Educational Data Mining using Chance Discovery from Discussion Board

GCCCE 2016 -

Preliminary Results Presentation

Introduction and Motivation

- E-learning systems have become more popular recently in the higher education sector producing a **large volume of data** created by students and teachers inside via collaborative online discussion forums, which are difficult to analyze without using data mining technologies.

- The forums contain a lot of useful data for **educational data mining** to extract insightful pattern so that educators can better understand the thinking patterns of students during the learning process.
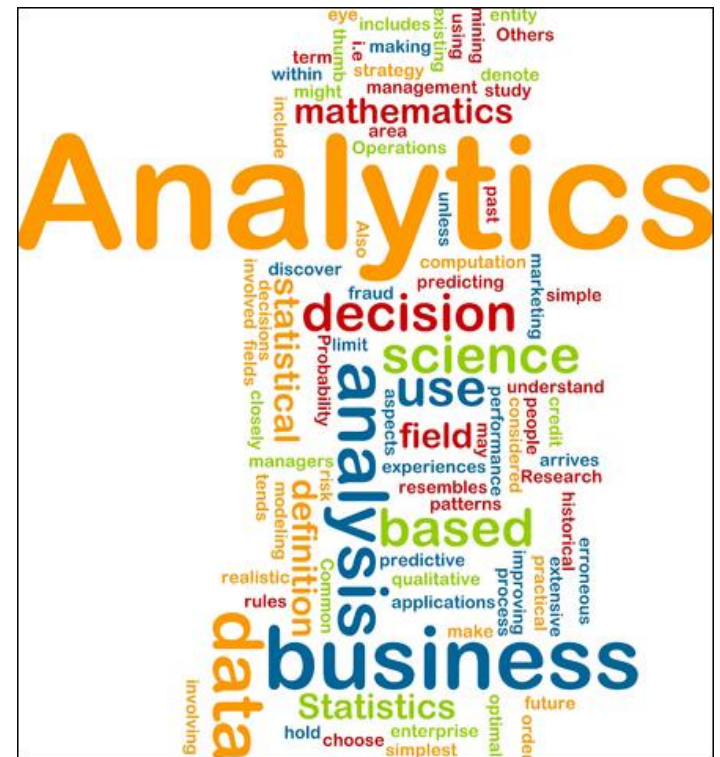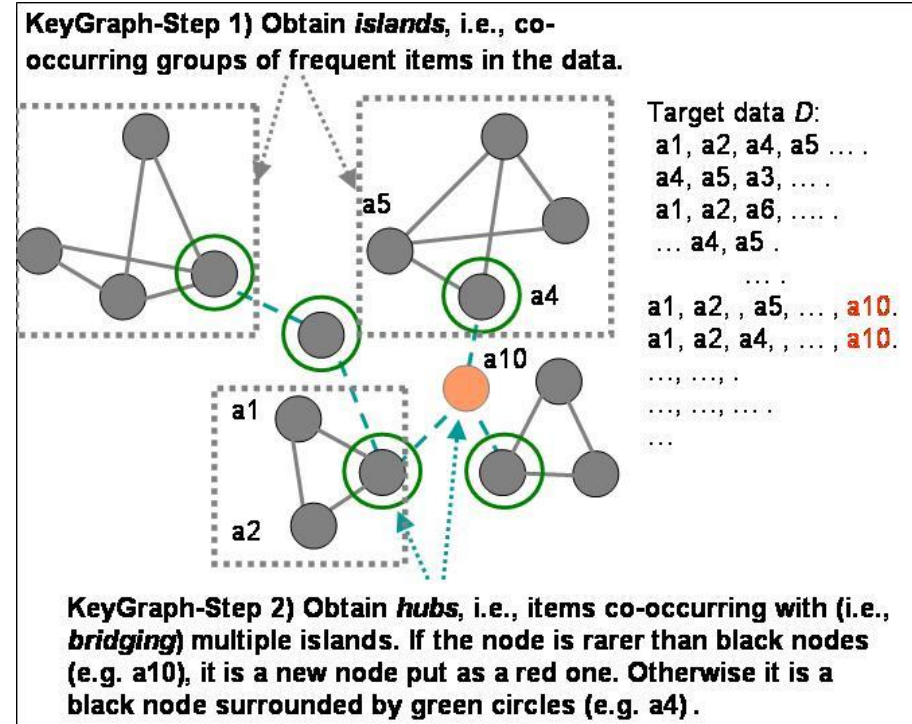
Project Objectives

- Develop innovative **analytical methods and tools** facilitating assessment of teachers to evaluate the students learning in the online learning environment to determine if students are able to meet **learning objective** and / or **generate new knowledge** beyond the expected learning activities.

- Present our preliminary research findings on educational data mining using student discussion forums for future research planning.

Prior Works - KeyGraph

- The black nodes and black links represent the frequent items and their co-occurrences, implying an established trend in the data.

- Black nodes and links form clusters representing concepts.

- The red nodes and red links are the **bridges** which connect multiple clusters or some phenomena such as transition of events from one to other clusters.

- Red nodes can be regarded as chances.

KeyGraph-Step 1) Obtain *islands*, i.e., co-occurring groups of frequent items in the data.

Target data *D*:
a1, a2, a4, a5 … .
a4, a5, a3, … .
a1, a2, a6, … . .
… a4, a5 .
 … .
a1, a2, , a5, … , a10.
a1, a2, a4, , … , a10.
…, …, .
…, …, … .
…

KeyGraph-Step 2) Obtain *hubs*, i.e., items co-occurring with (i.e., *bridging*) multiple islands. If the node is rarer than black nodes (e.g. a10), it is a new node put as a red one. Otherwise it is a black node surrounded by green circles (e.g. a4).

Source: Ohsawa, Y., Benson, N. E., & Yachida, M. (1998, April). KeyGraph: Automatic indexing by co-occurrence graph based on building construction metaphor.

Prior Works - Chance Discovery (CD)

- Three stages: Scenario Generation, IMG, and AHP.

- KeyGraph is the algorithm used to **generate scenarios** for **finding chances** from document to (a) *valuable result*(s).

- IMG, which aims at evaluating the chances using human intervention (with expert / prepared minds) to identify values, is the _**human model**_, which is enhanced by using the third stage AHP.

- Wang, H., Ohsawa, Y., & Nishihara, Y. (2013). A system method to elicit innovative knowledge based on chance discovery for innovative product design.

- <u>CD can produce serendipitous findings (as chances) from its computational or human models.</u>
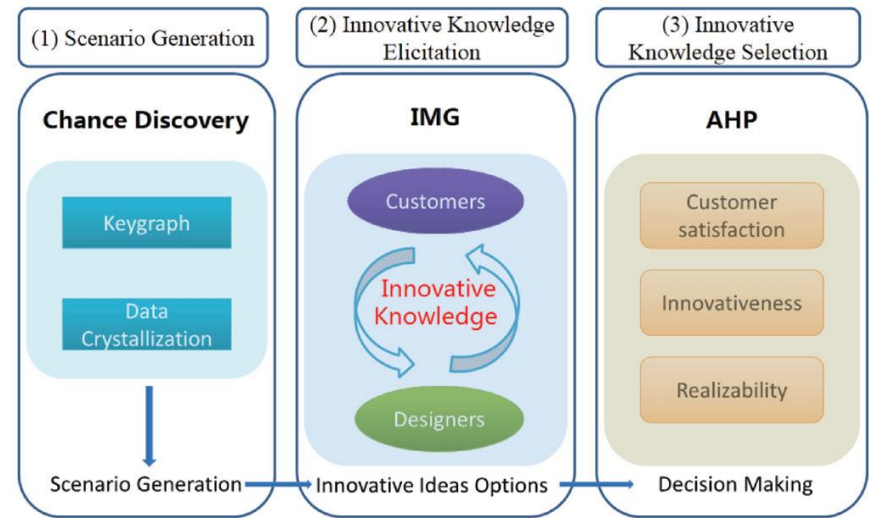
Scenario diagram generated by KeyGraph

serendipitous finding ?
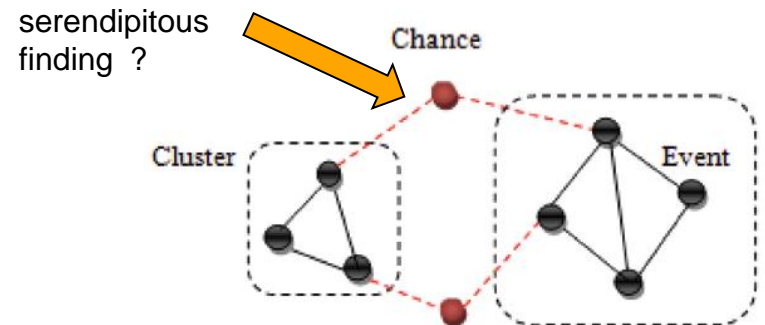
Experiment Design

- In this project, a total 24 undergraduate students in the HKIEd from the General Education course called "***Technology, Entertainment and Mathematics***" have been sampled for this preliminary experiment.

- One of the course requirements was to complete a *reflective posting* on an online discussion forum in Schoology.

- They were asked to watch a BBC documentary film called "***Beautiful Equations***" and other ***selected movies***.

- Afterwards they posted their reflections in the forum. Each student was also required to <u>comment on three self-selected peers</u>, which were extracted in our experiment as text files for analysis.

Reflective Blog: BBC (Beautiful Equations)

Write a comment

Highligh... ▾   107 Posts

behind things, including things that cannot be seen. The mathematical equation helps us to understand the theory behind the things around us and people can know the physics in the world. Things can be explained by mathematical equation.

Second, I believe the equations are beautiful after watching the documentary because the equations are universal language. There are many languages in the world but mathematical equation use symbols to represent. In the documentary, all people are discussing the beautiful equation by using the same language, mathematics. Though people have different languages, people share same mathematical ideas in equations so that people can communicate by using mathematical vocabulary.

Different people have different conception and definition about "beauty". Mathematical equation is beautiful because it is useful in reality and become a universal language. Although I am not a science student and cannot explain every details of the equation, I find that mathematical equation is beautiful after watching this documentary as mathematical equation give us ideas behind our imagination and facilitate human development.

Show Less

Like 😊 1 · Reply

↳ Hide All 5 Replies

Sun Feb 22, 2015 at 7:22 am

I do believe mathematics is an universal language apart from English as everyone in the world knows numbers. Many people do not classify mathematics can really communicate with others but scientists are sending numbers in binary system in order to find out is there any aliens outside the Earth. Mathematics is much more beautiful than we can imagine. The details of nature which full of mystery can be explained by several equations. What is more, these beautiful equations help people construct, entertain and improve their lives.

Show Less

Like 😊 1 · Reply

Sat Feb 28, 2015 at 9:25 am

I agree with you. Firstly, mathematical equation is beautiful because it can be used to explained the natural order. We can explain the world through a mathematical equation. It is very amazing. It can also apply to universal in order to explain and discover more about the world. Secondly, mathmathical equation is beautiful because it is the universial language. It can bring people together. People can share their idea through these equations. It is beautiful because mathematical equations can break the limitation among the world.

Show Less

Like 😊 1 · Reply

Experiment Design

- A software tools called "Polaris" from Ohsawa Laboratory was used for mining text from the following sources.

- Sources of Data (text format):

  – Online reflective discussion forum, etc.

- Performed analysis using KeyGraph to generate the visual patterns to identify:

  – The formulation of key concepts from black nodes and links

  – chances (red nodes and links) for the purposes of decision making and planning in the associated areas above.
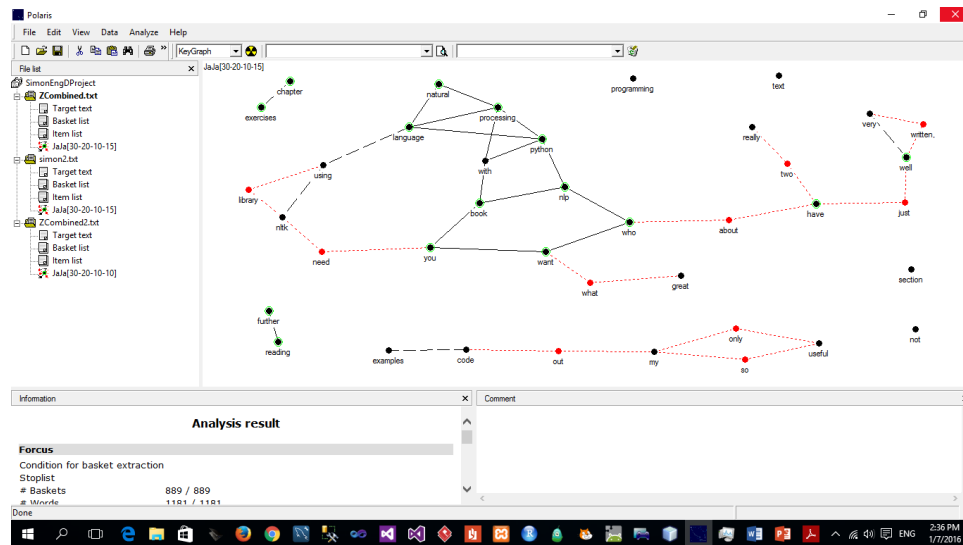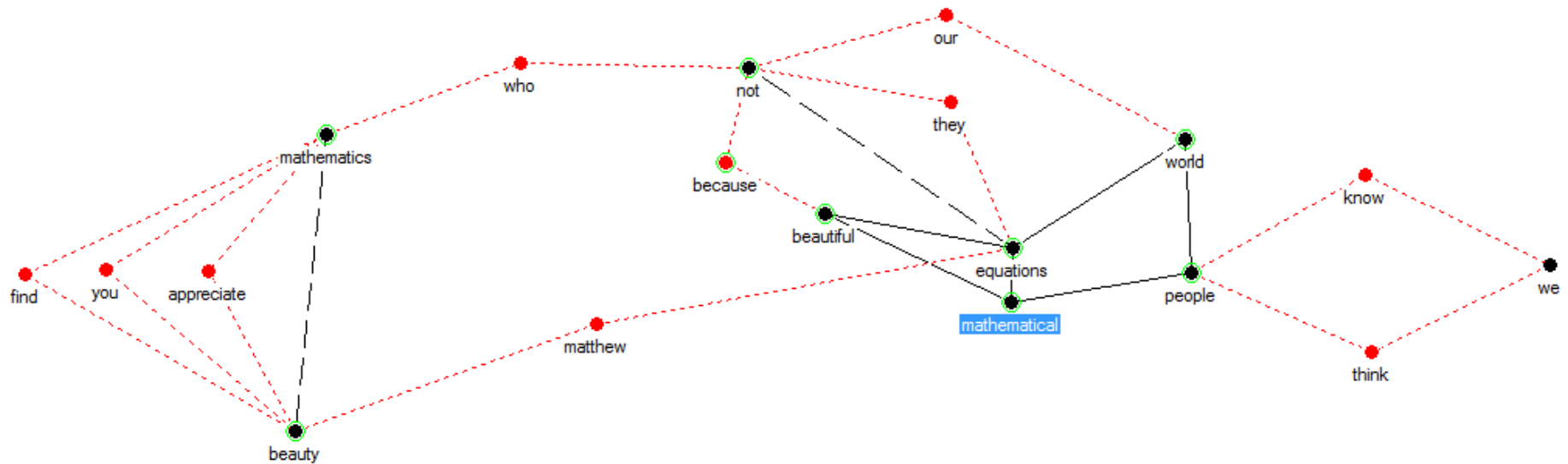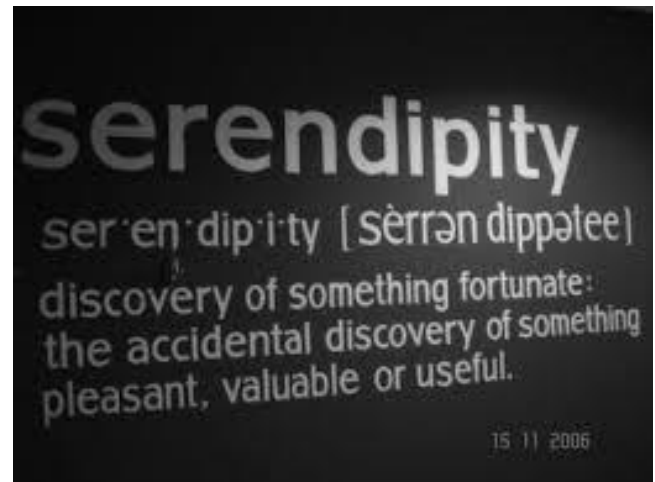
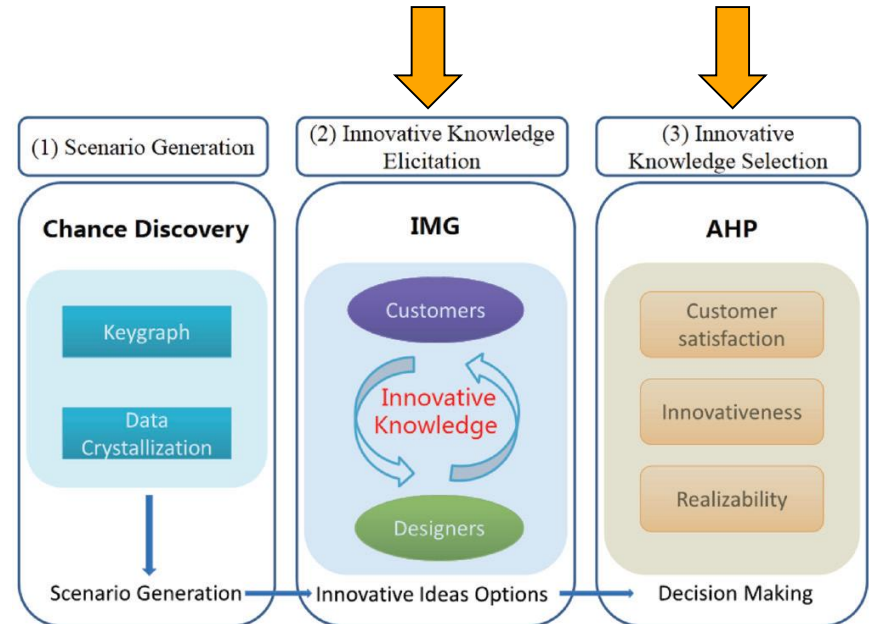# Preliminary Results – Selected Movie
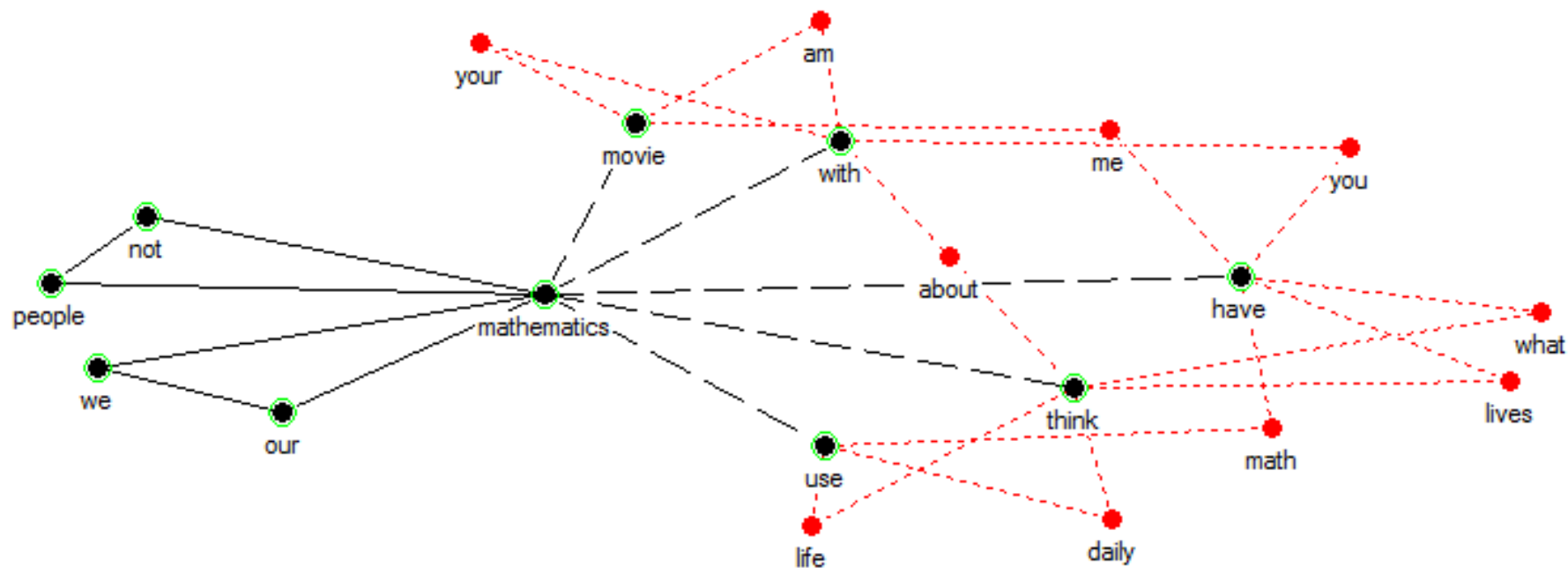
Potential Contribution

- Teachers can better understand the patterns of thinking of students during the learning process.

- The assessment of students can be facilitated through a systematic approach with effective pedagogic changes for particular students through the learning process to optimize their learning outcome

- There seems a huge potential of unexpected learning discoveries – serendipity, in collaborative online learning.

serendipity
ser·en·dip·i·ty [sèrrən dippətee]
discovery of something fortunate:
the accidental discovery of something
pleasant, valuable or useful.

15 11 2006

Follow-up

- Explore different perspectives to improve the analysis like

  - sample sizes,

  - multiple academic subjects,

  - writing styles,

  - and other qualitative factors.

- Proceed to the IMG and AHP stages of CD for evaluation of the chances already spotted.

End of Presentation – Thank You