Shahabi, [A PCA-based Similarity Measure for Multivariate Time Series, The Second ACM International Workshop on Multimedia Databases \(ACM-MMDB 2004\)](#), pp 65 - 74 , ISBN:1-58113-975-6, Washington D.C., U.S.A., November 2004
- Tim Marsh, Kiyoun Yang, and Cyrus Shahabi, [Vicariouly there: connected with and through our own and other characters](#), [AISB Virtual Social Characters Symposium: Empathic Interaction Theme](#), University of Hertfordshire, Hatfield, UK, April 2005
- Margaret McLaughlin, Yi-Shin Chen, Namkee Park, Weirong Zhu, and Hyunjin Yoon, [Recognizing User State from Haptic Data, The 10th International Conference on Information Systems Analysis and Synthesis](#), Orlando, Florida, USA, July 2004
- J. Eisenstein, S. Ghandeharizadeh, L. Huang, C. Shahabi, G. Shanbhag, and R. Zimmermann, [Analysis of Clustering Techniques to Detect Hand Signs](#), [International Symposium on Intelligent Multimedia, Video and Speech Processing](#), Hong Kong, May 2001
- Cyrus Shahabi and Donghui Yan, [Real-Time Pattern Isolation And Recognition Over Immersive Sensor Data Streams](#), [The 9th International Conference On Multi-Media Modeling](#), Taiwan, January, 2003
- M. McLaughlin, G. Sukhatme, J. Hespanha, C. Shahabi, A. Ortega, and G. Medioni, [The Haptic Museum](#), Proc. EVA 2000 Conference on Electronic Imaging and the Visual Arts, Florence, Italy, 2000
- C. Shahabi, L. Kaghazian, S. Mehta, A. Ghoting, G. Shanbhag, and M. McLaughlin, [Analysis of Haptic Data for Sign Language Recognition](#), [9th International Conference on Human Computer Interaction](#), New Orleans, August 2001
- Cyrus Shahabi, Leila Kaghazian, Soham Mehta, Amol Ghoting, Gautam Shanbhag, and Margaret L. McLaughlin, [Understanding of User Behavior in Immersive Environments](#), [Touch in Virtual Environments: Haptics and the Design of Interactive Systems](#), ISBN 0-13-065097-8, Margaret L. McLaughlin, Joao Hespanha, and Gaurav Sukhatme, Editors. Prentice Hall, December, 2001
- Cyrus Shahabi, Mohammad R. Kolaoudouzan, Greg Barish, Roger Zimmerman, Didi Yao, and Kun Fu, [Alternative Techniques for the Efficient Acquisition of Haptic Data](#), [ACM SIGMETRICS Performance Evaluation Review Volume 29, Issue 1 \(June 2001\)](#), June 2001
- Cyrus Shahabi, Mohammad R. Kolaoudouzan, Greg Barish, Roger Zimmerman, Didi Yao, and Kun Fu, [Alternative Techniques for the Efficient Acquisition of Haptic Data](#), [ACM SIGMETRICS/Performace 2001](#), Cambridge, June 2001
- Cyrus Shahabi, Antonio Ortega, and Mohammad R. Kolaoudouzan, [A Comparison Of Different Haptic Compression Techniques](#), [IEEE International Conference on Multimedia and Expo \(ICME\)](#), Lausanne, Switzerland, August 2002

[Top](#)

## Peer-to-Peer Search



The next generation of distributed database systems is a family of massive self-organizing data networks; we term Querical Data Networks (QDNs). QDNs consist of a dynamic and large set of peer autonomous nodes, a dynamic dataset that is naturally distributed among the nodes in extra fine-grain, and a transient-form communication infrastructure that interconnects the nodes. Peer-to-peer and sensor networks are well-known examples of QDN. Distributed query processing within a QDN is far more challenging as compared to that of the traditional distributed database systems. Our approach is to categorize QDNs as instances of "Complex Systems" and apply the Complex System Theory (CST) to study QDNs. As our focus problem, we study CST-based mechanisms for efficient data location (search) within QDNs. The search primitive is one of the most fundamental functionalities vital to QDN query processing.

Some QDNs are structurable/indexible whereas some others, due to the extreme dynamism of the QDN topology and the extreme autonomy of the QDN nodes, render any attempt to even semi-structurize the network to an index-like structure inefficient and/or impossible. For structurable QDNs, we study self-organizing mechanisms to structurize the topology of the QDN to a search-efficient topology. Our approach is inspired by the "small-world" models that try to explain efficient communication in a social network, which is a semi-structured complex system. For unstructurable QDNs, we study efficient query flooding mechanisms. Efficient flooding can be applied as a method to disseminate broadcast search queries as well as other unicast and multicast queries. We employ percolation theory, an analytical tool borrowed from CST, to formalize and analyze these flooding mechanisms.

### People Involved

- Farnoush Banaei-Kashani

### Publications

- Farnoush Banaei Kashani and Cyrus Shahabi, [Applications of Sensor Network Data Management](#), Encyclopedia of Database Systems 2009, 107-111,
- Ali Khoshgozaran and Cyrus Shahabi, [Private Buddy Search: Enabling Private Spatial Queries in Social Networks](#), [Symposium on Social Intelligence and Networking \(SIN09\)](#), In conjunction with IEEE

[International Conference on Social Computing \(SocialCom09\)](#), Vancouver, Canada, August 2009

- F. Banaei-Kashani and C. Shahabi, [Case study: Scoop up data from peer-to-peer databases](#), Book chapter in Handbook of Peer-to-Peer Networking, Eds: X. Shen, H. Yu, J. Buford, M. Akon; Publishers:Springer, March 2009
- Farnoush Banaei-Kashani and Cyrus Shahabi, [Partial Read from Peer-to-Peer Databases](#), [Journal of Computer Communications](#), pp. 332-345, DOI information: 10.1016/j.comcom.2007.08.013, February 2008
- Cyrus Shahabi and Farnoush Banaei-Kashani, [Modelling Peer-to-Peer Data Networks under Complex System Theory](#), [International Journal of Computational Science and Engineering \(IJCSE\)](#), Vol. 2, No 5/6, September 2006
- Cyrus Shahabi and Farnoush Banaei-Kashani, [Querical Data Networks](#), [Encyclopedia of Database Technologies and Applications](#), ISBN: 1-59140-560-2, Editors: Laura C. Rivero, Jorge Horacio Doorn, Viviana E. Ferraggine; Publisher: Idea Group Inc., August 2005
- Farnoush Banaei-Kashani and Cyrus Shahabi, [Partial Selection Query in Peer-to-Peer Databases](#), [22nd International Conference on Data Engineering \(ICDE'06\)](#), (Short Paper), Atlanta, Georgia, USA, April 2006
- Cyrus Shahabi, [Modelling Peer-to-Peer Data Networks Under Complex System Theory](#), Ewha Woman University, Seoul, Korea, May 2005
- Cyrus Shahabi, [Modelling Peer-to-Peer Data Networks Under Complex System Theory](#), Seoul National University, Seoul, Korea, May 2005
- Cyrus Shahabi, [Modelling Peer-to-Peer Data Networks Under Complex System Theory](#), Korea University, Seoul, Korea, May 2005
- Farnoush Banaei-Kashani and Cyrus Shahabi, [SWAM: A Family of Access Methods for Similarity-Search in Peer-to-Peer Data Networks](#), [ACM Thirteenth Conference on Information and Knowledge Management \(ACM CIKM'04\)](#), pp 304-313, Washington DC, USA, November 2004
- Cyrus Shahabi, [Modelling Peer-to-Peer Data Networks Under Complex System Theory](#), [Department of Computer Science, National Tsing Hua University](#), HsingChu, Taiwan, March 2005
- Cyrus Shahabi, [Modelling Peer-to-Peer Data Networks Under Complex System Theory](#), [4th International Workshop on Databases in Networked Information Systems \(DNIS 2005\)](#), Aizu-Wakamatsu, Japan, March 2005
- Cyrus Shahabi and Farnoush Banaei-Kashani, [Modelling Peer-to-Peer Data Networks Under Complex System Theory \(Invited Paper\)](#), [Workshop on Databases in Networked Information Systems \(DNIS 2005\)](#), LNCS, Volume 3433, ISBN: 3-540-25361-0, Aizu-Wakamatsu, Japan, March 2005
- Farnoush Banaei-Kashani and Cyrus Shahabi, [Criticality-based Analysis and Design of Unstructured Peer-to-Peer Networks as Complex Systems](#), [Third International Workshop on Global and Peer-to-Peer Computing \(GP2PC\) in conjunction with CCGrid 2003](#), pp 351-358, Tokyo, Japan, May 2003
- Farnoush Banaei-Kashani and Cyrus Shahabi, [Brief Announcement: Efficient Flooding in Power-Law Networks](#), [Twenty-Second ACM Symposium on Principles of Distributed Computing \(PODC'03\)](#), page 336, Boston, USA, July 2003
- Farnoush Banaei-Kashani, Ching-Chien Chen, and Cyrus Shahabi, [WSPDS: Web Services Peer-to-peer Discovery Service](#), [International Symposium on Web Services and Applications \(ISWS'04\)](#), pp 733-743, Las Vegas, Nevada, USA, June 2004
- Farnoush Banaei-Kashani and Cyrus Shahabi, [Searchable Querical Data Networks](#), [International Workshop on Databases, Information Systems and Peer-to-Peer Computing in conjunction with VLDB2003](#), LNCS Vol. 2944, pp 17-32, Berlin, Germany, September, 2003

[Top](#)

## Web Services Discovery



The Web Services infrastructure is a distributed computing environment for service-sharing. In this environment, resource discovery is required as a primitive functionality for the service requesters to be able to locate the services, the shared resources. A discovery service with centralized architecture, such as UDDI (Universal Discovery Description and Integration), restricts the scalability of this environment as it grows to the scales comparable with the size of the web itself. In addition, current extensively used web service standards (e.g. UDDI, WSDL), do not support discovery at a semantic level. Hence, we propose a fully decentralized and interoperable discovery service with semantic-level search capability, termed p2p-WSD for peer-to-peer web-service discovery. The first decentralization characteristic of p2p-WSD is in synch with autonomy and distribution design objectives and its semantic-level matching capability is based on our knowledge representation design principles.

In sum, we believe a decentralized architecture of the semantic-enabled p2p-WSD not only satisfies the design requirements for efficient and accurate discovery in distributed environments, but also is compatible with the nature of the Web Services environment as a self-organized federation of peer service-providers without any particular sponsor.

### People Involved

- Dr. Ching-Chien (Jason) Chen
- hahzad Majeed Tiwana

## Publications

- Farnoush Banaei-Kashani, Ching-Chien Chen, and Cyrus Shahabi, [WSPDS: Web Services Peer-to-peer Discovery Service](#), [International Symposium on Web Services and Applications \(ISWS'04\)](#), pp 733-743, Las Vegas, Nevada, USA, June 2004
- Shahram Ghandeharizadeh, Craig A. Knoblock, Christos Papadopoulos, Cyrus Shahabi, Esam Alwagait, Jose Luis Ambite, Min Cai, Ching-Chien Chen, Parikshit Pol, Rolfe Schmidt, Saihong Song, Snehal Thakkar, and Runfang Zhou, [Proteus: A System for Dynamically Composing and Intelligently Executing Web Services](#), [The 2003 International Conference on Web Services \(ICWS'03\)](#), Las Vegas, Nevada, USA, June, 2003

[Top](#)

## Query Customization



As the demand for digital information and the amount of data dramatically increase, the task of acquiring accurate information has become more and more essential in various domains. In most of information-locating systems, providing a customized result set based upon features of data and user preferences is the ultimate objective. The main challenge is that both the data features and the user preferences are represented with several dimensions. Hence, aggregation, similarity matching, clustering, and classification of multidimensional data sets need to be studied.

We study these issues in the context of three application domains: 1) a content-based image retrieval system, 2) an e-commerce recommendation system, and 3) a neuroscience knowledge management system. We propose a customized query model, named Yoda, which can uniformly incorporate physical characteristics (data features), semantic meanings (attributes), and users' preferences during the query process. Moreover, Yoda also utilizes the users' relevance feedback to improve query results by using genetic algorithms (GA).

## People Involved

- Dr. Yi-Shin Chen

## Publications

- Ugur Demiryurek, Farnoush Banaei-Kashani, and Cyrus Shahabi, [Efficient K-Nearest Neighbor Search in Time-Dependent Spatial Networks](#), [21st International Conference on Database and Expert Systems Applications \(DEXA'10\)](#), Bilbao, Spain, August 2010
- Ali Khoshgozaran, Ali Khodaei, Mehdi Sharifzadeh, and Cyrus Shahabi, [A Multi-Resolution Compression Scheme for Efficient Window Queries over Road Network Databases](#), [Spatial and Spatio-temporal Data Mining \(SSTD\)](#), Hong Kong, December 2006
- Gully Burns, Arshad Khan, Shahram Ghandeharizadeh, Mark O'Neil, and Yi-Shin Chen, [Tools and Approaches for the Construction of Knowledge Models from the Neuroscientific Literature](#), [NeuroInformatics](#), Vol. 1, Nos. 1, Pages 81-110, January, 2003
- Yi-Shin Chen and Cyrus Shahabi, [Improving User Profiles for E-Commerce by Genetic Algorithms](#), [E-Commerce and Intelligent Methods Studies in Fuzziness and Soft Computing](#), Vol. 105, VIII, J. Segovia, P.S. Szczepaniak, M. Niedzwiedzinski (Eds.), 2002
- Yi-Shin Chen, Cyrus Shahabi, and Gully A.P.C. Burns, [Two-Phase Decision Fusion Based On User Preference](#), [The Hawaii International Conference on Computer Sciences](#), Honolulu, Hawaii, January, 2004
- Cyrus Shahabi and Yi-Shin Chen, [Efficient Support of Soft Query in Image Retrieval Systems](#), [International Conference on Advances in Infrastructure for Electronic Business, Science, and Education on the Internet](#), Rome, Italy, August, 2000
- Cyrus Shahabi and Yi-Shin Chen, [Soft Query in Image Retrieval Systems](#), [In Proceedings of the SPIE Internet Imaging \(EI14\), Electronic Imaging 2000, Science and Technology](#), San Jose, California, January, 2000
- Yi-Shin Chen and Cyrus Shahabi, [Automatically Improving the Accuracy of User Profiles with Genetic Algorithm](#), [IASTED International Conference on Artificial Intelligence and Soft Computing](#), Cancun, Mexico, May, 2001
- Cyrus Shahabi and Yi-Shin Chen, [A Unified Framework to Incorporate Soft Query into Image Retrieval Systems](#), [International Conference on Enterprise Information Systems](#), Setubal, Portugal, July, 2001
- Cyrus Shahabi, Farnoush Banaei-Kashani, Yi-Shin Chen, and Dennis McLeod, [Yoda: An Accurate and Scalable Web-based Recommendation System](#), [Sixth International Conference on Cooperative Information Systems \(CoopIS 2001\)](#), Trento, Italy, September, 2001
- Yi-Shin Chen and Cyrus Shahabi, [Yoda, An Adaptive Soft Classification Model: Content-based Similarity Queries and Beyond](#), [ACM/Springer Multimedia Systems Journal, Special Issue on Content-Based Retrieval](#), Vol. 8, Issue 6, Pages 523-535, 2003
- Cyrus Shahabi and Yi-Shin Chen, [An Adaptive Recommendation System without Explicit Acquisition of User Relevance Feedback](#), [Distributed and Parallel Databases Journal](#), Vol. 14, No 3, Pages 173-192, Kluwer Academic Publishers, September, 2003
- Cyrus Shahabi and Yi-Shin Chen, [Web Information Personalization: Challenges and Approaches](#), [The 3rd International Workshop on Databases in Networked Information Systems \(DNIS 2003\)](#), Aizu-

[Top](#)

## Continous Media Servers



Continuous media (CM) servers are unique in that they must provide continuous, real-time delivery of media with low start-up latency to clients. In addition, CM servers must be able to support multiple concurrent clients. We have designed and developed Yima, a multi-node CM server that is built with commodity components and is fully compatible with industry standards. Yima distinguishes itself from other CM servers in the following: 1) elimination of single point of failure with its multi-node shared nothing architecture, 2) frame/sample accurate synchronization of several streams, 3) independence from the media format, 4) network-failure tolerance with selective re-transmission, and 5) efficient flow control with multi-threshold buffering to support VBR traffic.

Currently we are equipping Yima with a disk scaling component, SCADDAR, that allows on-line addition and/or removal of disks with zero down-time to scale up/down the storage capacity adaptively as required. The random block-placement mechanism employed by SCADDAR ensures uniform distribution of blocks among disks (for balanced load), while the number of re-distributed blocks is also minimal.

### People Involved

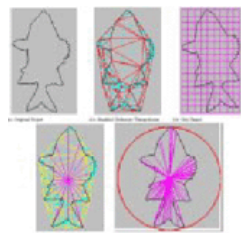
- Dr. Didi Yao
- Farnoush Banaei-Kashani
- Mehrdad Jahangiri

### Publications

- Shu-Yuen D. Yao, Cyrus Shahabi, and Roger Zimmermann, [BroadScale: Efficient scaling of heterogeneous storage systems](#), [International Journal on Digital Libraries](#), Volume 6, Number 1, Pages 98-111, ISSN: 1432-5012 (Paper) 1432-1300 (Online), Springer Berlin / Heidelberg, February 2006
- Roger Zimmermann, Cyrus Shahabi, Kun Fu, and Shu-Yuen Didi Yao, [Scalability evaluation of the Yima streaming media architecture](#), [Software: Practice and Experience](#), Volume 35, Issue 4, pp. 345 - 359, December 2004
- Shu-Yuen Didi Yao, Cyrus Shahabi, and Per-Ake Larson, [Hash-Based Labeling Techniques for Storage Scaling](#), [The VLDB Journal The International Journal on Very Large Data Bases by Springer-Verlag Heidelberg](#), ISSN: 1066-8888 (Paper) 0949-877X (Online), DOI: 10.1007/s00778-004-0124-6 , Issue: Online First, 2005
- Ali Dashti, Seon Ho Kim, Cyrus Shahabi, and Roger Zimmermann, [Streaming Media Server Design](#), , Prentice Hall PTR, 2003
- Roger Zimmermann, Kun Fu, Cyrus Shahabi, Didi Yao, and Hong Zhu, [Yima: Design and Evaluation of a Streaming Media System for Residential Broadband Services](#), VLDB 2001 Workshop on Databases in Telecommunications (DBTel 2001), Rome, Italy, September 2001
- Cyrus Shahabi, Roger Zimmermann, Kun Fu, and Shu-Yuen Didi Yao, [Yima: A Second Generation of Continuous Media Servers](#), IEEE Computer Magazine, Vol.35, No.6, Pages 56-64, 2002
- Ashish Goel, Cyrus Shahabi, Shu-Yuen Didi Yao, and Roger Zimmermann, [SCADDAR: An Efficient Randomized Technique to Reorganize Continuous Media Blocks](#), International Conference on Data Engineering, San Jose, California, February 2002
- Shahram Ghandeharizadeh, Seon Ho Kim, and Cyrus Shahabi, [On Disk Scheduling and Data Placement for Video Servers](#),

[Top](#)

## Shape Retrieval



We developed a new shape-based object retrieval technique based on minimum bounding circles (MBCs). Our MBC-based method can be used for efficient retrieval of 2D objects by utilizing three different index structures on features that are extracted from the object's MBC. Besides similarity retrieval, we also investigated how to support spatial queries (i.e. topological and direction queries) based on a subset of features extracted from the objects MBC. First, we focused on the support of topological relations as defined by the 9-intersection model. We proposed some filtering techniques based on special cases for MBC in order to reduce the number of false hits. Next, we proposed an extra filtering step to expedite direction relations between objects utilizing their MBCs based on cone-directions method. We evaluated and compared the performance of our method to different boundary based methods for shape representation and retrieval. The results showed that the similarity retrieval accuracy of our MBC-based method was better than other methods, while it has the lowest computation cost to generate the shape signatures of the objects. Moreover, it has low storage requirement, and a comparable computation cost to compute the similarity between two shape signatures. In addition, MBC-based method requires no normalization of the objects. Further more, our experiments showed the robustness of our method to the introduction of noise to the database and the query objects.

### People Involved

- Dr. Maytham Safar

### Publications



- Cyrus Shahabi and Maytham Safar, [An experimental study of alternative shape-based image retrieval techniques](#), *Multimedia Tools and Applications*, ISSN : 1380-7501 (Print) 1573-7721 (Online), Springer Netherlands, November 2006
- Cyrus Shahabi, [GeoDec: Enabling Geospatial Decision Making](#), ETRI, Korea, September 2006
- Maytham Safar and Cyrus Shahabi, [MBC-based shape retrieval: basics, optimizations, and open problems](#), *Multimedia Tools and Applications*, Volume 29, Number 2 , Pages 189-206, ISSN: 1380-7501 (Print) 1573-7721 (Online), Springer Netherlands, June 2006
- Maytham H. Safar and Cyrus Shahabi, [Shape Analysis and Retrieval of Multimedia Objects](#), , Kluwer Academic Publishers, 2002
- Cyrus Shahabi, Maytham Safar, and Hezhi Ai, [Multiple Index Structures for Efficient Retrieval of 2D Objects](#), IEEE ICDE99 (International Conference on Data Engineering)
- Maytham Safar and Cyrus Shahabi, [2D Topological and Direction Relations in the World of Minimum Bounding Circles](#), IDEAS 99, Montreal, Canada, August 1999
- Maytham Safar, Cyrus Shahabi, and Xiaoming Sun, [Image Retrieval By Shape: A Comparative Study](#), IEEE International Conference on Multimedia and Expo (ICME 2000), New York, July 2000
- Maytham Safar, Cyrus Shahabi, and Chung-hao Tan, [Resiliency and Robustness of Alternative Shape-Based Image Retrieval Techniques](#), International Database Engineering and Applications Symposium (IDEAS 2000), Yokohama, Japan, September 2000
- Maytham Safar and Cyrus Shahabi, [Two Optimization Techniques to Improve the Performance of MBC-Based Shape Retrieval](#), *Multimedia Information Systems (MIS 2000)*, Chicago, Illinois, October, 2000
- Maytham Safar and Cyrus Shahabi, [A Framework To Evaluate The Effectiveness of Shape Representation Techniques](#), *Journal of Applied Systems Studies, Special Issue on Distributed Multimedia Systems with Applications*, Vol. 2, No. 3, 2001

[Top](#)