FTGWS: Forming Optimal Tutor Group for Weak Students Discovered in Educational Settings

Yonghao Song[1,2], Hengyi Cai[1,2], Xiaohui Zheng[1,2], Qiang Qiu[1], Yan Jin[1], and Xiaofang Zhao[1](✉)

[1] Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China
{songyonghao,caihengyi,zhengxiaohui,qiuqiang,zhaoxf}@ict.ac.cn, jinyan@ncic.ac.cn

[2] University of Chinese Academy of Sciences, Beijing, China

Abstract. The task of experts discovering, as one of the most important research issues in social networks, has been widely studied by many researchers in recent years. However, there are extremely few works considering this issue in educational settings. In this work, we focus on the problem of forming tutor group for weak students based on their knowledge state. To solve this problem, a novel framework based on Student-Skill Interaction (SSI) model and set covering theory is proposed, which is called FTGWS. The FTGWS framework contains three major steps: firstly, building SSI models for each student and each skill he or she has encountered; then, discovering the top-k weak students based on their knowledge state; finally, forming the optimal tutor group for each weak student. We evaluate our framework on a real-word dataset which contains 28834 students and 244 skills. The experiments show that the framework is capable of producing high-quality solutions (for 93% of weak students, the size of the optimal tutor group can be decreased up to 2 students).

Keywords: Tutor group · Grouping students · Weak student · Cooperative learning · Student-Skill Interaction Model (SSI)

1 Introduction

With the booming popularity of web-based educational settings, such as Coursera, Khan Academy, and ASSISTment, e-learning has attracted much attention of educators, governments and the general public [1]. E-learning aims to make high quality online learning resources to the world, and has attracted a diverse population of students from a variety of age groups, educational backgrounds and nationalities [3]. Despite these successes, providing high quality online education is a multi-faceted and complex system [2]. Two particular problems that have vexed researchers and educators for a long time are how to identify students who are at risk of poor performance early and how to create tutor groups for

D. Benslimane et al. (Eds.): DEXA 2017, Part I, LNCS 10438, pp. 436–444, 2017.
DOI: 10.1007/978-3-319-64468-4_33

these weak students so that they can augment their knowledge with cooperative learning from each other [3–6].

In this research work, we explore how to identify weak students and how to form the optimal tutor group for weak students based on their knowledge state which is described as two skill sets. The formal definition of these two problems will be given in preliminary section. If a set of students that together have all of the required skills which a weak student has not mastered, through the cooperative learning between the weak student and this set of students can improve the performance of this weak student [5,8]. This set of students is defined as a tutor group of one specific weak student. Based on above idea, a FTGWS framework is proposed, where weak students are discovered based on their interaction records in system and the optimal tutor group will be generated based on students knowledge state. Our main contributions can be summarized as follows:

(1) We give the formal definitions of difficulty of skills and learning rate of students, then the concept of weak students is defined, and a algorithm called FKWS to discover top-k weak students has been designed.
(2) We introduce the formal definition of tutor group and convert the problem of forming optimal tutor group to the minimum set cover problem (SCP) which has been proved a NP-hard problem; a heuristic algorithm based on genetic algorithm is implemented to solve this problem.
(3) Extensive experiments on the real world data set[1] which contains 28834 students and 244 skills are carried out. The experimental results are capable of producing high-quality solutions (for 93% weak students, the size of the optimal tutor group can be decreased up to 2 students).

2 Preliminaries

2.1 Notations and Definitions

The mathematical denotations throughout this paper are listed in Table 1.

Definition 1 ***(Difficulty Coefficient of Skill).*** *Given a skill k_j, a set of students $S^j = \{s_1^j, s_2^j, \ldots, s_m^j\}$ who have exercised k_j and the matrix $SK_{M\times N}$. The difficult coefficient of skill k_j is*

$$d_j = 1 - \frac{\sum_{s_i^j \in S^j} P_{i,j}(T)}{||S^j||} \tag{1}$$

Definition 2 ***(Learning Ability of Student).*** *Given a student s_i, a set of skills $K^i = \{k_1^i, k_2^i, \ldots, k_n^i\}$ which s_i has exercised and the matrix $SK_{M\times N}$. The learning ability of student s_i is*

$$l_i = \frac{\sum_{k_j^i \in K^i} P_{i,j}(T)}{||K^i||} \tag{2}$$

[1] https://sites.google.com/site/assistmentsdata/home/2012-13-school-data-with-affect.

Table 1. Mathematical notations used in this paper

Notation	Description
M, N, D	Number of students, number of skills and data set, respectively
$S = \{s_1, s_2, \ldots, s_i, \ldots, s_M\}$	Set of students where s_i is the student i
$W = \{s_1, s_2, \ldots, s_i, \ldots, s_k\}$	Set of students where s_i is a poor performance student, $W \subseteq S$
$K = \{k_1, k_2, \ldots, k_j, \ldots, k_N\}$	Set of skills where k_j is the skill j
$K^i = \{k_1^i, k_2^i, \ldots, k_n^i\}$	Set of skills which s_i has exercised
$S^j = \{s_1^j, s_2^j, \ldots, s_m^j\}$	Set of students who have exercised k_j
$R^{i,j} = r_1^{i,j} r_2^{i,j} \ldots r_l^{i,j}$	Response sequence of s_i on k_j, e.g. 01001011111
$P_{i,j}(L_0)$	Probability that s_i masters skill k_j initially
$P_{i,j}(T)$	Probability that s_i transforms k_j from unlearned state to learned
$P_{i,j}(G)$	Probability that s_i guesses correctly on k_j
$P_{i,j}(S)$	Probability that s_i slips (make a mistake) on k_j
$s_i k_j$	SSI model of s_i for k_j, it is a four-tuple: $\{P_{i,j}(L_0), P_{i,j}(T), P_{i,j}(G), P_{i,j}(S)\}$
$SK_{M \times N}$	Matrix formed based on SSI model where $SK_{ij} = s_i k_j$
$DMap\langle k_j, d_j \rangle$	Collection contains coefficient of difficulty d_j of skill k_j
$LMap\langle s_i, l_i \rangle$	Collection contains learning rate l_i of student s_i
MS_{s_i}	Mastered skill set of student s_i
TS_{s_i}	Target skill set of student s_i
TG_{s_i}	Optimal tutor group for s_i

Definition 3 ***(Mastered Skill).*** *Given a student s_i, a skill k_j, a SSI model $s_i k_j = \{P_{i,j}(L_0), P_{i,j}(T), P_{i,j}(G), P_{i,j}(S)\}$, a response sequence $R^{i,j} = r_1^{i,j} r_2^{i,j} \ldots$ and a determining factor e. Let $n = ||R^{i,j}||$. If the following condition is satisfied, then k_j is a mastered skill of s_i.*

$$\begin{aligned}
P_{i,j}(L_{n-1}|r_{n-1}^{i,j} = 1) &= \frac{P_{i,j}(L_{n-1}) * (1 - P_{i,j}(S))}{P_{i,j}(L_{n-1}) * (1 - P_{i,j}(S)) + (1 - P_{i,j}(L_{n-1})) * P_{i,j}(G)} \\
P_{i,j}(L_{n-1}|r_{n-1}^{i,j} = 0) &= \frac{P_{i,j}(L_{n-1}) * P_{i,j}(S)}{P_{i,j}(L_{n-1}) * P_{i,j}(S) + (1 - P_{i,j}(L_{n-1})) * (1 - P_{i,j}(G))} \\
P_{i,j}(L_n) &= P_{i,j}(L_{n-1}) + (1 - P_{i,j}(L_{n-1})) * P_{i,j}(T)
\end{aligned} \tag{3}$$

and

$$P_{i,j}(L_n) \geq e \tag{4}$$

Definition 4 ***(Target Skill).*** *Given a student s_i, a skill k_j, and a determining factor ε, obtaining the coefficient of difficulty d_j of skill k_j from $DMap\langle k_j, d_j \rangle$, and obtaining the learning rate l_i of student s_i from $LMap\langle s_i, l_i \rangle$. If the following condition is satisfied, then k_j is a target skill of s_i.*

$$l_i \geq \varepsilon d_j \tag{5}$$

2.2 Problems Formulation

The major tasks of this research are discovering weak students who are at risk of poor learning performance, and seeking out the optimal tutor group for each of them based on students interaction records in e-learning system. Based on the notations and definitions provided, the problems to be solved in this paper are formulated as follows.

Problem 1 (Discovering Top-K Poor Performance Students, FKWS). Given a student s_i, the mastered skill set MS_{s_i} of s_i, and the target skill set TS_{s_i} of s_i, the function $f(s_i, Performance)$ is used to calculate the performance score of s_i.

$$f(s_i, Perf) = \frac{||MS_{s_i}||}{||TS_{s_i}||} \tag{6}$$

Based on Eq. 6, the top-k poor performance student can be sought out. Before the definition of forming tutor groups for weak students, the tutor skill set is defined as follows.

Definition 5 ***(Tutor Skill Set).*** *Given a weak student w_i and the other students set $S = \{s_1, s_2, \ldots, s_i, \ldots, s_{M-1}\}$ where $w_i \notin S$. Given mastered skill set MS_{w_i} and target skill set TS_{w_i} of w_i, and other students mastered skill sets $\{MS_{s_1}, MS_{s_2}, \ldots, MS_{s_i}, \ldots, MS_{s_{M-1}}\}$. Let $UMS_{w_i} = TS_{w_i} - MS_{w_i}$ and $I_i = TS_{w_i} \cap MS_{s_i}$, the tutor skill set of w_i is*

$$TutorSet_{w_i} = \cup_{i=1}^{M-1} I_i \ (I_i \in UMS_{w_i}) \tag{7}$$

Problem 2 (Forming the optimal tutor group for weak students, FTGWS). Given a weak student w_i and the other students $S = \{s_1, s_2, \ldots, s_i, \ldots, s_{M-1}\}$ where $w_i \notin S$. Given mastered skill set MS_{w_i} and target skill set TS_{w_i} of w_i, and other students mastered skill sets $\{MS_{s_1}, MS_{s_2}, \ldots, MS_{s_i}, \ldots, MS_{s_{M-1}}\}$. Based on Definition 5, the problem of forming the optimal tutor group for weak student w_i is to find a student set $S_{w_i} \subset S$, where

$$\cup_{s_i \in S_{w_i}} MS_{s_i} = TutorSet_{w_i} \, and \, S_{w_i} = argmin|S_{w_i}| \tag{8}$$

3 FTGWS Framework

In this section, the FTGWS framework is presented in detail. The design of our algorithm is inspired by a simple idea: each skill has an inherent difficulty and each student has an inherent learning ability, based on these hypotheses, the target skill set and the mastered skill set of each student can be obtained from this student's interaction records on skills. The criterion of poor performance students is defined according to these skill sets; then, the optimal tutor group can be formed for each weak student which is formulated by Problem 2 that is converted to minimum set cover problem (SCP). Based on the above idea, the FTGWS framework that employs SSI model and genetic algorithm is proposed, which contains three major steps (the pseudo code is shown in Algorithm 1).

To understand the work mechanism of FTGWS scheme, we give an illustrative example in Fig. 1, and each step of FTGWS is introduced in the following subsections.

Algorithm 1. FTGWS framework of forming tutor group for weak students

Initial Step: initialing the variables will be used in the following steps
  Obtain the set of students S and the set of skills K from D
  Obtain the response sequence $R^{i,j}$ from D
Step 1: Learning SSI Model for Each Student and Each Skill $SK_{M\times N}$
  $SK_{M\times N} \leftarrow LSSI(S, D, d)$ // d is the threshold of stopping learning
Step 2: Discovering Top-K Poor Performance Student W
  $W \leftarrow FKWS(SK_{M\times N}, S, D, k)$ // k is the number of weak students who are top-k
Step 3: Forming the optimal tutor group for each weak student
  $TG \leftarrow FOTG(S, W, MS, TS)$
Finalized Step:
  return TG

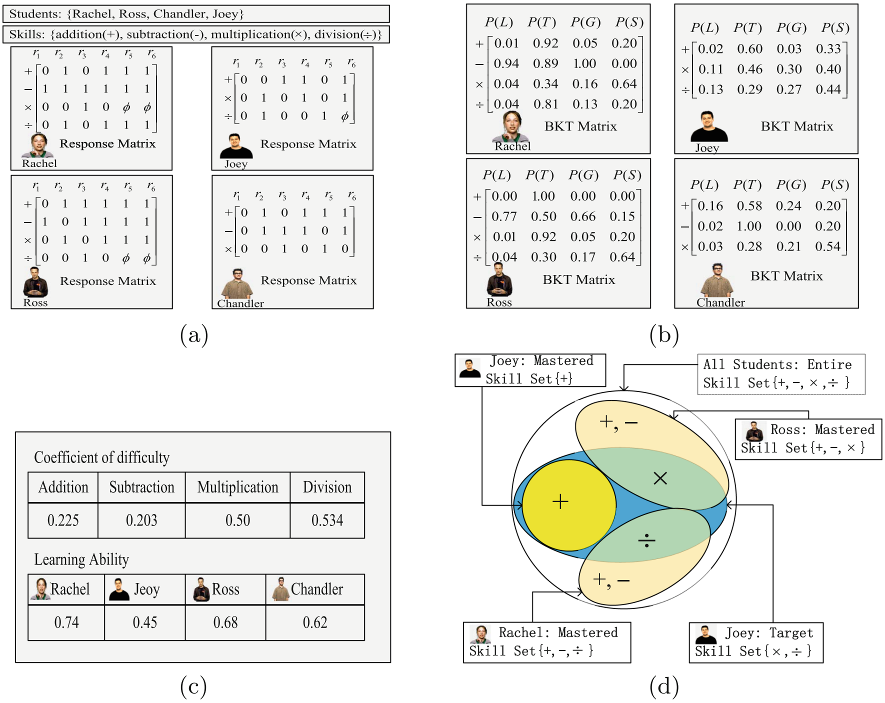

Fig. 1. The illustrative example of FTGWS framework

3.1 Learning SSI Model for Each Student and Each Skill

The Student-Skill Interaction (SSI) model proposed by Pardos & Heffernan is expanded based on standard BKT model which is a simple hidden markov model (HMM) [7,10]. The first step of SSI model is to learn student specific parameters by training all skill data of an individual student. The second step is to embed all students' specific parameter information which obtained from first step into

SSI model. The classical Baum-Welch algorithm is used to find the unknown parameters of a HMM.

Figure 1(a) shows that the entire skill set of student Rachel is $\{+, -, \times, \div\}$, for each of them, Rachel has a response sequence which obtains from interaction records by chronological order. For instance, the response sequence of Rachel on addition skill is $r_1 r_2 \ldots r_6 = [010111]$. The individual initial knowledge of Rachel is 15/22 for all skills. As showed in Fig. 1(b), the learning rate, guess rate and slip rate on addition is 0.92, 0.05, 0.20 respectively, which are learnt by SSI model.

3.2 Discovering Top-K Poor Performance Student

According to the definitions in the preliminary section, we utilize algorithm FTWS to discover top-k poor performance students. Firstly, the learning rate l_i for a specific student s_i and the difficulty of a specific skill k_j can be calculated. Next, the mastered skill set MS_{s_i} and the target skill set TS_{s_i} can be obtained based on Definitions 3 and 4. Lastly, we calculate the score of performance for each student and find the top-k weak students.

As shown in Fig. 1(c), the difficulty of $\{+, -, \times, \div\}$ is {0.225, 0.203, 0.50, 0.534} in which division is the hardest skill, and the learning rate of {Rachel, Joey, Ross, Chandler} is {0.74, 0.45, 0.68, 0.62} where Rachel has the best learning ability. Figure 1(d) illustrates that the top-1 weak student is Joey with mastered skill set $\{+\}$ and target skill set $\{+, \times, \div\}$, whose score of performance is 0.33 derived from Eq. 6.

3.3 Forming the Optimal Tutor Group for Weak Students

Now that the weak students have been discovered, the optimal tutor group needs to be formed for each weak student to augment their knowledge. Generally, for a weak student w_i, if $(TS_{w_i} - MS_{w_i}) \subset \cup_{s_i \in S'} MS_{s_i}$ where $S' \subset S$, the student set S' is a tutor group of w_i, we define the tutor group with the minimal size as the optimal tutor group. Based on the formal description of Problem 2 in preliminary section, the FTGWS problem is a minimum set cover problem which has been proved to be a NP-hard problem. In this paper, we employ a heuristic algorithm proposed by Beasley & Chu which is based on genetic algorithm to solve FTGWS problem [9]. The result of experiment shows that this heuristic algorithm is capable of producing high-quality solutions.

Figure 1(d) shows that the optimal tutor group obtained from FOTG algorithm is {Rachel, Ross} for weak student Joey. The mastered skill sets of Rachel and Ross are $\{+, -\div\}$ and $\{+, -, \times\}$, the tutor skill set of Joey is $TS_{Joey} - MS_{Joey} = \{+, \times, \div\}$ which can be covered by the mastered skill sets of Rachel and Ross.

4 Experiments

In this section, the proposed FTGWS framework is evaluated on the real-world data set assistments_2012_2013 published by ASSISTment platform. Specifically, we show and analyze the result of every step of FTGWS framework, which includes the coefficient of difficulty of skills, the learning rate of students and the optimal tutor group.

4.1 Experimental Data

The ASSISTments data set contains 46674 students, 265 skills and 4 problem types which are choose_1, algebra, fill_in and open_response. We preprocessed the data set by deleting the records in which skill_id is null and problem type is open_response, for the reason that open response problem is always marked as correct. The final experimental data set contains 28834 students and 244 skills.

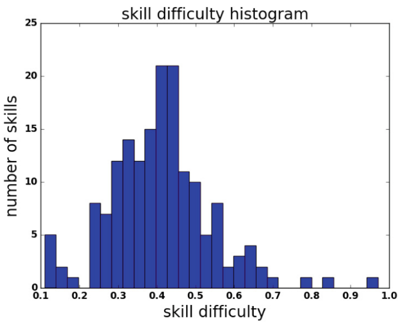

Fig. 2. DIST of skills difficulty

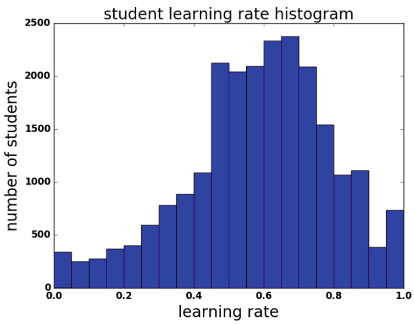

Fig. 3. DIST of students learning rate

4.2 Experimental Results and Analysis

Coefficient of Difficulty of Skills. In this group of experiments, the coefficient of difficulty for all 244 skills in the dataset were calculated based on Definition 1. Figure 2 shows that difficulty coefficients of most skills are less than 0.7, which represents these skills are relatively simple; the difficulty of skills follow the normal distribution which verified the rationality of Definition 1.

Learning Rate of Students. In this group of experiments, the learning rates of all 28834 students were calculated based on Definition 2. The student with higher learning rate tends to have better learning ability. Figure 3 shows that the mean learning rate of most students is 0.6 which indicates that most students has a normal learning ability, overall, the distribution of students learning rate fits the normal distribution represents that there are fewer prominent students or backward students.

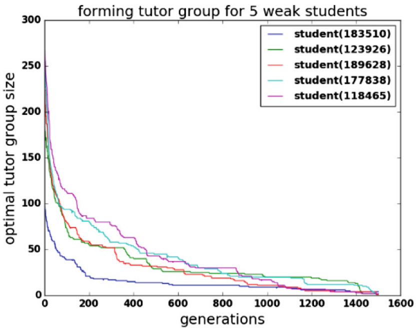

Fig. 4. Forming optimal tutor group

Optimal Tutor Group. The convergence and the stability of FOTG algorithm are evaluated and optimal tutor groups for top-100 weak students have been formed. Figure 4 demonstrates the iteration processes of FOTG algorithm, for all 5 weak students the size of their optimal tutor group can be converged to less than 3, which means the mastered skill sets of 3 students can cover the target skill set of one weak student.

5 Conclusion and Future Work

This paper proposed a novel FTGWS framework to form the optimal tutor group for weak students discovered in educational settings, which is based on BKT model and SCP theory. There are several possibilities to extend the research in the future. First, due to the high complexity of FOTG algorithm, a more effective substitutable algorithm needs to be designed to reduce the complexity of forming tutor group. Second, the FTGWS framework is not sufficiently sophisticated, an excellent student who is good at many skills maybe appears in every tutor group, this unbalance problem will be solved in the future work.

References

1. Anderson, A., Huttenlocher, D., Kleinberg, J.: Engaging with massive online courses. In: WWW, pp. 687–698 (2014)
2. Gillies, J., Quijada, J.: Opportunity to learn: a high impact strategy for improving educational outcomes in developing countries. In: USAID EQUIP (2008)
3. He, J., Bailey, J., Rubinstein, B., Zhang, R.: Identifying at-risk students in massive open online courses. In: AAAI, pp. 1749–1755 (2015)
4. Lakkaraju, H., Aguiar, E., Shan, C.: A machine learning framework to identify students at risk of adverse academic outcomes. In: KDD, pp. 1909–1918 (2015)
5. Agrawal, R., Golshan, B., Terzi, E.: Grouping students in educational settings. In: KDD, pp. 1017–1026 (2014)
6. Kim, B.W., Chun, S.K., Lee, W.G., Shon, J.G.: The greedy approach to group students for cooperative learning. In: Park, J., Yi, G., Jeong, Y.S., Shen, H. (eds.) UCAWSN & PDCAT 2016. LNEE, vol. 368, pp. 83–89. Springer, Singapore (2016). doi:10.1007/978-981-10-0068-3_10
7. Pardos, Z.A., Heffernan, N.T.: Modeling individualization in a Bayesian networks implementation of knowledge tracing. In: De Bra, P., Kobsa, A., Chin, D. (eds.) UMAP 2010. LNCS, vol. 6075, pp. 255–266. Springer, Heidelberg (2010). doi:10.1007/978-3-642-13470-8_24
8. Compton, J.I., Forbes, G.R.: Modeling success: using preenrollment data to identify academically at-risk students. In: Education Publications, No. 37 (2015)

9. Beasley, J.E., Chu, P.C.: A genetic algorithm for the set covering problem. Eur. J. Oper. Res. **94**, 392–404 (1996)
10. Song, Y., Jin, Y., Zheng, X., Han, H., Zhong, Y., Zhao, X.: PSFK: a student performance prediction scheme for first-encounter knowledge in ITS. In: Zhang, S., Wirsing, M., Zhang, Z. (eds.) KSEM 2015. LNCS, vol. 9403, pp. 639–650. Springer, Cham (2015). doi:10.1007/978-3-319-25159-2_58