Abstract No: SO-01 **Software Intensive Systems**

An approach to personalize learning using big data analytics for higher education

Asma Jabir* **and Chathura Rajapakse**

Department of Industrial Management, Faculty of Science,
University of Kelaniya, Sri Lanka
azinneera@gmail.com

The concept of *BYOD* (Bring Your Own Device) has gained popularity in student-centered learning and higher education institutions make significant investments on improving the wireless network to enhance this. Virtual Learning Environment and Learning Management Systems were introduced and personalization of learning becomes the next milestone. The huge streams of data produced by these Wi-Fi networks makes ground for Big Data analytics to identify opportunities in educational environments to adopt personalized learning.

The term 'Personalization' refers to the tailoring of content and recommending items by inferring what interests a user based on previous or current interactions with that user, and possibly other users. This research proposes an approach to personalize learning on an online learning platform by providing personalized recommendations of educational web resources, comparative feedback and allocate personalized bandwidths based on the concept of deprioritization (lowering priority ranks of heavy users).

Concepts of Big Data analytics and data mining techniques will be used to satisfy the objectives. The approach consists of offline phase (modelling phase) and online phase (recommendation /deprioritization) phase. In the offline phase, models will be developed for recommendation and deprioritization separately. For recommendation a hybrid filtering method will be used. *k-Nearest Neighbour*, a user-based collaborative filtering technique, will be used with correlation based similarity measure with demographic filtering based on demographic classifiers (faculty, year, General/Special/Honors, GPA) to eliminate the cold start problem. To increase the efficiency and accuracy, *k*-means clustering will be used as an intermediate step to determine usage clusters to group users exhibiting similar browsing patterns and page clusters to discover pages with similar access patterns. For this the access logs of the University of Kelaniya's Wi-Fi network will be utilized. The parameters for usage clustering would be the timestamp, web resource and category (education, social networking, gaming etc.) whereas the parameters for page clustering would be category and temporal concepts. In the online phase, first the cluster that the current active user belongs to will be identified and *k-NN* will be applied on that particular cluster to recommend web resources. These techniques also provide the basis for comparative feedback compared to top scorers of the same area of major. For personalized allocation of bandwidth a separate *k*-means clustering will be performed to identify heavy users during the offline phase. During the online phase deprioritization will be applied accordingly if the current user belongs to the heavy users cluster and there is a heavy traffic in the network. Cross validation will be used to evaluate the models.

Keywords: Big data, e-learning, Personalization, Network management, Web usage mining