



IEEE

Institute of Electrical and Electronic Engineers
www.ieeefortwayne.org

Fort Wayne Section Technical Meeting

Tuesday, May 6, 2014 6:30PM

Meeting Location

IPFW Engineering & Technology Building – Room ET 346

Email Reservation to: Kevin.vanzuilen@exelisinc.com

Or use Purdue supported Qualtrics service to [sign up](#) for this event.

By Monday May 5 – Seating is limited



**Pizza
Provided**

A Presentation On

Spatial Data Mining in the Era of Big Data

by

Jin Soung Yoo

Associate Professor

Department of Computer Science

Indiana University-Purdue University Fort Wayne

Abstract

So-called “Big data” is a fact of today’s world and brings not only large amounts of data but also various data types that previously would not have been considered together. Richer data with geolocation and date and time stamps is collected from numerous sources including mobile phones, personal digital assistants, social media, vehicles with navigation, GPS tracking systems, wireless sensors, and outbreaks of disease, disaster and crime. The spatial and spatiotemporal data are considered nuggets of valuable information. Spatial data mining is the process of discovering interesting and previously unknown, but potentially useful patterns from large spatial data and spatiotemporal data. As one task of data mining, spatial association mining serves as a useful tool for discovering correlations and interesting relationships among spatial events and/or features. However, the computation of spatial association analysis is inherently too demanding of both processing time and memory requirements. Furthermore, explosive growth in the spatial data emphasizes the need for developing new and computationally efficient methods tailored for analyzing “Big” data. This talk shows how the power of parallel/distributed processing in the cloud environment is leveraged to achieve higher spatial data mining processing efficiency. For large-scale spatial data processing on clusters of commodity machines, a MapReduce-based spatial association mining method is presented. The developed algorithm was evaluated on Amazon Elastic MapReduce platform with Hadoop using real world data as well as synthetic data.