Groupized Learning Path Discovery Based on Member Profile

Xiuzhen Feng[1], Haoran Xie[2], Yang Peng[3], Wei Chen[4], and Huamei Sun[5]

[1] Economics & Management School, Beijing University of Technology, Beijing, China
[2] Department of Computer Science, City University of Hong Kong, Hong Kong
[3] Department of Computing, The Hong Kong Polytechnic University, Hong Kong
[4] School of Computer Science & Technology, Beijing Institute of Technology Beijing, China
[5] School of Management, Harbin Institute of Technology, Harbin, China
xfeng@bjut.edu.cn, hrxie2@student.cityu.edu.hk, peterpeng204@gmail.com, wchen@bit.edu.cn, shuamei@hit.edu.cn

Abstract. With the explosion of knowledge nowadays, it is urgent for people to learn new things quickly and effectively. To meet such a requirement, how we can find a suitable path for learning has become a crucial issue. Meanwhile, in our daily life, it is important and necessary for people from various backgrounds to achieve a certain task (eg. survey, report, business plan, etc.) collaboratively in the form of the group. For these group-based task, it often requires members to learn new knowledge by using e-learning system. In this paper, we focus on addressing the problem on discovering an appropriate study path to facilitate a group of people rather than a single person for effective learning under e-learning environment. Furthermore, we propose a group model to capture the expertise of each member. Based on this model, a groupized[1] learning path discovering (GLPD) algorithm is proposed in order to help a group of learners to grasp new knowledge effectively and efficiently. Finally, we conduct a practical experiment whose result verifies the soundness of our approach.

1 Introduction

With the explosion of knowledge nowadays, it is urgent for people to learn new things quickly and effectively. To meet such a requirement, how we can find a suitable path for learning has become a crucial issue. Meanwhile, in our daily life, it is important and necessary for people from various backgrounds to achieve a certain task (eg. survey, report, business plan, etc.) collaboratively in the form of the group. For the group-based task, it often requires the whole group to learn a particular topic of knowledge effectively through e-learning. Hence, in this paper, we focus on addressing the problem on discovering an appropriate study path to facilitate a group of people rather than a single person for effective learning under e-learning environment. Compared with a single person, some characteristics of the group should be mapped in the group learning task.

[1] We use the term "groupized" to versus "personalized".

X. Luo et al. (Eds.): ICWL 2010 Workshops, LNCS 6537, pp. 301–310, 2011.

- Knowledge Diversity. A group in the real life or virtual environment often gets its members with various knowledge backgrounds and personal characteristics. For the group learning task in the e-learning system, diversity indicates that members might have different pre-knowledge and levels.
- Preference Variety. Members in a learning group may have different learning preferences when they are learning a particular new topic of knowledge. So that they would have the particular preference in different parts of a new topic.

Therefore, it is inadequate and unsuitable to use conventional shortest learning path selection [4] and personalized path generation [5] to solve this problem. To best of our knowledge, this is the first piece of work on group learning path discovery. The contributions of this paper are listed as follows.

- We propose a group model to capture the characteristic of group members.
- Based on the group model, a groupized learning path discovering (GLPD) algorithm is proposed to help a group of learners to grasp new knowledge effectively and efficiently.
- We conduct an experiment in two classes, which consist of learning groups, and the result verifies the soundness of our approach.

The rest of paper is structured as follows. In Section 2, a survey on related works is given. Subsequently, in Section 3 and 4, we introduce the proposed group model and groupized learning path discovering (GLPD) algorithm correspondingly. The process of how the experiment is conducted and its result are discussed in the Section 5. Finally, we conclude this work and discuss the possible future directions in the Section 6.

2 Related Works

For the past few years, many researchers have focused on developing the learning management system. Meanwhile, for the learning path discovering, the related researches could be divided into two areas; one is explicit, which means providing the learning method (eg., learning path, learning plan, etc.) to users directly, the other is implicit, which means offering the learning method through the learning materials (eg., courseware recommendation, courseware delivery, etc.). Before we introduce our GLPD approach, some related researches which focused on learning path delivery and courseware personalization for LMS are discussed as below:

Learning Path Delivery. Researches on learning path delivery are mainly focused on providing an appropriate learning path individually in order to improve the learning performance. In [1], Chen et al proposed a genetic-based personalized learning path generation scheme for individual learners and proved that it can promote learner's learning effectiveness during learning processes. An approach [2] on the use of concept maps for deriving prerequisite relations and structures based on CbKST has been generated by Steiner and Albert, with the purpose to achieve personalization in web-based learning. Madhour and Forte [3] presented the Lausanne Model and introduced the ACO algorithm which is based on the User model and other user's experience so as to provide the best-possible personalized learning path to users. Another recent work done by Zhao and Wan [4], in which an algorithm for selecting the shortest learning path to learn the target knowledge was proposed to save the time and efforts. Elvis and Li [5] established a dynamic conceptual network mechanism for personalized study plan generation, which can formulate different study plans for different students to meet their personalized needs.

Courseware Recommendation. Courseware recommendation have been widely adopted in current e-learning system to provide users with an implicity learning path. A personalized CAI courseware system [6] was introduced by Wei et al., and it can provide the same courseware tree to the students within the same group to improve the performance. Ge et al. [7] presented an algorithm in the courseware recommendation module, which combines content filtering (recommend from a single user information) and collaborative filtering (recommend from other users perspectives) to reflect the users full interests in courseware selection. Moreover, in the research done by Li et al. [8], a three-tier profiling framework has been proposed, in which course content is modeled by using concept nodes, and it could offer a unified way for modeling and handling a variety of student learning needs and the different factors that affect course material relevance.

3 Group Modeling

3.1 Topic Graph

For group-based learning, it usually needs members to learn a new topic of knowledge, which may contain several knowledge units. We formally define the **topic graph** to represent a new topic of knowledge as shown in Fig. 1, which is in the form of two elements tuple as follows.

Definition 1. A **topic graph** for a new topic of knowledge , denoted by TG, which is a two elements tuple as follows:

$$TG =< U, R >$$

where U is a set of weighted knowledge unit nodes in the graph and the weight t_{u_i} is the time to learn this unit and $R \subseteq U \times U$ is a relation set which represent the edges between the knowledge units. The relation indicates the learning sequence for the topic graph. R also can be represented as a matrix whose row and column representing all knowledge unit nodes and entries by using binary value to indicate whether there is a sequential relationship between two units or not(if there is a relationship, 1 is given. Otherwise, 0 is given).

People usually have different knowledge backgrounds and learning preferences when they are learning a new topic of knowledge. To acquire a learners' preknowledge and levels, there are mainly two kinds of strategies.

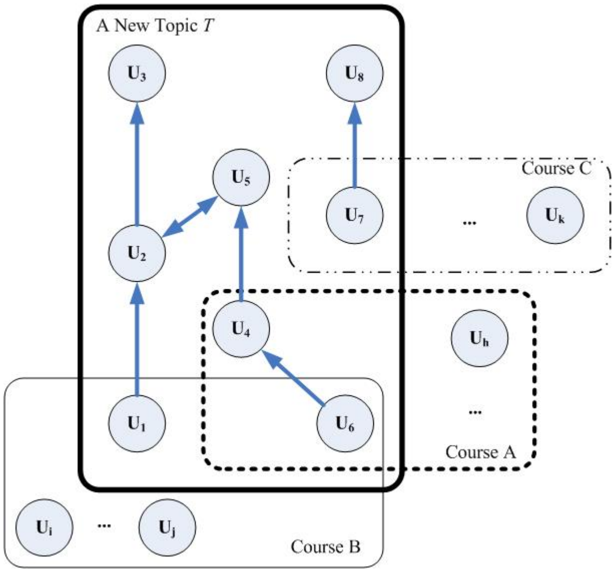

Fig. 1. The relationships among a new topics, past courses and knowledge units

3.2 Objective Pre-knowledge Level Acquisition

To some extend, a learner's objective data such as past course grades, pre-test scores may reflect his/her per-knowledge and levels for a new topic. A higher course grade may indicate the learner is in a higher level in related pre-knowledge. We use an **objective pre-knowledge level vector** to represent a learner's pre-knowledge levels, which are acquired by objective data.

Definition 2. An **objective pre-knowledge level vector** for a learner i, denoted by $\overrightarrow{L_i}$, which is a vector of pre-knowledge unit-value pairs as follows:

$$\overrightarrow{L_i} = (u_1 : v_{i,1}, u_2 : v_{i,2}, ..., u_n : v_{i,n})$$

where u_x is a pre-knowledge units[2] and n is the total number of pre-knowledge units for a new topic, $v_{i,x}$ denotes the level value of u_x of learner i. The higher value of $v_{i,x}$ is, the more knowledge of the learner has for this pre-knowledge unit.

To convert past course grades to the level value of $v_{i,x}$, we observe the relationships among a new topic, past courses and knowledge units as illustrated in Fig. 1. Some past courses may overlap with some knowledge units, which may reflect the level of a learner in these knowledge units. In Fig. 1, course A overlaps to a new topic T by unit 4 and 6, noted that a knowledge unit in topic TG might overlap with multiple courses (e.g. unit 6 overlaps with course B and A).

[2] We use terms "pre-knowledge units" for learner and it can be regarded "knowledge units" as well.

The level value $v_{i,x}$ for pre-knowledge units u_x can be obtained by a formula as follows:

$$v_{i,x} = \frac{\sum_{u_x \in K}^{\forall K} S_K}{\varepsilon \cdot N}$$

where K denotes the course containing knowledge unit u_x and S_K represents the course grade, N denotes the total number of courses and ε is normalized parameter to represent the length of course grade value range to normalized the value degree to $[0, 1]$ (e.g. $\varepsilon = 100$ for hundred mark system).

3.3 Subjective Pre-knowledge Level Acquisition

A learner's past grade may not reflect all his/her pre-knowledge completely, because a learner may be interested in some knowledge units that are not appeared in the course. Therefore, it is necessary for learners to specify their pre-knowledge and corresponding levels explicitly based on their own understanding. Similarly, we use a **subjective pre-knowledge level vector** to denote a leaner's pre-knowledge levels, which is acquired by his/her subjective feeling.

Definition 3. A **subjective pre-knowledge level vector** for a learner i, denoted by $\overrightarrow{L_i'}$, which is a vector of pre-knowledge unit-value pairs as follows:

$$\overrightarrow{L_i'} = (u_1 : v'_{i,1}, u_2 : v'_{i,2}, ..., u_n : v'_{i,n})$$

where u_x is a pre-knowledge units and n is the total number of pre-knowledge units for a new topic, $v'_{i,x}$ denotes the level value of u_x of learner i. The higher value of $v'_{i,x}$ is, the more knowledge of the learner has for this pre-knowledge unit.

3.4 Learning Preference Acquisition

For acquiring learning preferences from a learner, two possible ways are provided. The first one is letting the learner to specify a value [0,1] to indicate how much he/she is interested in past courses that overlap with this new topic of knowledge. The other is selecting the interested knowledge units explicitly by the learner himself. We use an **learning preference vector** to represent a learner's preferences for a new topic of knowledge.

Definition 4. A **learning preference vector** for a learner i, denoted by $\overrightarrow{P_i}$, which is a vector of pre-knowledge unit-value pairs as follows:

$$\overrightarrow{P_i} = (u_1 : w_{i,1}, u_2 : w_{i,2}, ..., u_n : w_{i,n})$$

where u_x is a pre-knowledge units and n is the total number of pre-knowledge units for a new topic, $w_{i,x}$ denotes the preference degree of u_x of learner i. The higher value of $w_{i,x}$ is, the more preference of the learner has for this knowledge unit. The $w_{i,x}$ can be obtained as follows.

$$w_{i,x} = \begin{cases} 1, & \text{once } u_x \text{ is selected explicitly} \\ \frac{\sum_{u_x \in K}^{\forall K} D_K}{N}, & \text{if } u_x \text{ is contained by some courses} \\ 0, & \text{otherwise} \end{cases}$$

where u_x is a knowledge units and n is the total number of knowledge units for a new topic, K denotes the course containing knowledge unit u_x and D_K represents the learner specify value of interesting degree for this course, N denotes the total number of courses. Note that the $w_{i,x}$ will be given as "1" once it is selected by the learner explicitly. Because it is quite sure that the learner has high preference on this units instead of other methods.

3.5 Member Profile

We aggregate subjective, objective pre-knowledge level vectors into **unified pre-knowledge level vector**. Then we defined **member profile** to represent the complete pre-knowledge and levels for a member (learner) and his/her learning preference in a group.

Definition 5. A **member profile** for a member (learner) i, denoted by $\overrightarrow{M_i}$, which is a tuple containing two vectors as follows:

$$\overrightarrow{M_i} = (\overrightarrow{L_i''}, \overrightarrow{P_i})$$

$$\overrightarrow{L_i''} = (u_1 : v''_{i,1}, w_{i,1}; u_2 : v''_{i,2}, w_{i,2}; ..., u_n : v''_{i,n}, w_{i,n})$$

where $v''_{i,x}$ denotes the level value of u_x of member i, $\overrightarrow{P_i}$ is the learning preference vector. The $v''_{i,x}$ is obtained by the aggregating subjective and objective pre-knowledge level values as follows.

$$v''_{i,x} = \begin{cases} v'_{i,x}, & \text{if } v_{i,x} = 0 \text{ and } v'_{i,x} \neq 0 \\ v_{i,x}, & \text{if } v'_{i,x} = 0 \text{ and } v_{i,x} \neq 0 \\ \alpha \cdot v_{i,x} + (1-\alpha) \cdot v'_{i,x}, & \text{otherwise} \end{cases}$$

where α is a parameter to adjust the effect of the level values from subjective and objective pre-knowledge vectors, $v''_{i,x}$ is the linear combination of level values from subjective and objectives pre-knowledge vectors except one of them is zero.

A learning group is consisted of some group members and we defined the **learning group** based on its members.

Definition 6. A **learning group**, denoted by G, which are two $i \times n$ matrix of member profiles as follows:

$$G = A_{i\times n}, B_{i\times n} = \begin{pmatrix} 0.5 & 0 & \ldots & 0.2 \\ 0.3 & 0.1 & \ldots & 0.7 \\ \vdots & \vdots & \vdots & \vdots \\ 0.9 & 0.4 & \ldots & 0 \end{pmatrix}, \begin{pmatrix} 0.2 & 0 & \ldots & 0.2 \\ 0.3 & 0 & \ldots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 0.4 & 1 & \ldots & 0 \end{pmatrix}$$

where $A_{i\times n}$ denotes all group members pre-knowledge level matrix and $B_{i\times n}$ denotes their learning preference matrix correspondingly.

4 Groupized Learning Path Discovery

In this section, we devise the groupized learning path discovery (GLPD) algorithm to find the suitable path for a group. GLDP approach mainly considers two levels of aims. The first one is the **group-based aim**, which requires that the whole group rather than a single member can grasp the new topic efficiently. In other words, it means the union of the grasped knowledge units for all group members should be as many as possible. The other one is the **member-based aim**, which is to select a suitable learning path by taking his/her preferences as more as possible.

For example, given a group containing two members Tom and Kate, Tom is a programmer and Kate is a business school graduate. If they need to learn a new topic "business information system", the most efficient way may be let Tom to learn technical units and Kate to learn business units. However, Tom might have more preference to learn units related to management and Kate is more interested in units on the decision-making support system. The group-based aim is focused on the efficiency aspect for the whole group while the member-based aim is focused on the interest aspect for the individual member. In addition, we observe these two aims associated with the learning time. Back to our example, if the group is given limited time to learn, efficiency needs to be considered, firstly; but if enough learning time is provided, preferences could be taken into account, since no matter how their path is selected, the group-based aim can be achieved.

Input: Matrix $A_{i \times n}$, and vector C
Output: T_{Lower}, $InitialPath_i$ for member i, and cost Matrix $CM_{i \times n}$
for $j = 1$; $j \leq C.length$; $j++$ **do**
  for $k = 1$; $k \leq n$; $k++$ **do**
    for $s = 1$; $s \leq i$; $s++$ **do**
      $a[s,k] = a[s,k] \times t_{u_j}$;
      $CM_[i,n] = a[s,k]$;
    end
  end
end
for $k = 1$; $k \leq n$; $k++$ **do**
  for $s = 1$; $s \leq i$; $s++$ **do**
    $Value_s = a[s,k] + RowValue_s$;
  end
  if $Value_s == Min(Value_1 \ to \ Value_i)$ **then**
    $RowValue_s = a[s,k] + RowValue_s$;
    Add u_k to $InitialPath_s$;
  end
end
$T_{Lower} = \text{Max}(RowValue_1$ to $RowValue_s)$ $InitialPath_i$ is a sequence of knowledge units. $CM_{i \times n}$ is initialized.

Algorithm 1. Lower Time Boundary Discovery

Therefore, our GLDP approach includes two major steps. Firstly, we discover two temporal boundaries for a learning group. Then, according to learning time for the group, a corresponding strategy is to select to find the suitable learning path.

Input: T,T_{Lower}, T_{Upper},$InitialPath_i$ for member i, cost Matrix $CM_{i\times n}$, $B_{i\times n}$, Relations Matrix $R_{n\times n}$ in topic graph
Output: $FinalPath_i$ for member i
if $T \le T_{Lower}$ **then**
  for $j = 1$*;* $j \le InitialPath_i.length$*;* $j++$ **do**
    for *Each two adjacent nodes* u_a *and* u_{a+1} *in* $InitialPath_i$ **do**
      if *entry* $R[a+1, a] == 1$ *and* $R[a, a+1] == 0$ **then**
        $u_a \leftrightarrow u_{a+1}$ in $InitialPath_i$;
      end
    end
  end
end
if $T_{Lower} < T < T_{Upper}$ **then**
  for $j = 1$*;* $j \le InitialPath_i.length$*;* $j++$ **do**
    while $T_{Lower} + T_{Cost} < T$ **do**
      for *Each node in* u_a *in* $InitialPath_j$ **do**
        Compare u_a with u_1 to u_n;
        if $r[a, b] == 1$ **then**
          Find Max u_b in j-th row in $B_{i\times n}$;
          Add u_b to $InitialPath_j$;
          $T_{Cost} = cm[i, b] + T_{Cost}$;
        end
      end
    end
  end
end
if $T \ge T_{Upper}$ **then**
  for *Each knowledge unit* u_a *to Memberi* **do**
    if $b[a, i]$ *is maximal value in row* i *in* $B_{a,i}$ *and* $b[a, i] == 1$ **then**
      Add u_a to $InitialPath_i$;
    end
  end
end
$FinalPath_i = InitialPath_i$

Algorithm 2. Learning Path Discovery

4.1 Temporal Boundaries Discovery

For the first step, we need to discover two temporal boundaries for different groups. We name the maximum time for every member to learn all knowledge units as "upper boundary time" (denoted by T_{Upper}), and we name the minimum time for the whole group to grasp all knowledge units as "lower boundary time" (denoted by T_{Lower}). To find out T_{Upper} and T_{Lower}, the time cost for each knowledge units is needed, and we name the **time cost vector** to represent the time cost for each unit in a topic graph. We use $C = (t_{u_1}, t_{u_2}, ...t_{u_n})$ to denote time cost vector, where t_{u_x} is defined in **Definition 1** to represent time cost for learning unit u_x. T_{Lower} can be obtained by Algorithm 1 and T_{Upper} is calculated by the following formula.

$$T_{Upper} = max((J_{i\times n} - A_{i\times n}) \times C^T)$$

where $J_{i\times n}$ is a matrix of ones, $max()$ denotes the maximum value in a vector, C^T is the transpose of C.

4.2 Learning Path Discovery

Given a learning time T, we compare it with T_{Lower} and T_{Lower} to determine which strategy is to use. There are three cases as follows and their algorithms are unified in Algorithm 2.

1. $T \leq T_{Lower}$. With limited time, we mainly consider pre-knowledge level and relations in the knowledge units to achieve the group-based aim firstly.
2. $T \geq T_{Upper}$.If the time is adequate to finish both two aims, the member preference is most important aspect for learning path discovery.
3. $T_{Lower} < T < T_{Upper}$. If the time is enough to finish the group-based aim, the member preference can be also considered partially.

5 Experiment

We conduct two experiments on students in two classes. The number of students is 32 (Class I) and 29 (Class II). Since it is an elective course that is open for all year 2 to year 4 undergraduate students in the university from different majors, it is suitable for illustrating our problem. Then we let 3 or 4 students to form a group so that we get 8 groups in Class I and 7 groups in Class II. Afterward, we give a topic of "business decision support system" which includes about 6 knowledge units, and it would take about 5 hours to digest them normally. We give them 1 hour to learn and then give a test that includes 20 multiple choice questions across all the points in 5 units to each group. For Class I, we generate the learning path for them according to each group member profile. For Class II, they learn by their own decision. The second test is that we provide another topic, which is consisted of 7 units and required 8 hours to learn, and we give 4 hours for each group. The setting is the same with the previous one. In addition, we test each group member separately and collect the right answers to regard as the group mark. The experiment results are shown in Table I. From results, the Class I in the three tests have always outperformed Class II. These results have verified the soundness of our approach.

Table 1. Average grade results in Test 1 and 2

	Test 1	Test 2 (Group)	Test 2 (Member)
Class I	67.3	80.6	72.3
Class II	58.4	74.2	63.0

6 Conclusion

In this paper, we have proposed the groupized learning path discovery (GLPD) algorithm, which is based on the group model. It offers a shortcut for a group of people to learn new knowledge under the e-learning environment as well as meets the different preference of each group member in his learning process. Different with conventional learning path discovery approaches, what we offered is a suite of learning paths for group learning according to different learning time limits. Through the experiments, the soundness of our approach is also verified. For future works, one of the possible directions is that identifying some learning style patterns for group to explore how to discover learning path in different learning styles.

Acknowledgement

The research described in this paper has been supported primarily by National Natural Science Foundation of China (Project 70971032).

References

1. Chen, C.-M.: Intelligent web-based learning system with personalized learning path guidance. Computer & Education 51, 787–814 (2008)
2. Steiner, C.M., Albert, D.: Personalising learning through prerequisite structures derived from concept maps. In: Leung, H., Li, F., Lau, R., Li, Q. (eds.) ICWL 2007. LNCS, vol. 4823, pp. 43–54. Springer, Heidelberg (2008)
3. Madhour, H., Forte, M.W.: Personalized learning path delivery: Models and example of application. In: Woolf, B.P., Aïmeur, E., Nkambou, R., Lajoie, S. (eds.) ITS 2008. LNCS, vol. 5091, pp. 725–727. Springer, Heidelberg (2008)
4. Zhao, C., Wan, L.: A shortest learning path selection algorithm in e-learning. In: ICALT 2006 (2006)
5. Leung, E.W.C., Li, Q.: A Dynamic Conceptual Network Mechanism for Personalized Study Plan Generation. In: Zhou, W., Nicholson, P., Corbitt, B., Fong, J. (eds.) ICWL 2003. LNCS, vol. 2783, pp. 69–80. Springer, Heidelberg (2003)
6. Wei, C., Fang, Z., Zhang, Y., Jin, Y.: A personalized CAI courseware system. In: CISE 2009 (2009)
7. Liang, G., Weining, K., Junzhou, L.: Courseware Recommendation in E-Learning System. In: Liu, W., Li, Q., Lau, R. (eds.) ICWL 2006. LNCS, vol. 4181, pp. 10–24. Springer, Heidelberg (2006)
8. Li, F.W.B., Lau, R.W.H., Dharmendran, P.: A Three-Tier Profiling Framework for Adaptive e-Learning. In: Spaniol, M., Li, Q., Klamma, R., Lau, R.W.H. (eds.) ICWL 2009. LNCS, vol. 5686, pp. 235–244. Springer, Heidelberg (2009)