

Analysis of Multimodal Data in the domain of eLearning and Insider Attacks

Abstract

This talk provides a brief overview of designing various systems to aid eLearning and prediction of insider threat. We will also analyze the multimodal data gathered while using these systems.

The last decade has seen major shifts in the field of education - in both the generation of content as well as the mode of delivery. With the affordable network infrastructure across the world as well as the ubiquitous availability of smart-phones, tablets and personal computers, there is a growing demand for online educational material. These materials come in the form of text-based websites or educational blogs as well as multimodal learning systems consisting of lecture videos, images, text and combinations of them. Using more than one medium in the learning process typically provides better illustrations of the study materials, thus making the concepts more perceivable. These multimodal learning systems, as an effective method for helping students to understand complex concepts, have attracted much research interest recently. The crowd sourced web-based encyclopaedia, Wikipedia is a source of educational material that is referred to by students on a regular basis, sometimes as frequently as daily. These Wikipedia articles, together with the public videos available on YouTube, Videolectures.Net, etc. provide important supplementary reference materials for e-learning.

In the first study, we use topic modelling to map such videos and articles in the common semantic space of topics and recommend the videos with high similarity values for the further explanation of the concepts present in the articles. This integrated framework reduces the users' efforts in searching for relevant e-learning videos. Moreover, it provides a mechanism for searching lecture videos based on the course content present in them.

The increase in user-generated videos hosted on video-sharing sites like YouTube, Videolectures.net, has resulted in an increase in the number of lecture videos also. More and more universities sharing their lectures online have also added to the content duplication in the field of lecture videos. Often there are too many videos on the same topic hosted on these video-sharing sites. However, even though these videos are similar to each other in terms of academic content, there are differences in the learning experience for the students when presented with the videos. This is because the videos vary in their image resolution, mode of lecture delivery, the speed of lecture delivery and other such factors. Hence in the second study, we explore the various multimodal features of the lecture videos which affect the users' experience. We also conduct a user survey to validate our hypothesis and finally design a ranking system to rank videos based on the quality of their multimodal features.

Insider threats pose a challenge to all companies and organizations. Identification of culprit after an attack is often too late and result in detrimental consequences for the organization. It has been proposed that certain personality vulnerabilities place individuals to be at risk to perpetuating insider threats should the environment and opportunity arise. To that end, our study utilizes a game-based approach to simulate a scenario of intellectual property theft and investigate behavioural and personality differences of individuals who exhibit insider-threat related behaviour. Features were extracted from games, text collected through implicit and explicit measures, simultaneous facial expression recordings, and personality variables. We applied ensemble machine learning algorithms and show that they produce an acceptable balance of precision and recall. Our results showcase the possibility of harnessing personality variables, facial expressions and linguistic features in the modelling and prediction of insider-threat.

Speaker



Dr Subhasree Basu
PhD in Computer Science at National University of Singapore

Subhasree Basu is a research fellow at School of Electrical and Electronic Engineering, NTU. She did her MSc in Computer Science from Chennai Mathematical Institute, India and PhD from School of Computing, NUS. Her research interest lies in Text Mining, Data Mining, Multimedia Analytics, Multimodal Data Analysis. She has experience in machine learning and text mining to solve problems from various domains like e-learning, psychology, etc.

Details

Thursday, April 18, 2019
5:00 PM - 6:00 PM

Meeting Room 4-4, Level 4

School of Information Systems, Singapore Management University, 80
Stamford Road, Singapore 178902, Singapore

We look forward to seeing you at this research seminar.