Exploring Self-Organizing Learning Groups in an Online Platform

Thierno Bah
Borough of Manhattan Community College, CUNY

Abstract
Online learning platforms such as Classroom Salon[1] help capture students' interpretation of documents they read and videos they watch. The documents and video annotations reveal the student's interaction with the content and their specific interests. Such a platform enables students to learn from each other by self-organizing into groups. However, this task may prove hard to accomplish if the platform has a large number of users. A student will not be able to browse through thousands of annotations to find users of similar interest. In this study, we look for various text analysis tools and methods to automatically find groups that can collaborate more effectively. We are using well known algorithms and tools such as Wordnet[2], DocuScope[3] and a topic modeling system with which users can find each other for possible project collaborations in their respective fields.

Background
Topic models are algorithms that uncover the hidden thematic structure in document collections. They help us develop new ways to search, browse, or summarized large archives of text[4]. One of these algorithms is LDA (Latent Dirichlet Allocation)[5]. Topic model analyzes the count of words and finds topics which are word proportion and topic assignments for each document at the same time.

Motivation
Through the use of LDA (Latent Dirichlet Allocation) algorithm, we want to group users of an online platform in self-organizing groups of students with the similar interest.

Methodology

- Extract all the students' comments
- Find the underlying topics from the comments
- Find the topics proportion per document (comment)
- Infer the topic preference for each student
- Compute the correlation between topics which yield a graph
- Find cliques with k maxsize such that the sum of weight of edges is maximized

Result
We have come up with an algorithm that enable us to identify the self-organizing groups that form based on the topics that the students are interested in. The graphs below represent the students correlations. Figure 4a shows the correlation between students with more than 70% similarity. Figure 4b shows the groups that forms.

Discussion
Our goal is to identify the groups that form based on topic interest. Our hypothesis is that if students comment usually on similar topics they must have similar interest and questions. Therefore, by self-organizing, they can collaborate more effectively. Now that we have these groups that are formed, our next move is to validate our result. Moving forward, we are going to use this algorithm that we have come up with. We are going to test it in an actual class while the students are still taking the course and see how much of our hypothesis is verified.

Figure 1: Sample of Topics students' comments

Figure 2: Students topic preference

Figure 3 comment proportion

Figure 4 a: Correlation between students

Figure 4 b: self-organizing groups

Reference

Acknowledgements to the Louis Stokes Alliance for Minority Participation (LSAMP) NYC and Mid-InfraRed Technologies for Health and the Environment (MIRTHE).

This material is based upon work supported by the National Science Foundation under Grant No. EEC-0540832.