WRIGHT STATE UNIVERSITY

Department of Computer Science and Engineering

Ph.D. Dissertation Defense

"PERSONALIZED AND ADAPTIVE SEMANTIC INFORMATION FILTERING FOR SOCIAL MEDIA"

By Pavan Kapanipathi

Wednesday, April 6, 2016 – 8:00 am
366 Joshi

Ph.D. Committee: Drs. Amit Sheth, Advisor, T.K. Prasad, Derek Doran, and Prateek Jain (Nuance Research Labs)

ABSTRACT:

Social media has experienced immense growth in recent times. These platforms are becoming increasingly common for information seeking and consumption, and as part of its growing popularity, information overload pose a significant challenge to users. For instance, Twitter alone generates around 500 million tweets per day and it is impractical for users to have to parse through such an enormous stream to find information that are interesting to them. This situation necessitates efficient personalized filtering mechanisms for users to consume relevant, interesting information from social media.

Building a personalized filtering system involves understanding users interests and utilizing these interests to deliver relevant information to users. These tasks primarily include analyzing and processing social media text which is challenging due to its shortness in length, and the real-time nature of the medium. The challenges include: (1) Lack of semantic context: Social Media posts are on an average short in length, which provides limited semantic context to perform textual analysis. This is particularly detrimental for topic identification which is a necessary task for mining users interests; (2) Dynamically changing vocabulary: Most social media websites such as Twitter and Facebook generate posts that are of current (timely) interests to the users. Due to this real-time nature, information relevant to dynamic topics of interest evolve reflecting the changes in the real world. This in turn changes the vocabulary associated with these dynamic topics of interest making it harder to filter relevant information; (3) Scalability: The number of users on social media platforms are significantly large, which is difficult for centralized systems to scale to deliver relevant information to users. This dissertation is devoted to exploring semantic techniques and Semantic Web technologies to address the above mentioned challenges in building a personalized information filtering system for social media. Particularly, the necessary semantics (knowledge) is derived from crowd sourced knowledge bases such as Wikipedia to improve context for understanding short-text and dynamic topics on social media.