Ranking Concept Maps and Tags to Differentiate the Subject Experts in a Collaborative E-Learning Environment

Ahmad A. Kardan, Fatemeh Hendijanifard, Solmaz Abbaspour

Advanced E-Learning Technology Laboratory
Department of Computer Engineering and Information Technology
Amirkabir University of Technology, Tehran, Iran
Email: {aakardan, hendijani, s_abbaspour}@aut.ac.ir

Abstract

Members of a collaborative learning environment need to refer to the subject experts. Therefore, it is necessary to identify the subject experts and to introduce them to the other members. To achieve this goal, one approach is to make use of concept map evaluation by means of ranking methods. Another approach is to utilize tagging methods for finding subject experts in a collaborative learning environment. In this article a new approach for estimating the knowledge level of the members in a virtual environment is introduced, which is based on the concept mapping and tagging. In construction of the concept maps, concepts could be linked to any type of related resources. The labels associated to these links could be assumed as tags for those resources. Therefore, tagging methods could be used as a measure for ranking the quality of the resources and the expertise of members. In the proposed method, four parameters are considered for ranking the subject experts: concept map ranking, tag ranking, tag and resource relevancy, and the relation between the number of tags and the number of concepts. This paper presents the required algorithms which examine these parameters to determine the subject experts. These algorithms and the evaluation method will be discussed in detail.

Keywords: Elearning, Concept map, tag, ranking method, collaborating environment

1 Introduction

It can be not easy to get a satisfactory answer to a problem by using search engines. Instead, one may prefer to find and ask someone who has related expertise; online communities and collaborative learning environments have emerged as one of the most important places for people to seek advice or help (Jun Zhang, Mark S. Ackerman & Lada Adamic, 2007). Thus a common issue in collaborative learning environments is finding experts. Some works have been done on ranking the knowledge of people with the help of concept maps, or utilizing tagging methods to identify subject experts in a collaborative learning environment.

Concept map is a graphical representation of a human's understandings of a domain of knowledge. Concept maps represent this understanding by using a two-dimensional network in which nodes correspond to concepts and links demonstrate concept relationships. In a concept map, *concepts* are the labels used to refer to objects and *linking phrases* (the text on the links) are usually verbs (Alejandro Valerio *et al*, 2008).

Concept mapping is used to enable individuals to make new knowledge, externalize knowledge, share and compare knowledge. Given that each person's understanding of a domain is different, even on the same topic, the maps constructed by everyone are different, reflecting their personal knowledge structures (Alejandro Valerio *et al*, 2008).

Another means that is used in collaborative environment is tagging. Collaborative tagging systems provide web users a new means of organizing and sharing resources such as bookmarks on the Web (M. G. Noll & C. Meinel, 2008). Such systems also allow users to search for documents relevant to a particular topic or for other users who are experts in a particular domain. However, identifying relevant documents and knowledgeable users is not a trivial task. Thus new approaches by tags are used to rank resources and users (R. Wetzker & C. Zimmermann & C. Bauckhage 2008).

In this article a new approach for approximating knowledge of members in such environments is introduced. In this approach concept maps and tagging methods are applied to identify experts in particular subjects and is discussed in detail in the following sections. Section 2 of this article includes related works. System architecture with details of each stage is represented in section 3. And the last section is about evaluation and future works.

2 Related works

Tag Ranking Methods: "Studies in psychology have shown that experts involve in the ability to select the most relevant information to achieve a goal" (P. J. Feltovich *et al.*, 2006). In the context of collaborative tagging, users assign tags to resources to facilitate retrieval of resources. "Therefore, it is believed that an expert should be someone who not only has a large collection of documents annotated with a particular tag, but tends to add high quality documents to their collections. In other words, there is a relationship between the expertise of a user and the quality of a document." In tag ranking methods, usually spammers are omitted to find users that have used high quality tags and rank documents. Koutrika *et al.* (2007) are the first to discuss methods of tackling spamming activities in collaborative tagging. There are also proposals for detecting spammers in tagging systems based on machine learning approaches (A. Madkour *et al.*, 2008; R. Krestel & L. Chen, 2008). In (Michael G. Noll *et al.*, 2009) the proposed algorithm named SPEAR in addition to demoting spammers in the ranked list of users instead of detecting their presence; it finds experts. They believe that different types of methods, including detection, demotion, and also prevention is complementary in tackling spammers (P. Heymann *et al.*, 2007).

Concept Maps Ranking Methods: "Despite the variety of concept maps that arise from the differences among map builders, some maps can be considered "better" than others, based on a variety of criteria. One concept map could show a "deeper understanding" of a topic than another, perhaps reflecting that the first was constructed by an expert and the second by a novice" (Alejandro Valerio *et al*, 2008). Therefore different methods have been done on evaluating concept maps.

An evaluation may involve making qualitative and/or quantitative remarks (John R. Mcclure *et al.*, 1999). Topological features and semantic features are two features that assess the quality of a concept map (Alejandro Valerio *et al*, 2008).

Qualitative relational methods assess the accuracy of each proposition; quantitative structural methods score the valid elements in comparison to a criterion map. “Ruiz-Primo pointed to three scoring systems: (1) of proposition accuracy, (2) of convergence with a criterion map and (3) of salience which is the “proportion of valid propositions out of all the propositions in the student's map”. Ruiz-Primo and colleagues found correlations between this convergence score of construct-a-map procedures and learners’ explanatory skills which gives evidence, that concept mapping assessment is “in fact measuring what is claimed”.” Mc Clure and colleagues compared holistic, relational and structural scoring methods without and with the use of a master map unveiling a high reliability for the latter (Steffen Schaal, 2008). Some other methods which were evaluated are: holistic, holistic with master map, relational, relational with master map, structural, and structural with master map (John R. Mcclure *et al.*, 1999). In (Steffen Schaal, 2008) structural attributes of concept maps are scored. The relevant structural attributes were the ‘volume’ of a concept map, the ‘ruggedness’ which is the division of a concept map into un-connected sub-maps and the amount of accurate propositions in relation to the volume (Steffen Schaal, 2008).

3 System Architecture

3.1 System Outline

In the proposed system shown in figure 1, learners in a collaborative environment construct concept maps to show their knowledge for a specific subject. Learners can link resources to their maps to further explain their contents (concepts or linking phrases). Concepts associated to these links are single words or in the form of a phrase, and we can assume these concepts as tags drawn on the resources. Tags for any resource show how much the learner understood the main concepts of that source. As stated in (R. Wetzker & C. Zimmermann & C. Bauckhage, 2008) “the simplest way to assess the expertise of a user is by the number of times he has used a tag (or a set of tags) on some documents. However, this does not take into consideration the facts that quantity does not imply quality, and that there exist spammers who indiscriminately tag a large number of documents” (R. Wetzker & C. Zimmermann & C. Bauckhage, 2008). It is believed that an expert should be someone who tends to add to his collection high quality documents. “Thus there is a relationship of mutual support between the expertise of a user and the quality of a document” (A. Madkour *et al.*, 2008).

Consequently tagging methods are used for knowledge evaluation and expert identification. To achieve this, system extracts the linked concepts – tags – from the maps for each learner, and to check relevancy of tags to the resources, compares tags with the keywords of resources. In this way quality of tags are ensured. From this part, at the first stage the score of tags will be achieve; and in the second stage a concept map scoring system will be applied on the maps. Therefore the Knowledge ranking of the learners is achieved from these two steps and then the subject experts are introduced. Details of each part are followed in section 3.2.

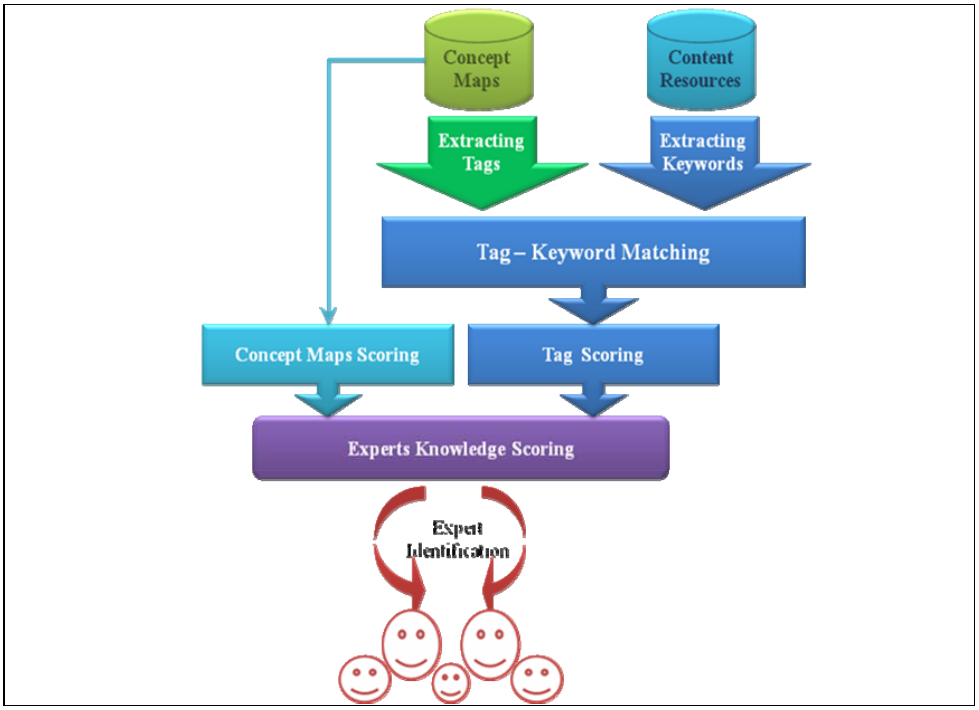

Figure 1. System Architecture

3.2 Stages in Detail

Tag – Keyword Matching: In this module, the linked concepts of every concept map are extracted. The list of these concepts is considered as a set of tags for each map, ***T***= $\{t_1, t_2, \dots, t_m\}$. The keywords for each content resource are determined. Hence a keyword set ***K***= $\{k_1, k_2, \dots, k_n\}$ will be attained. Tags ***T*** and keywords ***K*** are checked to investigate the equivalency between them. WordNet is used to go with the keywords and tags. WordNet is a lexical database that groups English words into sets of synonyms called *synsets*, provides short, general definitions, and records the various semantic relations between these synonym sets. This matching is used in the next stage for ranking tags of each concept map.

Tag Ranking: In this part, two parameters for tag ranking are applied. The first parameter is the similarity of tags to the keywords. For obtaining this similarity, tag set ***T*** for a map is used to compare with keyword set ***K*** with the help of WordNet. The result will give a grade to the tags and is shown in the algorithm1.

It is remarkable to note that the most popular tags are considered too. These tags are attached to a resource by many users. So these words are added to the keyword set ***K***.

Algorithm1. Determining parameter

//Key word set of all content resources: ***K****= {*$k_1, k_2, \dots, k_n$*}*
//Tag set of concept map i: T_{ci}*= {*$t_{1ci}, t_{2ci}, \dots, t_{mci}$*}*
//Set of all concept maps: ***C****= {*$c_1, c_2, \dots, c_p$*}*
For each concept map ***c***
For each tag in T_{ci}
Compare tag with keywords in ***K*** *with help of WordNet*
//use one of the below functions to grade the tag
If tag is exactly one of one of the keywords in ***K*** *then* $s_1=\alpha$
Else if tag is synonym for one of the keywords in ***K*** *then* $s_1=\beta$
Else if tag is hyponym for one of the keywords in ***K*** *then* $s_1=\Omega$
Else $s_1=\zeta$
//S_{1ci} *is the similarity of tags and keywords and shows* T_{ci} *score for this //similarity in concept map* ***ci****.*
$S_{1ci} = \{s_{11}, s_{12}, \dots, s_{1mci}\}$
//End

The second factor for ranking tags is leveraging. Tags are classified to specific, general and not related tags. This classification and grade for each class is determined by an expert in that subject. The result grade is used to rank tags of concept maps. Here is the algorithm for this part:

Algorithm2. Determining s_2 *parameter*

//Tag set of each concept map ***ci****:* T_{ci}*= {*$t_{1ci}, t_{2ci}, \dots, t_{mci}$*}*
//Set of all concept maps: ***C****= {*$c_1, c_2, \dots, c_p$*}*
//Categories of tags determined by expert: Cat = {$specific, general, not\ related$}
For each concept map ***c***
For each tag in T_{ci}
Determine each tag belongs to which category of ***Cat***
//use one of the below functions to grade the tag
If tag is specific then $s_2=\gamma$
Else if tag is general then $s_2=\delta$
Else $s_2=\eta$
//S_{2ci} *shows how specific or general the tags in set* T_{ci} *of concept map* ***ci*** *are.*
$S_{2ci} = \{s_{21}, s_{22}, \dots, s_{2mci}\}$
//End

Parameter s_2 is considered as a weight for s_1 for every concept map. It shows how general or specific the concepts are, and this is an important factor for recognizing knowledge of people in a particular subject. Hence total rank for the tags is calculated as below:

- mci: number of tags in concept map ***i***
- $Sci_{tagTotal}$: for every concept map ***ci***, is summation of $s_1 * s_2$ for each tag in set T_{ci}, divided by number of tags in a concept map.

[1] $$Sci_{tagTotal} = \frac{(s_{1ci} \times s_{2ci})}{mci}$$

Note: grades for s_1 and s_2 in each ***if*** can be determined by comparing with what experts usually use.

Concept Map Ranking: A reliable knowledge structure is necessary for conceptual understanding. Thus, "the interrelationship of concepts is seen as a fundamental attribute of knowledge" (Steffen Schaal, 2008). Of course, the semantic content is always more important than the topological structure, but a "well structured" concept map is considered better than a badly structured map, even if their contents are "equivalent" (Alejandro Valerio *et al*, 2008). The topological taxonomy classifies concept maps into seven levels of increasing structural complexity. In the taxonomy, five features are used to describe the structure of a concept map: "the existence of hierarchical structure, size of concept labels, presence of linking phrases, number of branching points, and number of cross links. Values for these features determine the level of complexity" (Alejandro Valerio *et al*, 2008).

In this part, a topological classifier method described in (Alejandro Valerio *et al*, 2008) is used to categorize concept maps into six levels of expertise. Level -1 defines the default level. The classification is determined by checking the specifications described in table 1 for concept maps.

Table 1. Required conditions for classification of concept maps (Alejandro Valerio et al, 2008)

Level #	**Conditions by level as indicated in Canas, Novak et al. (2006)**	**Conditions evaluated by to classify concept map *M*, which has concept *c* and linking phrase *l* as the nodes of *M*.**
Level -1	No conditions	(default)
Level 0	At least 4 connected concepts Mostly long concept labels Empty linking-phrases 0 to 1 branching points	$\lvert\{c \mid concept\}\rvert \geq 4$ (default) (default) (default)
Level 1	More concepts than long concept labels Half or more missing linking-phrases 0 to 1 branching points	$\lvert\{c \mid concept \mid labelSize(c) < 12\}\rvert \geq \lvert\{c \mid concept\}\rvert/2$ (default) (default)
Level 2	More concepts than long concept labels Less than half missing linking-	(checked at Level 1) $\lvert\{l \mid linking\ Phrase \mid labelSize(l) > 0\}\rvert \geq$

Level #	Conditions by level as indicated in Canas, Novak et al. (2006)	Conditions evaluated by to classify concept map *M*, which has concept *c* and linking phrase *l* as the nodes of *M*.
	phrases At least 2 branching points	$\lvert\{l: linking\ Phrase\}\rvert/2$ $\lvert branchingPoints(M)\rvert \geq 2$
Level 3	No long concept labels No linking-phrases missing At least 3 branching points Less than 3 hierarchy levels	$\forall c: concept, lableSize(c) < 12$ $\forall l: linkingPhrase, lableSize(l) > 0$ $\lvert branchingPoints(M)\rvert \geq 3$ (default)
Level 4	No long concept labels No linking-phrases missing At least 5 branching points At least 3 hierarchy levels	(checked at Level 3) (checked at Level 3) $\lvert branchingPoints(M)\rvert \geq 5$ $\exists c: concept, depth(c, M) > 3$
Level 5	No long concept labels No linking-phrases missing At least 5 branching points At least 3 hierarchy levels At least 1 cross-link	(checked at Level 3) (checked at Level 3) (checked at Level 4) (checked at Level 4) $\lvert crossLinks(M)\rvert \geq 1$
Level 6	No long concept labels No linking-phrases missing At least 7 branching points At least 3 hierarchy levels At least 3 cross-links	(checked at Level 3) (checked at Level 3) $\lvert branchingPoints(M)\rvert \geq 7$ (checked at Level 4) $\lvert crossLinks(M)\rvert \geq 3$
Conditions labeled with (default) are met if the map fails a condition of a following level. Conditions labeled with (checked at Level N) are revised at a previous level.		

Experts Knowledge Ranking: In part three of section 3.2 concept maps are ranked and each map is assigned to a class shown in table 1. Each group of maps in every level of table 1 will be classified again with regard to the rank of tags. Another parameter should be considered here: the ratio of number of tags in the whole map. As stated before, the simplest way to assess the expertise of a user is by the number of times he has used a tag. Thus, this parameter shows how much a person is expertise in one subject and how many resources one has used for different aspects of his knowledge. The quality of tags is considered in the tag rankings. The proportion of tags to the whole map is shown in formula (2):

- *m= number of tags*
- *n= number of nodes (concepts) in the concept map*

[2] $$vol = m/n$$

With the help of vol from formula (2) and due to the previous grades of concept maps and tags for each member, experts in this learning environment could be introduced. So members from the top level with the greatest tag grades and greatest vol will be introduced as subject experts.

4 Evaluation and Future Works

For evaluation of this work, a group of concept maps which are linked to some text resources can be used. By applying the algorithms on each map and determining the grades, experts will be determined. In some concept map repositories, such as repositories of CmapTools, there are expert and novice concept maps. With the help of this information and comparing the results of algorithms, system will be evaluated.

For future works we can determine and verify quantities of $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\Omega}$, $\boldsymbol{\gamma}$, $\boldsymbol{\delta}$ and $\boldsymbol{\eta}$ during different experiments to use in algorithms of part two of section 3.2.

In addition, the vol quotient in formula (2) is believed to be different in the maps related to the experts and usual maps. This parameter can be certified by analyzing maps constructed by experts, such as professors' maps.

Considering vol for experts maps as a threshold, another work is to change map levels pointed in table 1, with regard to vol. For example check validity of stage amending of one's map from level $\boldsymbol{n}$ to level $\boldsymbol{n+1}$ if vol is $\boldsymbol{\theta}$ percent of this threshold.

RESOURCES

A. Madkour, T. Hefni, A. Hefny, and K. S. Refaat (2008), Using semantic features to detect spamming in social bookmarking systems. *Proc. of ECML PKDD Discovery Challenge Workshop, Belgium*

Alejandro Valerio, David B. Leake,, Alberto J. Cañas (2008), Automatic Classification Of Concept Maps Based On A Topological Taxonomy And Its Application To Studying Features Of Human-Built Maps. *Proc. of The Third Int. Conference On Concept Mapping,* Tallinn, Estonia & Helsinki, Finland

G. Koutrika, F. A. E_endi, Z. Gy ongyi, P. Heymann, and H. Garcia-Molina (2007), Combating spam in tagging systems. *Proc. of Int'l Workshop on Adversarial information retrieval on the web*, pages 57-64

John R. Mcclure, Brian Sonak, Hoi K. Suen (1999), Concept Map Assessment Of Classroom Learning: Reliability, Validity, And Logistical Practicality. *Journal Of Research In Science Teaching, Vol. 36, No. 4,* 475–492

Jun Zhang, Mark S. Ackerman, Lada Adamic (2007), Expertise Networks in Online Communities: Structure and Algorithms. *WWW 2007,* Banff, Alberta, Canada

M. G. Noll, C. Meinel (2008), Exploring social annotations for web document classi_cation. *Proc. of ACM Symposium on Applied Computing*, pages 2315-2320, Fortaleza, Brazil

Michael G. Noll, Ching-man Au Yeung, Nicholas Gibbins, Christoph Meinel, Nigel Shadbolt (2009), Telling Experts from Spammers:Expertise Ranking in Folksonomies. *Proc. of the 32nd international ACM SIGIR conference on Research and development in information retrieval*

P. Heymann, G. Koutrika, and H. Garcia-Molina (2007), Fighting spam on social web sites: A survey of approaches and future challenges. *IEEE Internet Computing*, 11(6):36 -45

P. J. Feltovich, M. J. Prietula, K. A. Ericsson (2006), Studies of expertise from psychological perspectives. *The Cambridge Handbook of Expertise and Expert Performance*, pages 41-68. Cambridge University Press, USA

R. Krestel, L. Chen (2008), Using co-occurrence of tags and resources to identify spammers. *In Proceedings of ECML PKDD Discovery Challenge Workshop, collocated with ECML/PKDD*

R. Wetzker, C. Zimmermann, C. Bauckhage (2008), Analyzing social bookmarking systems: A del.icio.us cookbook. *Proc. of Mining Social Data Workshop, collocated with ECAI,* pages 26-30

Steffen Schaal (2008), Concept Mapping In Science Education Assessment: An Approach To Computer-Supported Achievement Tests In An Interdisciplinary Hypermedia Learning Environment. *Proc. of The Third Int. Conference On Concept Mapping,* Tallinn, Estonia & Helsinki, Finland