OMB No. 0925-0001 and 0925-0002 (Rev. 09/17 Approved Through 03/31/2020)

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: Shashi Shekhar

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

POSITION TITLE: McKnight Distinguished University Professor

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 |
| --- | --- | --- | --- |
| Indian Inst. Of Tech., Kanpur, India | B.S. | 1985 | Computer Science |
| University of California, Berkeley | M.S. | 1987 | Computer Science |
| University of California, Berkeley | M.S. | 1989 | Business Administration |
| University of California, Berkeley | Ph.D. | 1990 | Computer Science |

**A. Personal Statement**

I am a member of the Screening, Prevention, Etiology and Cancer Survivorship research area at the Masonic Cancer Center. I have a broad background in Computer Science, with specific training and expertise in spatial databases, spatial data mining and Geographic Information Science. A key research focus is on *geo-spatial approaches to cancer control* including geo-spatial surveillance. We have active collaborations with American Cancer Society (Dr. Liora Sahar, Strategic Director, Geospatial Research) to evaluate methods for identifying geographic clusters of cancer and their spatial relationship with screening centers and rurality. Within the Masonic center, we are collaborating with Prof. Logan Spector (Division Director, Epidemiology and Clinical Research) for understanding spatial interactions between childhood cancers and birth defects. As PI or co-Investigator on several university-, NSF-, USDOD- and NIH-funded grants, I laid the groundwork for the proposed research by developing effective measures and scalable algorithms to analyze spatial and spatio-temporal data for interesting, novel and non-trivial patterns such as statistically significant clusters, hotspots, colocations, spatial anomalies in context of societal applications such as public health, public safety, emergency evacuation and transportation. Many of my previous experiences are interdisciplinary and I am aware of the importance of communication among project members from different disciplines and I know how to construct a realistic research plan, timeline, and budget in a collaborative research project. In summary, I have the expertise, leadership, training, and motivation necessary to successfully carry out the proposed research project.

**B. Positions and Honors**

## Positions and Employment

1989-1995 Assistant Professor, University of Minnesota, Minneapolis, MN

1995-2000 Associate Professor, University of Minnesota, Minneapolis, MN

2001-present Professor, University of Minnesota, Minneapolis, MN

2005 – 2007 Director, Army High Performance Computing Research Center, University of Minnesota

2005-present Distinguished McKnight Univ. Professor, University of Minnesota, Minneapolis, MN

2015-present Distinguished Teaching Professor, University of Minnesota, Minneapolis, MN

## Other Experience and Professional Memberships

## 2017 – 2018 President, University Consortium for Geographic Information Science (UCGIS)

## 2016 - present Member, Masonic Cancer Center, University of Minnesota.

## 2016 Presented a talk in the Conference on Geospatial Approaches to Cancer Control and Population Sciences, September 12-14, NIH National Cancer Institute, Bethesda, MD, 2016.

2016-2019 Chair, Board of Directors, Symp. on Spatial and Temporal Databases (SSTD) Endowment.

2016-2019 Member, Board of Directors, Computing Research Association (CRA).

2015-2016 Member, National Academies Committees on “From Maps to Models:

Augmenting Nation’s Geospatial-Intelligence Capabilities”

## 2013 Distinguished Colloquium on “From GPS and Virtual Earth to Spatial Computing 2020”, Division of Cancer Control and Population Sciences, NIH National Cancer Institute, Dec. 12th, 2013.

## 2012-2013 Invited speaker in three NIH-AAG workshops on Geospatial Frontiers in Health and Social Environment: Towards a Unified Vision, National Inst. of Health & Am. Asso. of Geographers.

2012-2015 Member, CRA Computing Community Consortium Council.

2012-2013 Member, National Academies Committee on Geo-targeted Alerts and Warning,

2011-2012 Member, National Academies’ Committee on Future Workforce for Geospatial Intelligence.

2004-2005 Member, National Academies) Committee to review basic and applied research at the

USDOD National Geo-spatial Intelligence Agency (NGA).

2003-2009 Member, Mapping Sciences Committee, National Academy of Sciences.

## Honors

2015 Education Award, University Consortium on Geographic Information Science

2015 Post-baccalaureate, Graduate and Professional Teaching Award, University of Minnesota

2011 Elected a Fellow of the Institute on Environment, University of Minnesota

2006 Elected an AAAS Fellow for distinguished research, service and teaching contributions to the advancement of science in the fields of spatial databases, spatial data mining and GIS.

2006 Awarded the IEEE Computer Society Technical Achievement Award for distinguished contributions to research, teaching and service in spatial databases.

2005 McKnight Distinguished University Professorship, University of Minnesota

2003 Elected an IEEE Fellow for contributions to spatial database storage methods, data mining, and geographic information systems.

**C. Contributions to Science**

1. **Cluster Analysis:** Our group has experience in cluster analysis, particularly hotspots that are geographic areas with large number of events. Further, the group has extended SaTScan style statistically significant circular hotspots to statistically significant linear, ring shaped and elliptical hotspots. For example, we have conducted case study on a New York Bronx Legionnaires’ diseases outbreak dataset, and found ring-shaped hotspots to be effective in identifying the most relevant water tank (Reference a. below).
   1. *Significant Linear Hotspot Discovery, IEEE Transactions on Big Data, 3(2), pp.140-153, 2017 (w/ Xun Tang, Emre Eftelioglu, Dev Oliver).*
   2. *Detecting Isodistance Hotspots on Spatial Networks: A Summary of Results, International Symposium on Spatial and Temporal Databases (selectivity: 30%),* Springer LNCS: 10411, *pp. 281-299, 2017 (w/ Xun Tang, Emre Eftelioglu).*
   3. *Ring-Shaped Hotspot Detection*, IEEE Transactions on Knowledge and Data Engineering, 28(12):3367-3381, 2016 (w/ Emre Eftelioglu, James Kang, Christopher Farah).
   4. *Mining Network Hotspots with Holes: A Summary of Results, International Conference on Geographic Information Science (selectivity: 27%),* Springer LNCS: 9927,*pp. 51-67, 2016. (w/ Emre Eftelioglu, Yan Li, Xun Tang, James Kang, Christopher Farah)*
2. **Colocation Detection:** Our research group pioneered colocation patterns, which represent subsets of event-types that frequently occur together in geographic space. For example, lung cancer cases in 1950s often occurred near the ship building factories possibly due to use of asbestos. We proposed novel interest measures and algorithms.
   1. Cascading spatio-temporal pattern discovery. IEEE Transactions on Knowledge and Data Engineering, 24(11), 2012, pp 1977-1992 (w/ Pradeep Mohan, James A. Shine, James P. Rogers).
   2. [*Mixed-drove spatiotemporal co-occurrence pattern mining*](http://ieeexplore.ieee.org/abstract/document/4522550/)*,* IEEE Transactions on Knowledge and Data Engineering, 2008, 20(10), (1322-1335) (w/ Mete Celik, James P Rogers, James A. Shine)
   3. *Discovering Colocation Patterns from Spatial Data Sets: A General Approach*, IEEE Transactions on Knowledge and Data Engineering, 16(12), pp. 1472-1485, 2004 (with Yan Huang et al.).
   4. *Discovering spatial co-location patterns: A summary of results,* Symposium on Spatial and Temporal Databases (selectivity 30%), Springer LNCS: 2121, pp. 236-256. 2001 (w/ Yan Huang).
3. **Spatial Data Mining:** Beyond hotspot and colocation, my research has also explored other spatial pattern families in the broad area of spatial data mining, which is concerned with detection of useful, interesting, and non-trivial patterns from geospatial data. Examples include spatial outlier detection and spatial decision tree learning.
   1. *Transdisciplinary Foundations of Geospatial Data Science, ISPRS International Journal of Geo-Informatics, 6(12), 2017. doi:10.3390/ijgi6120395. (with Y. Xie, E. Eftelioglu, R. Ali, X. Tang, Y. Li, and R. Doshi).*
   2. *Spatiotemporal Data Mining: A Computational Perspective*, Special Issue on Advances in Spatio-Temporal Data Analysis and Mining, ISPRS International Journal of Geo-Information, MDPI, 4(4), 2306-2338, 2015. (w/ Zhe Jiang, Reem Y. Ali, Emre Eftelioglu, Xun Tang, Venkata M. V. Gunturi, and Xun Zhou)
   3. *Focal-test-based spatial decision tree learning*, IEEE Transactions on Knowledge and Data Engineering, 27(6), pp.1547-1559, 2015 (w/ Zhe Jiang, Xun Zhou, Joseph Knight, and Jennifer Corcoran.).
   4. *Identifying patterns in spatial information: A survey of methods*, Wiley Interdisciplinary Review: Data Mining and Knowledge Discovery 1(3), 2011. (w/ Michael R. Evans, James M. Kang, Pradeep Mohan).
4. **Spatial Computing:** Beyond spatial data mining, my work has also explored spatial databases and routing algorithms.
5. *Encyclopedia of GIS*, 2nd Ed. Springer, 2017, (w/ Hui Xiong, Xun Zhou).
6. *Spatial Computing*, Communications of the ACM (Association for Computing Machinery), cover article, 59(1), Jan. 2016 (w/ Steven K. Feiner, Walid G. Aref).
7. *Capacity-constrained network-voronoi diagram*, IEEE Transactions on Knowledge and Data Engineering, 27(11), pp.2919-2932, 2015 (w/ Kwangsoo Yang, Apurv Hirsh Shekhar, Dev Oliver).
8. *Spatial Databases: A Tour*, Prentice Hall, 2003 (w/ S. Chawla).

## Complete List of Published Work in My Bibliography:

<http://www.cs.umn.edu/~shekhar/bio/2016/fullreport.pdf>

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

## Ongoing Research Support

USDA Current Research Information System (2017-51181-27222) 09/2017 – 08/2021

*Increasing low-input turfgrass adoption through breeding, innovation, and public education*: The long-term goal of this project is to increase the use of well-adapted fine fescue cultivars in sustainable landscapes by the provision of new tools for consumers to use when making grass seed purchasing decisions, new turfgrass seed cultivars with improved low-input adaptation, new knowledge about the stress tolerance of fine fescues, new tools for public and private plant breeders to use when selecting fine fescues.

## Role: Co-Principal Investigator

NSF Computer and Network Systems (1737633) 08/2017 – 09/2020

*Connecting the Smart-City Paradigm with a Sustainable Urban Infrastructure Systems Framework to Advance Equity in Communities*: This project will investigate a smart urban infrastructure systems framework for advancing access and wellbeing in cities. With transformative new infrastructures (e.g., smart electricity grid, urban farms) on the horizon, this research will provide new perspectives on how the future spatial deployment of these new infrastructures in cities will shape wellbeing, health, and the environmental sustainability of outcomes in the different areas of cities.

Role: Principal Investigator

ARPA-E NEXTCAR 02/2017 – 02/2020

*Cloud-Connected Delivery Vehicles: Boosting Fuel Economy Using Physics-Aware Spatio-temporal Data Analysis and Real-Time Powertrain Control*: The project aims to develop technology to improve the fuel efficiency of delivery vehicles through real-time vehicle dynamic and powertrain control optimization using two-way vehicle-to-cloud (V2C) connectivity. The project developments aim to enable at least an additional 20% reduction in energy consumption of future connected and automated delivery vehicles, highly efficient domestic vehicle fleet, reduction in sector emissions, help in the improvement of urban air quality and decreasing the sector's carbon footprint.

Role: Co-Principal Investigator

## NSF CISE/Expeditions 08/2010 – 12/2018

## *Expedition: Understanding Climate Change: A Data Driven Approach*: This project addresses key challenges in the science of climate change by developing methods that take advantage of the wealth of climate and ecosystem data available from satellite and ground-based sensors, the observational record for atmospheric, oceanic, and terrestrial processes, and physics-based climate model simulations.

## Role: Senior Personnel, Investigator

## NSF OCI/Datanet 10/2011 – 09/2018

## *Datanet: Terra Populus: A Global Population Environment Data Network:* Terra Populus will integrate the world’s population and environmental data, including population censuses and surveys; land cover information from remote sensing; climate records from weather stations; and land use records from statistical agencies.

## Data that are interoperable across time, space, and scientific domain will allow us to understand the dramatic transformation of the earth’s inhabitants and their environment. This infrastructure will make it easier for researchers to use data describing people along with data describing the places they inhabit.

## Role: Co-Principal Investigator

## NSF CISE/IIS/III 09/2012 – 08/2018

## *III: Small: Towards Spatial Database Management Systems for Flash Memory Storage*: The goal of this research project is to design and develop highly efficient spatial and spatio-temporal database systems on flash memory storage. Specific aims include investigation of efficient spatial indexing on flash memory, efficient spatial query processing and optimization, spatio-temporal indexing and querying, storage hierarchy of flash and magnetic disks to exploit the full potential of both storage media.

## Role: Co-Principal Investigator

## USDOD NGA (HM0210-13-1-0005) 05/2013 – 05/2018

## *Identifying and Analyzing Patterns of Evasion*: The goal of this project is to investigate novel and computationally efficient methods for quantifying and mining spatio-temporal patterns of evasion.

## Role: Principal Investigator

## NSF CISE/IIS/III (IIS-1320580) 08/2013 – 08/2018

## *III: Small: Investigating Spatial Big Data for Next Generation Routing Services*: The goal of this project is to study emerging spatial big datasets (e.g., GPS track, engine measurements, temporally detailed roadmaps) to rank alternative routes on greenhouse gas emissions and fuel consumption.

## Role: Principal Investigator

## Completed Research Support

<http://www.cs.umn.edu/~shekhar/bio/2016/fullreport.pdf>