Abstract:
The volume of available geospatial data increased tremendously. Such data includes but is not limited to weather maps, geo-tagged social media, and Internet of Things (IoT) sensors. Making sense of the geospatial properties hidden in the data may benefit many disciplines such as climate change analysis, urban planning, and transportation engineering. Such data-intensive spatial analytics applications rely on the underlying data management systems to efficiently manipulate, query and manage data.

In this talk, I will present my research efforts on crafting data systems to support scalable and interactive analytics over big spatial data. I will first tackle the scalability issue by introducing a cluster computing system called GeoSpark which offers scalable spatial queries, visualization, and simulation. Then I will present a sampling middleware system called Tabula to address the interactivity issue. Tabula sits between the data management system and the front-end visualization dashboard (such as Tableau) to uphold interactive spatial visual analytics. Third, I will describe my work on lightweight database indexing mechanisms, such as Hippo index and Hermit index, which reduce the storage and maintenance overhead in big data systems. Finally, I will briefly present my ongoing and future work on building new data system components for large-scale geospatial streaming data analytics and dynamic visualization.

Bio:
Jia Yu is a PhD candidate in the Department of Computer Science at Arizona State University. He is the recipient of 2019 ASU Fulton Engineering Graduate Fellowship, 2017 ACM SIGSPATIAL Student Research Competition bronze medal, and 2019 SSTD Best Demo Paper Runner-up. Jia’s research focuses on large-scale database systems and geospatial data management. His research outcomes have appeared in the most prestigious database / GIS conferences and journals, including SIGMOD, VLDB, ICDE, SSTD and Geoinformatica Journal. He is the main architect of several open-source systems which include GeoSpark project that receives 10,000 downloads per month and has users / contributors from major companies (e.g., Facebook, Uber, AT&T, and MoBike).