

ISE-INFORMS Research Seminar

A Network-based Model for Conversational Components in Text Mining

SPEAKER INFORMATION

Dr. Onur Seref

Associate Professor

Business Information
Technology

Virginia Tech

SEMINAR SESSION INFORMATION

DATE: WEDNESDAY, April 12

TIME: 12:15pm

LOCATION: Durham 260

PROVIDED: Pizza and Soda

Abstract: We present a network-based model to identify and cluster conversational components in different classes of text data using prevalence scores of n-gram structures and their connections. Our framework extends the “bag-of-words” models by creating sub-graphs of connected n-grams. The paths in these sub-graphs represent sequences of words, which form conversational phrases with richer contextual meaning. We apply the proposed framework to different domains and compare its effectiveness to standard text-mining models such as LDA.

Biography: Onur Seref is an Associate Professor of Business Information Technology in the Pamplin College of Business at Virginia Polytechnic Institute and State University. He received a Ph.D. in Operations Research from the University of Florida and an M.S. and B.S. in Industrial Engineering from the Middle East Technical University in Ankara, Turkey. His primary research interests involve combinatorial and nonlinear optimization methods in data mining and machine learning with applications in biomedicine, network optimization, and text mining. His research has been published in a number of journals including Operations Research, INFORMS Journal on Computing, Annals of Operations Research, and various IEEE Transactions journals. Dr. Seref is a member of INFORMS and currently serving as the chair of the INFORMS Data Mining Section.

Faculty and Students are invited