BIOGRAPHICAL SKETCH

Provide the following information for the Senior/key personnel and other significant contributors.
Follow this format for each person. **DO NOT EXCEED FIVE PAGES.**

NAME: Shahabi, Cyrus

eRA COMMONS USER NAME (credential, e.g., agency login): shahabi

POSITION TITLE: Professor and Chair, Department of Computer Science -- Director, Integrated Media Systems Center (IMSC)

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable. Add/delete rows as necessary.)*

INSTITUTION AND LOCATION	DEGREE *(if applicable)*	Completion Date MM/YYYY	FIELD OF STUDY
Sharif University of Technology, Tehran, Iran	B.S.	05/89	Computer Engineering
Univ. of Southern California, Los Angeles, CA	M.S.	05/93	Computer Science
Univ. of Southern California, Los Angeles, CA	Ph.D.	05/96	Computer Science

A. Personal Statement

My research is in the area of information management. I have been PI on eleven and co-PI on two NSF grants. I have also received funding from a number of other federal agencies such as NGA, NIJ, NASA, NIH, DARPA, AFRL, and DHS as well as several industries such as Chevron, Google, HP, Intel, Microsoft, NCR, Oracle and NGC. My most relevant research to this effort includes my contributions to the field of spatiotemporal information management. Under this area, I have made contributions in several subareas, notably: Multivariate Time-series analytics, spatial indexing, participatory sensing, geo-social environments and most recently human performance monitoring for health applications. I was also one of the two co-founders of two spin-off companies: 1) Geosemble Technologies that was acquired in July 2012, and 2) ClearPath (rebranded later as TallyGo), which was axquired in March 2019.

Currently, in my role as the director of the Integrated Media Systems Center (IMSC), a graduated NSF Engineering Research Center at USC in the area of multimedia, I oversee several multidisciplinary projects spanning multiple schools at USC, such as Viterbi School of Engineering; Annenberg School for Communication & Journalism; Keck School of Medicine; School of Policy, Planning and Development; and Dornsife College of Letters, Art and Sciences. Some of these projects also include collaboration with industry partners (e.g., Microsoft, Oracle) and international institutes (e.g., National Informatics Institute of Japan).

Under my leadership, IMSC's focus has expanded to the field of data science by developing several BigData systems for real-world decision-making applications such as TransDec [1], GeoDec [2], ProDA [3], GeoSIM, MediaQ and Janus [4]. With each of these systems, we have addressed various major Data Science challenges, for example, multimedia data acquisition in MediaQ, sensor data collection and repository in TransDec, data analysis in ProDA, data integration in Janus, and multimedia data visualization in GeoDec and GeoSIM. Building on these experiences, we are well positioned to undertake the aims of this proposal, which are inline with IMSC's vision in the area of healthcare.

1. U. Demiryurek, F. Banaei-Kashani and C. Shahabi. TransDec: A Spatiotemporal query processing framework for transportation systems (demo paper). In IEEE ICDE, 2010.
2. C. Shahabi, F. Banaei-Kashani, A. Khoshgozaran, L. Nocera and S. Xing, GeoDec: A Framework to Effectively Visualize and Query Geospatial Data for Decision-Making, *IEEE Multimedia*, ISSN: 1070-986X, pp. 14-23, July-September 2010 (vol. 17 no. 3).

3. C. Shahabi, M. Jahangiri, and F. Banaei-Kashani, ProDA: An End-to-End Wavelet-Based OLAP System for Massive Datasets, IEEE Computer Magazine, Vol.41, No.4, ISSN: 0018-9162, April 2008.
4. C. Shahabi, S. Ho Kim, L. Nocera, G. Constantinou, Y. Lu, Y. Cai, G. Medioni, R. Nevatia and F. Banaei-Kashani, Janus - Multi Source Event Detection and Collection System for Effective Surveillance of Criminal Activity, in Journal of Information Processing System, 2014.

B. Positions, Scientific Appointments, and Honors

Positions and Employment

1996-2000	Research Assistant Professor, Department of Computer Science, USC.
1998-	Director of the Information Laboratory, USC.
2000-2004	Assistant Professor, Department of Computer Science, USC.
2001-2007	Research Area Director, Integrated Media Systems Center, USC.
2004-2012	Chief Technology Officer, Geosemble Technologies, Los Angeles, CA.
2004-2009	Associate Professor, Department of Computer Science, USC.
2009-	Professor, Department of Computer Science, USC.
2010-	Director, Integrated Media Systems Center, Viterbi School of Engineering, USC.
2012-	Professor, Department of Electrical Engineering, USC (joint appointment).
2012-2019	Chief Scientist & Founder, TallyGo, Los Angeles, CA.
2014-2018	Director, Informatics Program, Viterbi School of Engineering, USC.
2016-	Professor, Spatial Sciences, USC Dornsife College of Letters, Arts and Sciences.
2017-	Chair, Department of Computer Science, Viterbi School of Engineering, USC.
2019-	Helen N. and Emmett H. Jones Professor, Viterbi School of Engineering, USC.

Other Experience and Professional Memberships

1998-	Program Committee Member of VLDB, IEEE ICDE, ACM SIGMOD, ACMGIS, SSTD, ACM SIGKDD, IEEE ICDM, ACM Multimedia, ACM CIKM, etc.
1999	Founding Program chair of the ACM Workshop on WIDM
2004-	Member of the editorial board of ACM Computers in Entertainment magazine
2005	Founding Program chair of the IEEE Workshop (NetDB'05)
2005	Program Committee chair, ACM Workshop on Advances in GIS
2005-2009	Associate editor of IEEE Transactions on Parallel & Distributed Systems (TPDS)
2007	General co-chair of ACM Symposium on Advances in GIS
2008	General co-chair of ACM Conference on Advances in GIS
2009-	Associate editor of Journal of Spatial Information Science
2009-	Associate editor of Journal of Very Large Databases (VLDB)
2010-2013	Associate editor of IEEE Transactions on Knowledge & Data Engineering (TKDE)
2011	Chair of the ACM SIGSPATIAL Election Committee
2013-	Associate editor of ACM Transactions on Spatial Algorithms and Systems (TSAS)
2013	Program Committee co-Chair of the IEEE Mobile Data Management (MDM13).
2013	Program Committee co-Chair of the First IEEE Conference on Big Data 2013.
2013	Demonstration Chair of ACM SIGMOD 2013.
2015	Program Committee co-Chair of the DASFAA 2015 conference.
2015	General co-Chair of Symposium on Spatial & Temporal Databases 2015 (SSTD)
2016	Program Committee co-Chair of the 3rd International Conference on Big Data and Smart Computing 2016 (BigComp 16).
2016	Program Committee co-Chair of the IEEE Mobile Data Management (MDM'16).
2017	Program Committee co-Chair the Asia Pacific Web and Web-Age Information Management Joint Conference on Web and Big Data (APWeb-WAIM) 2017.
2017-2020	Chair of ACM SIGSPATIAL
2018-2021	PC Chair: SIGSPATIAL Spatial Gem Workshop
2020	Program Committee Vice-Chair: IEEE ICDE 2020.
2021	Senior PC Member: International Conference DASFAA 2021

Honors

2014	IEEE Fellow
2012-13	One of the eight recipients of the Vietnam Education Foundation (VEF) U.S. Faculty Scholar Grant to offer a course on geospatial information management (with hands-on mobile-application development projects) to two universities in Ho-chi-minh city, Vietnam.

2011-12	One of the four recipients of the Vietnam Education Foundation (VEF) U.S. Faculty Scholar Grant to offer a course on geospatial information management (with hands-on mobile-application development projects) to the Hanoi University of Science and Technology
2011	An organizer of the National Academy of Engineering "Japan-America Frontiers of Engineering" program
2010	An invited speaker in the National Research Council (of the National Academies) Committee on New Research Directions for the National Geospatial-Intelligence Agency
2009	Association of Computing Machinery (ACM) Distinguished Scientists Award
2005	Participant in National Academy of Engineering "Frontiers of Engineering" program
2003	U.S. Presidential Early Career Awards for Scientists and Engineers (PECASE)
2002	NSF CAREER award
2001	Okawa Foundation Research Award for Information and Telecommunications

C. Contributions to Science

I. My earliest research, in 1996, right after my graduation, pioneered a new area of research in web usage mining [1,2,3]. The focus of the work was to monitor and acquire users' web-navigations and then analyze that data to understand or predict users' behavior. This information can then be used for many commercial applications such as in recommendation systems, target advertising and content customization, to name a few. Currently, it is easy to see the importance of this research, but note that in 1996, the era before Google and Amazon, this was indeed a risky and pioneering research. In fact, my first paper in this area [1] was cited by the majority of the subsequent papers and patents in this area (cited 526 times according to Google Scholar).

1. C. Shahabi, A. Zarkesh, J. Adibi, and V. Shah, Knowledge discovery from users web-page navigation, In Proceedings of the IEEE Seventh International Workshop on Research Issues in Data Engineering (RIDE'97), April 1997.
2. A. Zarkesh, J. Adibi, C. Shahabi, R. Sadri and V. Shah, Analysis and Design of Server Informative WWW-sites, In Proceedings of the ACM Sixth International Conference on Information and Knowledge Management (CIKM'97), November 1997.
3. C. Shahabi and F. Banaei-Kashani, Efficient and Anonymous Web Usage Mining for Web Personalization, INFORMS Journal of Computing, Special Issue on Web Data Mining, v15, n2, pages 123-147, Spring 2003.

II. However, I am most well-known for my work in the area of spatial databases. The focus of this topic is to design and develop special purpose data structures and algorithms to efficiently index and access real-world geospatial data. The main challenge is to index the location and/or proximity of objects in real-world spaces such as road-networks and/or land surfaces as the objects (e.g., cars) move in this space and as the distances (e.g., time to travel) changes dynamically (e.g., due to variation of traffic flow). The purpose of these index structures is to expedite access to large (i.e., disk resident) real-world datasets for sophisticated queries, which require a complementary effort in developing query processing algorithms to exploit the designed index structures effectively. Our major accomplishment in the first topic has been the design of a series of indexing techniques based on Voronoi Diagrams [1,2] that can efficiently index objects in road-networks and land surfaces. In the process, we have introduced a new type of spatial queries, termed *spatial skyline [3]*, which was well received by the community. Our first paper proposing a Voronoi-based index structure for road-networks appeared in 2004 [4] with 592 citations so far.

1. M. Sharifzadeh and C.Shahabi, Approximate Voronoi Cell Computation on Spatial Data Streams, In The VLDB Journal: The International Journal on Very Large Data Bases, volume 18, n1, pp57-75, 2009.
2. A. Akdogan, U. Demiryurek, F. Banaei-Kashani and C. Shahabi, Voronoi-based Geospatial Query Processing with MapReduce, The 2nd IEEE International Conference on Cloud Computing Technology and Science (CloudCom 2010), Indianapolis, Indiana, December 2010 (Best Paper Award).
3. M. Sharifzadeh and C. Shahabi, The Spatial Skyline Queries, 32nd International Conference on Very Large Data Bases (VLDB'06), Seoul, Korea, September 2006.
4. M. R. Kolahdouzan and C. Shahabi, Voronoi-Based K Nearest Neighbor Search for Spatial Network Databases, 30th International Conference on Very Large Data Bases (VLDB'04), Toronto, Canada, September 2004.

III. I have also been involved in entrepreneurial activities resulted from one of my research topics. The focus of this topic is to devise automatic techniques to accurately align various geospatial data sources such as aerial imagery, maps and road vector data with each other to generate new fused geospatial products that are more useful than each of the individual components. The challenge here is that to align one source to the other, one cannot apply a global transformation on one source and instead multiple local transformations are needed. Consequently, the process requires the human in the loop to identify several matching features across the sources to which local transformations can apply. Our major accomplishments include a series of algorithms that can automatically align vector data and raster map data with aerial imagery. These algorithms use the knowledge from vector data (e.g., road intersections) to both narrow the search space and focus the search process in the aerial imagery to find the matching features across the two sources. We have not only published several papers on this topic but also were awarded several patents [1,2,3,4]. This last set of technologies has been transferred to industry through a company startup, called Geosemble Technologies, in 2004. I was the CTO and co-founder of Geosemble Technologies. The company provides automatic techniques for extracting and fusing geospatial data sets, including maps, aerial imagery, lists, events, databases, businesses and more. In April 2009, Geosemble received a strategic investment, technology development program, and licensing agreement with In-Q-Tel, the independent strategic investment firm that identifies innovative technology solutions to support the mission of the broader U.S. Intelligence Community. Geosemble has been acquired in July 2012 by another In-Q-Tel portfolio company, Terrago Technologies.

1. Processing time-based geospatial data: with Ching-Chien Chen and Craig A. Knoblock, Patent No. 8,953,887, United States, February 10, 2015.
2. Precisely locating features on geospatial imagery; with Ching-Chien Chen, Dipsy Kapoor, and Craig A. Knoblock, Patent No. 8,675,995, United States, March 2014.
3. Systems and methods for linking content to individual image features; with Craig A. Knoblock, Ching-Chien Chen, and Dipsy Kapoor, Patent No. 8,670,617, United States, March 2014.
4. Dynamically linking relevant documents to regions of interest; with Craig A. Knoblock, Dipsy Kapoor, and Ching-Chien Chen, Patent No. 8,635,228, issued in Jan 2014.

IV. One of my more recent work is the introduction of a new area of research called Spatial Crowdsourcing [1]. Spatial crowdsourcing engages individuals, groups, and communities in the act of collecting, analyzing, and disseminating urban, social, and other spatiotemporal information. This new paradigm of data collection has shown to be useful when traditional means fail (e.g., due to disaster), are censored or do not scale in time and space. The wide applicability of spatial crowdsourcing primarily became possible due to the broad availability of mobile devices. With spatial crowdsourcing, the goal is to efficiently outsource a set of spatiotemporal tasks (i.e., tasks related to time and location) to a set of workers, which requires the workers to perform the tasks by physically traveling to those locations. Hence, spatial crowdsourcing strategies must be designed to take advantage of large population of human workers for ad-hoc spatiotemporal tasks -- they must consider the environment's dynamism (i.e., tasks and workers come and go) and scale [2] as well as users' considerations such as trust (i.e., not all workers are trustworthy) [3] and privacy [4] (i.e., not all workers want to share their location information).

1. L. Kazemi and C. Shahabi, GeoCrowd: Enabling Query Answering with Spatial Crowdsourcing, ACM SIGSPATIAL GIS (ACMGIS 2012), Redondo Beach, CA, November 2012.
2. A. Alfarrarjeh, T. Emrich and Cyrus Shahabi, Scalable Spatial Crowdsourcing: A study of distributed algorithms, 16th IEEE International Conference on Mobile Data Management, Pittsburgh, PA, June 2015.
3. L. Kazemi, C. Shahabi and L. Chen, GeoTruCrowd: Trustworthy Query Answering with Spatial Crowdsourcing, ACM SIGSPATIAL GIS (ACMGIS 2013), Orlando, Florida, November 2013.
4. H. To, G. Ghinita, and C. Shahabi, A Framework for Protecting Worker Location Privacy in Spatial Crowdsourcing, Proceedings of VLDB 7(10): 919-930, August 2014.

V. Most recently, my research has focused on improving people's data privacy using privacy-preserving machine learning approaches. Privacy attacks occur more frequently and with far greater ramifications than before, as illustrated by recent high-profile privacy breaches that affected prominent players in the social media industry, including Google, Facebook and Yahoo. My recent work focuses on protecting data privacy related to the whereabouts of users. While mobile apps allow users to enjoy personalized services and receive information tailored to their geographical location—such as finding local interests, connecting with nearby friends or requesting rideshare services—sharing location data with service providers could lead to privacy concerns. For

instance, specific movement details may leak sensitive information about an individual's health status, political orientation or entertainment preferences. In one study [1] we investigate a technique that employs an encryption scheme, so the data can only be read if decrypted, which significantly increases data security. In another study [2], the goal is to prevent an online service from learning fine-grained user locations, while still allowing it to provide services based on approximate location information. The user can add controlled random noise to perturb their position before sending it to the location-centric application. As such, the user's locations never leave their mobile phones unprotected. Lastly, we consider what happens when service providers use the vast amounts of sensitive trajectory data to train machine learning models [3]. We introduce data management techniques and careful tuning of the model to ensure that the private model's utility remains high, while leaving out the esoteric, unique details of users in the training data.

1. S. Shaham, G. Ghinita, and C. Shahabi: An Efficient and Secure Location-based Alert Protocol using Searchable Encryption and Huffman Codes, Proceedings of the 24th International Conference on Extending Database Technology (EDBT 2021), Nicosia, Cyprus, March 2021.
2. Ritesh Ahuja, Gabriel Ghinita, and C. Shahabi: A Utility-Preserving and Scalable Technique for Protecting Location Data with Geo-Indistinguishability, Proceedings of the 22nd International Conference on Extending Database Technology (EDBT 2019), Lisbon, Portugal, March 2019.
3. R. Ahuja, G. Ghinita, and C. Shahabi, Differentially-Private Next-Location Prediction with Neural Networks, Proceedings of the 23rd International Conference on Extending Database Technology (EDBT 2020), Copenhagen, Denmark, March 2020.

Complete List of Published Work in DBLP: http://dblp.uni-trier.de/pers/hd/s/Shahabi:Cyrus

D. Additional Information: Research Support and/or Scholastic Performance

Selected Ongoing Research Support

CNS-2027794 (Shahabi) 05/2020 – 04/2022 0.20 Summer
National Science Foundation $83,185
RAPID: REACT: Real-time Contact Tracing and Risk Monitoring via Privacy-enhanced Mobile Tracking
The major goal of the project is to develop techniques for real time contact tracing and risk monitoring via privacy-enhanced mobility tracking.
Role: PI

IIS-1910950 (Shahabi) 09/2019 – 10/2022 0.25 Summer
National Science Foundation $306,814
II: Small: Collaborative Research: PE4GQ - Practical Encryption for Geospatial Queries on Private Data
The project's objective is to build a framework for secure and efficient processing of spatial queries using searchable encryption.
Role: PI

PS36665000 (Giuliano) 07/2017-12/2021 0.60 Summer
Los Angeles County Service Authority for Freeway Emergencies $587,624
Archived Data Management System Research and Technical Support Services
The goal of this project is to maintain the Advanced Data Management System (ADMS), which we developed for traffic and transit data for storing, querying and analysis of the data in real-time.
Role: Co-PI

Selected Completed Research Support

CNS- 1461963 (Shahabi) 04/01/2015-03/31/2019 0.50 Summer
National Science Foundation $ 299,976
BDD: Human-Centered Situation Awareness Platform for Disaster Response and Recovery
The goal was to develop a spatial crowdsourcing method for efficient visual situation awareness in disasters.

RFA-ES-15-004 (Ambite/Gilliland) 10/01/15-09/30/19 0.9 Academic
NIH $999,394
Pediatric Research using Integrated Sensor Monitoring Systems (PRISMS): Data and Software Coordination and Integration Center (DSCIC)
The major goals of the PRISMS-DSCIC are (1) to integrate environmental, physiological, and behavioral data produced by the sensor-based health monitoring systems funded under PRISMS and data from relevant external sources, and (2) facilitate access, analysis, and dissemination of such data and software generated by PRISMS.