UNIVERSITY OF CENTRAL FLORIDA
CENTER FOR RESEARCH IN COMPUTER VISION

Dr. Shashi Shekhar

University of Minnesota

"What is special about mining spatial and spatio-temporal datasets?"

Friday, March 21, 2014 · 2:00PM · HEC 101

ABSTRACT

The importance of spatial and spatio-temporal data mining is growing with the increasing incidence and importance of large datasets such as maps, virtual globes, repositories of remote-sensing images, the decennial census and collections of trajectories (e.g. gps-tracks). Applications include Environment and Climate (global change, land-use classification), Public Health (e.g. monitoring and predicting spread of disease), Public Safety (e.g. crime hot spots), Public Security (e.g. common operational picture), M(obile)-commerce (e.g. location-based services), etc.

Classical data mining techniques often perform poorly when applied to spatial and spatio-temporal data sets because of the many reasons. First, these dataset are embedded in continuous space, whereas classical datasets (e.g. transactions) are often discrete. Second, patterns are often local whereas classical data mining techniques often focus on global patterns. Finally, one of the common assumptions in classical statistical analysis is that data samples are independently generated. When it comes to the analysis of spatial and spatio-temporal data, however, the assumption about the independence of samples is generally false because such data tends to be highly self-correlated. For example, people with similar characteristics, occupation and background tend to cluster together in the same neighborhoods. In spatial statistics this tendency is called autocorrelation. Ignoring autocorrelation when analyzing data with spatial and spatio-temporal characteristics may produce hypotheses or models that are inaccurate or inconsistent with the data set.

Thus new methods are needed to analyze spatial and spatio-temporal data to interesting, useful and non-trivial patterns. This talk surveys some of the new methods including those for discovering interactions (e.g. co-locations , co-occurrences, tele-connections), detecting spatial outliers and location prediction along with emerging ideas on spatio-temporal pattern mining.

BIOGRAPHY

Shashi Shekhar is a Mcknight Distinguished University Professor at the University of Minnesota (Computer Science faculty). For contributions to geographic information systems (GIS), spatial databases, and spatial data mining, he received the IEEE-CS Technical Achievement Award and was elected an IEEE Fellow as well as an AAAS Fellow. He was also named a key difference-maker for the field of GIS by the most popular GIS textbook. He has a distinguished academic record that includes 280+ refereed papers, a popular textbook on Spatial Databases (Prentice Hall, 2003) and an authoritative Encyclopedia of GIS (Springer, 2008). Shashi is serving as a member of the Computing Community Consortium Council (2012-15), a co-Editor-in-Chief of Geo-Informatica: An International Journal on Advances in Computer Sciences for GIS (Springer), a series editor for the Springer-Briefs on GIS, and as a member of the National Research Council (NRC) committee on Geo-targeted Disaster Alerts and Warning (2013). Earlier, he served on multiple NRC committees including Future Workforce for Geospatial Intelligence (2011), Mapping Sciences (2004-2009) and Priorities for GEOINT Research (2004-2005). He also served as a general or program co-chair for the Intl. Conference on Geographic Information Science (2012), the Intl. Symposium on Spatial and Temporal Databases (2011) and ACM Intl. Conf. on Geographic Information Systems (1996). He also served on the Board of Directors of University Consortium on GIS (2003-4), as well as the editorial boards of IEEE Transactions on Knowledge and Data Eng. and IEEE-CS Computer Sc. & Eng. Practice Board. In early 1990s, Shashi's research developed core technologies behind in-vehicle navigation devices as well as web-based routing services, which revolutionized outdoor navigation in urban environment in the last decade. His recent research results played a critical role in evacuation route planning for homeland security and received multiple recognitions including the CTS Partnership Award for significant impact on transportation. He pioneered the research area of spatial data mining via pattern families (e.g. collocation, mixed-drove co-occurrence, cascade), keynote speeches, survey papers and workshop organization. Shashi received a Ph.D. degree in Computer Science from the University of California (Berkeley, CA). More details are available from http://www.cs.umn.edu/~shekhar.