Statement of Research Interests

Junjun Yin

As a GIScience researcher, my research interests center on computational geography to capture and understand the geospatial complexity in human-urban environment interactions. I value computational geography not just as another approach but a new data science paradigm driven by spatial thinking and geospatial Big Data. I focus on *geospatial Big Data analytics* to model urban dynamics concerning urban mobility, accessibility, and sustainability. I am fascinated by different interplays between human/citizen activities and spatial settings of the urban structures. Especially, I look into spatial interactions through the lens of human mobility patterns and its connections to dynamic urban structures, with research questions such as whether urban mobility is shaped by underlying urban structures or the real-world human interactions reveal a complete image of a city (Lynch 1960). To reach the research goals, my research attempts to utilize emerging methodologies, including advanced spatial data mining techniques, geovisual-analytics methods, and high-performance computing approaches, to delineate urban geography with geospatial Big Data. I am particularly interested in mining and analyzing *large-scale mobility data* (e.g., geo-located tweets and check-ins extracted from Location Based Social Media, taxi trip records, and GPS logs) and developing new *geovisual-analytics* methods in my research agenda. The remainder of this document illustrates the details of my current research and my vision to extend it for future research.

Current Research: *Spatial Big Data for Urban Studies: Through the lens of human mobility to understand urban dynamics*

Urban environments are real-world complex systems, where the embedded dynamics are interrelated to various types of human activities. To gain insights on the human-urban environment interactions at a fine spatial scale requires the synthesis of timely updated data that cover various aspects of the city, especially the detailed information related to human activities. My current research attempts to exploit human mobility patterns from large-volume and fine-grained movement data to better model and understand how citizens interact with their surrounding urban environment. It aims to address research questions on whether the spatial configurations of urban environment confine human activities or the social dynamics embedded in human activities change the functions of urban regions and therefore define the urban structures. Connecting mobility research and its applications to understand dynamic urban features takes a data driven strategy, which is essentially a “bottom-up” approach. It aims to address the methodological issues associated with partitioning urban and human activity space. My investigations of the spatial interaction behaviors of individuals reveal spatiotemporal patterns of human activities and uncover some interesting urban geography delineated by the collective human activities. The developed methods further assist to measure environmental exposures and the accessibility to resources for individuals. These efforts lead to applications for evaluating various neighborhood effects on public health. For instance, the methods are able to integrate geographic context to the GPS logs tracking individual movements across space and measure the environmental exposures during the course of daily activities.

Geospatial Big Data Synthesis: A GIS platform for the integration and synthesis of multi-layer geospatial data sources

Geospatial Big Data provide unique opportunities to look into “*informed urbanization*”. Despite the need for methodological investigations in innovative computational strategies for data handling, much current

research into spatial Big Data is limited in its scope as it is often constraint by a single data source. It is likely to contain multitude uncertainties to answer specific scientific questions. To reduce uncertainties, I look into augmenting geographic contexts to the existing datasets by spatial Big Data synthesis and integration with multi-layer geospatial information. For example, I integrated the geo-located Twitter data with information from detailed land use maps, which adds an extra dimension as semantics to the recorded user locations. This minimizes the gap often found between purely data-driven analytics and theoretical based approaches. The output from this research line is a prototype GIS-based geospatial Big Data platform, which utilizes high-performance, distributed computing for multi-source data integration and synthesis; built-in capabilities for incorporating advanced data mining algorithms; and a visual-analytic framework for presenting scientific findings.

Vision for Future Research

My vision for advancing my research career for the next 5 years starts with building my own research team (Geo-Complexity Lab as a working name) with a research agenda focusing on developing novel methods and models to better understand human-urban environment interactions (i.e., geo-complexity in urban and human dynamics). To be specific, my envisioned research agenda extends my current research and consists of three core components: (1) Continue developing a holistic GIS-based spatial Big Data synthesis platform, which leverages high-performance computing for spatial Big Data integration, processing, analysis, and visualization. (2) Develop a set of comprehensive visual and analytical methodologies on top of the platform to tackle specific scientific questions related to the theme geospatial complexity in human-urban environment interactions. (3) Seek interdisciplinary collaborations based on the developed platforms and methods to address societal issues concerning the urban environment and people, such as sustainable development of cities, quality of life in the environment for “smarter cities” application, and public health, to name a few.

To lay out tangible plans for the envisioned research, the first component has been equipped with a prototype GIS-based geospatial Big Data platform, which is able to chain multiple layer of geospatial information for studying geographic context-aware spatial interactions. Future efforts will tailor the implementations to link domain specific knowledge for urban related studies. Cyber infrastructure support to this platform will come from university-wide or external computational resources that I have successfully acquired before. Importantly, this research line shows promises for future pursuit as it has already enabled several interdisciplinary collaborations with funded research opportunities (the third component). For examples, it is used to study mobility and activity patterns affected by multiple environmental factors for health related research, using GPS logs and urban infrastructures to evaluate environmental exposures to address public health issues, and using Twitter data for dynamic population density estimation and mapping in the urban environment. While the first component aims to drive technical innovations for geospatial Big Data research, the second component is dedicated to advance the methodological and theoretically development to expand the definition of “spatial interaction” for studying urban and human dynamics. Spatial interaction essentially connects human activities with the physical environment. This thinking serves as a guideline when comes to collaborative research efforts and applications for urban related studies. I firmly believe computational geography plays a critical role in promoting GIScience in the age of spatial Big Data. I am confident that my proposed research line will encourage connections and collaborations from colleagues with different research perspectives. The research output will be consolidated for teaching, publications and research grant opportunities.