

MODELING URBAN BEHAVIOR BY MINING GEOTAGGED SOCIAL DATA

Abstract:

Data generated on location-based social networks provide rich information on the whereabouts of urban dwellers. Specifically, such data reveal who spends time where, when, and on what type of activity (e.g., shopping at a mall, or dining at a restaurant). That information can, in turn, be used to describe city regions in terms of activity that takes place therein. For example, the data might reveal that citizens visit one region mainly for shopping in the morning, while another for dining in the evening. Furthermore, once such a description is available, one can ask more elaborate questions. For example, one might ask what features distinguish one region from another – some regions might be different in terms of the type of venues they host and others in terms of the visitors they attract. As another example, one might ask which regions are similar across cities. In this paper, we present a method to answer such questions using publicly shared Foursquare data. Our analysis makes use of a probabilistic model, the features of which include the exact location of activity, the users who participate in the activity, as well as the time of the day and day of week the activity takes place. Compared to previous approaches to similar tasks, our probabilistic modeling approach allows us to make minimal assumptions about the data – which relieves us from having to set arbitrary parameters in our analysis.