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 where as 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.