Data Visualization and Data Mining/Processing

Luke Zhang and John Yen

Our research interests lie in combining data-oriented methods and user-centered design to support information exploration and knowledge discovery in large datasets. The rapid growth of data volume and the increasing complexity of data types in business and government intelligence analyses pose new challenges for understanding large-scale, heterogeneous, ill-structured data sets (e.g., fragmental intelligence information embedded in number, text, image, video, and social network). Effective decision-making and analytical systems must combine data-oriented computational methods (e.g., data-mining, machine-learning), which can identify useful patterns hidden in massive data, with user-centered designs, which consider the perceptual, cognitive, and task characteristics of human users in computing systems.

We take a "work-centered" approach that uses advanced algorithms (e.g., clustering and agent-based learning) to identify useful patterns hidden in massive and complex data and then support evidence-based decision-making and analysis with interactive visualization. Our innovative approach emphasizes the importance of both data and user tasks, and tries to maximize the strengths of both machine computation and human decision-making. We focus on building models that concern the interaction between human information analysis activities and machine algorithms through such modules as analytics tasks, information patterns, and task workflows. Based on these models, we develop systems and tools to support the analyses of the problems involving complex data in various situations, such as the discovery of optimal designs and alternatives in engineering design challenges, the impact of organization affiliation on scientific collaboration, the roles of organizational structures on technology innovation.