Forgetting Punished Recommendations for MOOC

Yanxia Pang[1,2](✉), Liping Li[2], Wenan Tan[2], Yuanyuan Jin[1], and Ying Zhang[1]

[1] East China Normal University, Shanghai, China
yxpang@sspu.edu.cn
[2] Shanghai Polytechnic University, Shanghai, China

Abstract. Prerequisite inadequacy tends to cause more drop-out of MOOC. Recommendation is an effective method of learning intervene. Existing recommendation for MOOC is mainly for subsequent learning objects that have not been learned before. This paper proposes a solution called Forgetting-punished MOOC Recommendation (FMR). FMR combines the forgetting effect on learning score as a main feature for recommendation. It provides Prerequisite Recommendation (PR) for the unqualified learning objects and Subsequent Recommendation (SR) for the qualified objects. Experiments verify the accuracy improvement of PR and SR.

Keywords: MOOC · Recommendation · Prerequisite · Subsequent Location

1 Introduction

MOOC (Massive Open Online Course) develops rapidly in recent years, but the drop-out rate reaches 90% [1]. Kizilcec found that frustration is an important factor affecting learners' persistence in learning [2]. Pappano believes that MOOC learners are often frustrated for the inadequacy of prerequisite. The learner fails to keep pace and tends to drop out [3].

Prerequisite relationship between learning objects plays an important role for MOOC learning. Recommendation can effectively guide learners to learn. It is an effective mean to intervene in MOOC learning.

MOOC platforms pay effort on prerequisite for better learning. Figure 1 shows the learning content of math subjects on Khan Academy (https://www.khanacademy.orgn) which is one of the most popular MOOC platform. Usually, learners learn in order one by one. The previous knowledge provides prerequisites for further learning. Coursera (https://www.coursera.org) lists the prerequisite in course introduction. Khan Academy (https://www.khanacademy.org) lists the subject of the course according to the grade level of the target learners. Learners are asked to have a test. By this way, a suitable starting point will be found for them. But the MOOC platforms do not provide personalized recommendation on prerequisite. Existed MOOC recommendation is mainly about learning objects that were not learned before.

X. Chen et al. (Eds.): CSoNet 2018, LNCS 11280, pp. 415–426, 2018.
https://doi.org/10.1007/978-3-030-04648-4_35

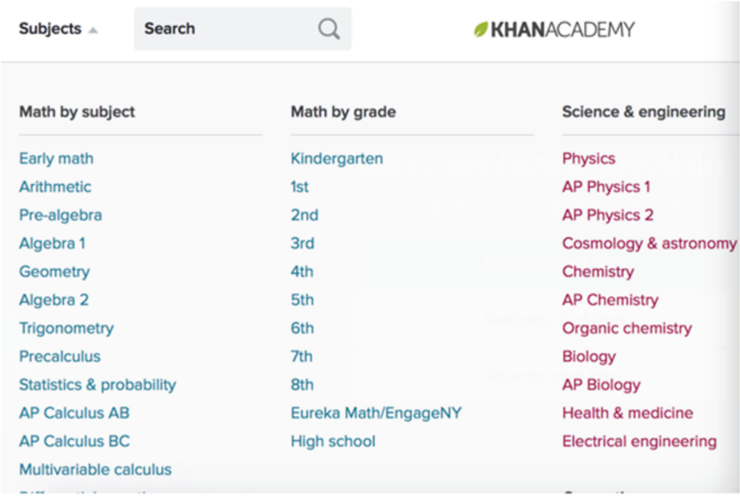

Fig. 1. Knowledge prerequisite of math subject on Khan Academy

This paper proposes a solution for MOOC recommendation on prerequisite and subsequent learning objects. Recommendations on Forgetting-punished MOOC Recommendation (FMR) recommends according to learners' learning situation. It diagnoses both qualified and unqualified location points on learning series (learning behaviors on the time series of the learner). Forgetting effect is combined for both correlation coefficient and recommendation feature measurement. For the unqualified learning objects, FMR recommends prerequisite according to learning series of learners who are qualified with the object. For qualified learning object, FMR recommends objects that take the qualified learning objects as prerequisite.

The main contributions are as follows:

– Learning location helps for adaptive recommendation according to learners' performance. The recommendation aims at the located qualified and unqualified learning objects.
– The forgetting effect is considered as punishment of learning score. It modifies the learning score with consideration on time decay for forgetting. Learning scores are adopted as features for recommendation. It is used to measure the prerequisite correlation. It reflects the effect of forgetting with time on.
– Experiments on realworld data show the improvement of FMR in accuracy. Especially the precision is improved obviously.

Section of Related Work is about research work of prerequisite and recommendation on learning series. The following section introduces FMR according to the work flow of recommendation. It includes prerequisite coefficient calculation and recommendation for prerequisite and subsequent learning. Experiments list the dataset and result of comparison with different recommendation methods. The last section is the summary.

2 Related Work

Prerequisite plays an important role in MOOC learning. In application, prerequisite is usually defined by expert labeling. Polyzou predicts academic performance based on the prerequisite relationship between courses which is achieved by expert annotation [4]. But manual labeling depends much on the experts. It cannot support massive recommendation.

In most research, prerequisite correlation is mainly calculated through knowledge based concepts analysis. Yang builds the concept map through the prerequisite relationship of the existing curriculum, which is used to predict the prerequisite [5]. Liu studies the learning dependence between knowledge points through text analysis [6]. Some research is based on the analysis of the concept map to establish the prerequisite relationship between the knowledge [7–9]. Wikipedia's content is mostly used for prerequisite training. Liang defines the prerequisite relationship between knowledge on links between pages of Wikipedia [10,11]. Wang adopts Wikipedia's links between knowledge concepts and establishes a concept map for teaching materials [12]. Agrawal extracts key concepts in the textbook and calculates prerequisite values between two concepts through the frequency and sequence of them [13]. These methods are all based on content. They are not personalized.

Sequential learning data is used for recommendation. Lu uses the association rule mining method to recommend courses and trains on other learners' learning paths [14]. Sun analyses learning path through metapath method, enriching the learner's portfolio [15]. Chen compares the homogeneity between the user and the item's image by path similarity [16]. Yu learned a similar user's behavioral sequence through collaborative filtering to make sequential recommendations [17]. These methods focus on prerequisite of knowledge, and do not recommend according to the situation of learners.

Yueh proposes a Markov-based recommendation on learning sequences and analyses the learning path from learner history [18]. Mi makes recommendations based on the context tree, focusing more on solution design than implementation [19]. Yu uses collaborative filtering to recommend in a game with storyline through other users' sequential actions [17]. Lee learns the sequence of behavioral learning courses through learners [20]. The recommendation considers only on subsequent recommendations and does not consider on relationship of prerequisite. And the feature is mainly on preference without consideration on learning performance.

We propose the solution FMR to recommend for prerequisite and subsequent learning objects with forgetting punishment.

3 Forgetting-Punished Recommendation on Prerequisite and Subsequent Learning Objects

Adaptive learning responds to learners according to their learning situation. It helps for less frustration and less drop-out [26]. To support adaptive learning better, RFP recommends according to situation of the learner. The situation is measured with learning location. Based on the location, the qualified and unqualified learning objects are detected. RFP recommends prerequisite learning objects for the unqualified learning objects and subsequent learning objects. Correlation of prerequisite on learning scores are adopted as features for recommendation. The effect of forgetting is combined to model the real learning better. According to the working flow of RFP, forgetting effect, prerequisite correlation and recommendation on prerequisite and subsequent learning are introduced in sequence.

3.1 Symbols

Before further discussion, some related symbols are listed with description in Table 1.

Table 1. Symbols

Symbol	Description
se_{si}	Score of learner s on learning object i
dis	Distance on time
$q(i, d_1)$	Prerequisite correlation coefficient between learning objects i and d_1
d_{si,sd_1}	Time distance between 2 learning behaviors of learner s on learning object i and d_1
I_{pr}	Learning object set for prerequisite recommendation
d_1	First unqualified learning object
d_2	Last qualified learning object
d_1	Any qualified learning object
sim_{sr}	Similarity between learner s and r
p_{ri}	Recommendation value of learning object i for learner r

3.2 Punishment of Forgetting Effect on Learning Score

German psychologist H. Ebbinghaus found that forgetting begins immediately after the learning behavior. The knowledge maintenance goes down with time on. The process of forgetting is not uniform. Ebbinghaus believes that the maintenance of mastered learning content is a function of time [22]. Table 2 lists Ebbinghaus's experimental results:

With the data of Table 2, Ebbinghaus proposes the forgetting curve as Fig. 2. It can be found that the memory is divided into short-term memory and long-term memory. The first memory zone is 5 min, the second memory zone is 30 min, and the third memory zone is 12 h. The first 3 memory zones belong to the category of short-term memory [23]. The fourth memory zone is 1 day, the fifth memory zone is 2 days, the sixth memory zone is 4 days, the 7th memory zone is 7 days, the 8th memory zone is 15 days, the last 5 memory zones are long-term memory [24].

Even for knowledge of science or engineering, although you will not forget so fast, the proficiency will low down like forgetting. Considering on the necessary of review, the learner still need practice repeatedly to strengthen the skills.

The learning score indicates knowledge maintenance. It is punished with time for forgetting. Even for the learning objects of science, the proficiency needs review with time on.

Table 2. Time points of Ebbinghaus's experimental results.

Days	Knowledge maintenance
0	0.33
0.33	0.582
1	0.442
8	0.358
24	0.337
48	0.278
144	0.254
720	0.211

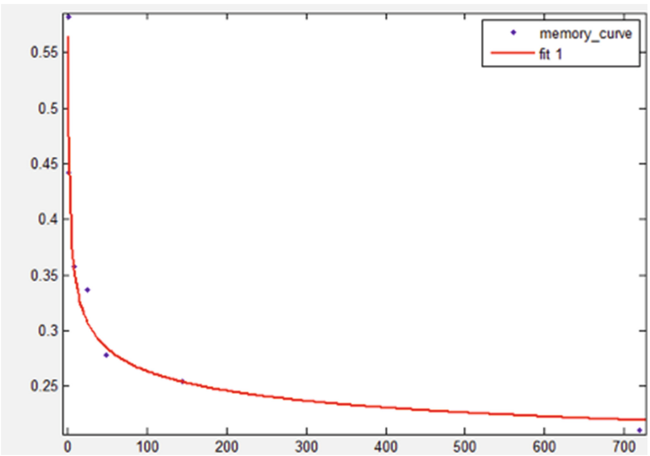

Fig. 2. Ebbinghaus forgetting curve

Matlab tools are used to model the forgetting function. By fitting the data of Table 2, the corresponding mathematical equation is as (2). It represents the score $se(si)$ decay with forgetting on time distance dis.

$$f(se_{si}, dis) = se_{si} * (0.34 * dis^{-0.2} + 0.13) \tag{1}$$

3.3 Prerequisite Correlation Coefficient Measuring

Breese made collaborative filtering recommendations based on correlation coefficients, vector comparisons, and Bayesian statistics. Correlation coefficient was found more accurate [25]. RFP defines the prerequisite correlation coefficient measurement with learning scores.

For two learning objects i and d_1, if scores of i and d_1 is positive correlated, the coefficient should be positive too. We calculate the correlation coefficient $q(i, d_1)$ by the Pearson correlation coefficient by learners' scores on the 2 learning objects.

The correlation between two learning behaviors is also affected by the time distance. If the time distance is long, the knowledge maintenance will decrease for forgetting. We suppose the behavior of learning object i takes place first. The maintenance of learning score on i after forgetting is punished at the time of learning behavior about d_1. $d_{(si,sd_1)}$ is the time distance between the two learning behaviors of learner s on learning objects i and d_1.

In order to keep the correlation values between 0 and 1, logic regression is adopted. The prerequisite correlation coefficient is calculated as (2) shows.

$$q(i, d_1) = \frac{1}{1 + e^{(\frac{\sum_{s=0}^{n_s} (f(se_{si}, d_{si,sd_1}) - \overline{se_l}) * (\overline{se_{d_1}} - se_{d_1})}{\sqrt{\sum_{s=0}^{n_s} (f(se_{si}, d_{si,sd_1}) - \overline{se_l})^2 \sum_{s=0}^{n_s} (se_{d_1} - \overline{se_{d_1}})^2}})}} \tag{2}$$

3.4 Prerequisite Recommendation (PR) for the Unqualified Learning Object

According to the learning location, the first unqualified learning object d_1 needs prerequisite recommendation to the learner. RFP recommends through learning path of learner neighbors that is qualified in d_1. Their learning objects before d_1 become prerequisite candidates to be recommended as Fig. 3 shows.

The algorithm is shown in Algorithm 1. The algorithm recommends prerequisite learning object set I_{pr} for the target learner r. The first layer of the cycle goes through each similar learner s among qualified learner neighbor set S who are qualified in d_1. Learning objects of them before d_1 are adopted as recommendation candidates i. The second layer of the cycle is for each candidate I of the learner neighbor s. se_{ri} is the learning score of the qualified learner neighbor s on learning object i. sesi contribute to the recommendation value. The prerequisite correlation $q(i, d_1)$ is multiplied as weight of the learning score feature. After the cycle, the recommendation value is normalized.

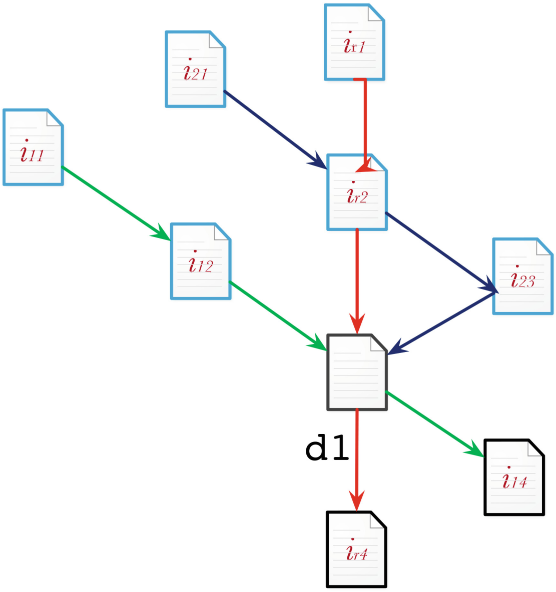

Fig. 3. Prerequisite recommendation candidates for d_1

seri is the learning score of the target learner r on learning object i. Considering on the necessity of review, it is in negative correlation with the recommendation value. For the forgetting after learning, it is punished with forgetting function $f(se_{ri}, dis(time(r, d_1), t_n))$. t_n is the system time of recommendation. $se_{r,d1}$ is the score of the target learner r on unqualified location d_1. $f(se_{ri}, dis(time(r, d_1), t_n))$ shows the inverse correlation of the score. It is in negative correlation with recommendation value for the review necessity. seri and ser, d1 are both considered for recommendation.

3.5 Subsequent Recommendation (SR) for the Qualified Learning Objects

The latest qualified learning object of the learner is defined as d_2. It means learning objects of the target learner before d_2 are all qualified. They are all indicated by symbol b. The subsequent learning objects with b as prerequisite should be recommended. The first learning object learned by qualified learner neighbors after d_2 is adopted as recommendation candidate as Fig. 4 shows. So are the qualified learning objects b.

The recommendation value is calculated according to the learning series of learner neighbors who are qualified in both d and following learning objects.

Qualified learner neighbors' learning scores on learning objects following b is adopted as one of the features for recommendation. The learning scores have similarity and prerequisite correlation coefficient as weights. The qualified learner

Algorithm 1. Prerequisite recommendation

Require: learner vectors $L\{l_1, l_2, \ldots, l_m\}$, the target learner r, unqualified location d_1;

Ensure: prerequisite recommendation result I_{pr};

get top similar d_1 qualified learner set $Ss_1, s_2, ..., s_{k1}$;

for EACH $s \in S$ **do**

  for EACH $i \in s\{i_1, i_2, \ldots, i_{index(d_1)}\}$ **do**

    $p_{ri} += sim_{sr} \times se_{si} \times q(i, d_1)$;

    $dec_{ri} += sim_{sr} \times q(i, d_1)$;

$p_{ri} /= dec_{ri}$;

$p_{ri} += w_1 \times (100 - f(se_{ri}, dis(time(r, d_1), time(r, i)))) + w_2 \times (100 - se_{r,d_1}) + w_3 \times p_{ri}$;

select top k_2 p_{ri} for learner r, add i to I_{pr}

return I_{pr};

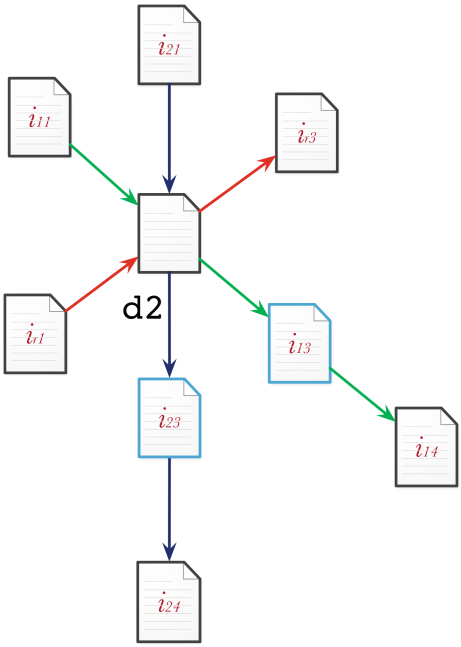

Fig. 4. Subsequent recommendation candidates for d_2

neighbor's learning score on prerequisite objects b and the learning score of the target learner on b are both combined as features for better performance of recommendation. The learning score of the target learner on b is punished by the forgetting function for better modeling of reality.

4 Experiment

The experiment is conducted on data recorded by the mic-video platform of ECNU[1]. It includes 686 learners, 136 mic-videos as learning objects and 7,163 related learning records.

The accuracy of recommendation is compared by precision, recall and f1-score. Experiments verify the improvement of accuracy on PR and SR.

k_1 is the number of selected top similar qualified learner neighbors for recommendation. k_2 is the top recommended items. The parameters of k_1 and k_2 were separately adjusted to test the performance. Weight parameters are assigned as 1 without loss of generality.

Different recommendation methods are compared under various k_1 and k2 combinations. One is collaborative filtering recommendation on interest CFPreference, and the other is a collaborative filtering recommendation on learning scores CFscore.

Figure 5 compares the precision between different k1 and k2 combinations. SR has the best performance in precision. It decreases the range of candidate learning objects. The learner neighbor's learning series is used for candidate selection. Only the first learning object after the qualified location is selected as a candidate for recommendation. The recommended results are more accurate. The precision of PR is better, The recommendation on prerequisite correlation has better performance.

Figure 6 is a comparison of recall under different combinations of k1 and k2. The performance of PR is relatively better. Its recommendation candidates cover all possible prerequisite learning objects of learner neighbors. The candidate of

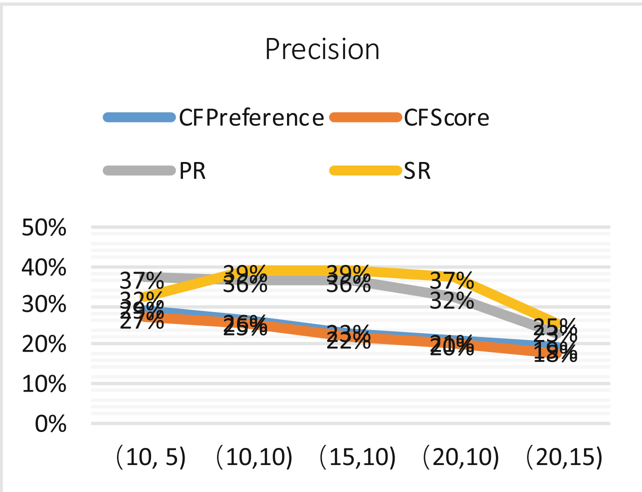

Fig. 5. Precision comparison between different recommendations

1 http://jclass.pte.sh.cn.

SR cover only one learning object immediately following d. CF methods consider all learning objects of learner neighbors as candidates for recommendation. Compared with CF methods, the candidates of PR and SR is decreased. But the recall is not decreased. It shows the accuracy of PR and SR.

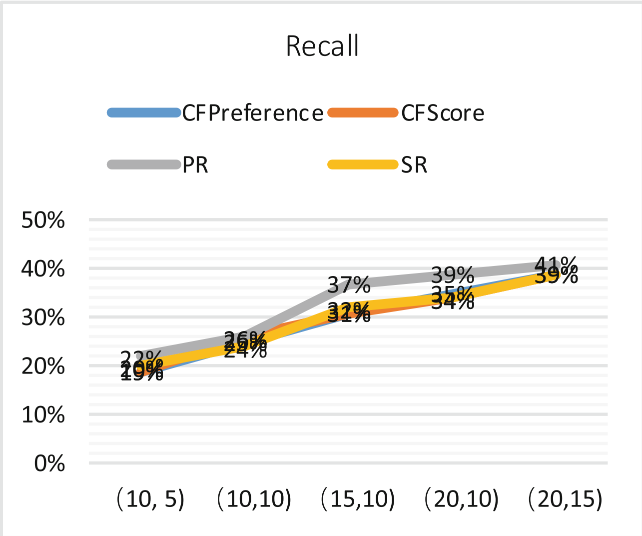

Fig. 6. Recall comparison of different recommendation

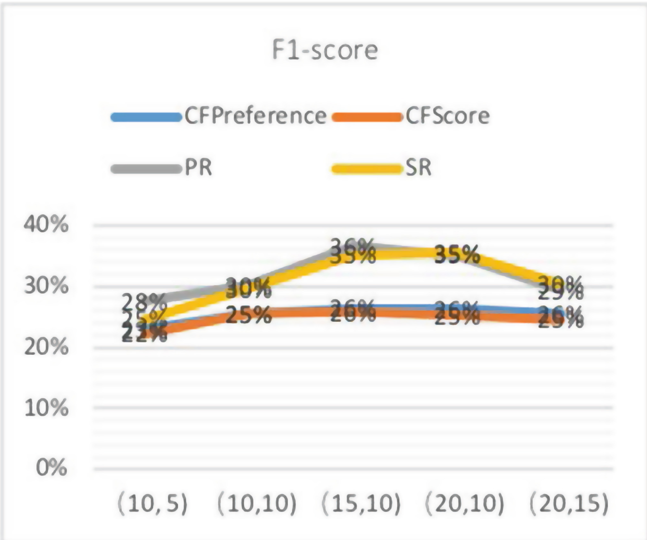

Fig. 7. f1-score comparison between different recommendations

Figure 7 compares f1-scores between different combinations of k1 and k2. Because the different performance on precision and similar performance on recall, and the comparison result on f1-score is similar to that on precision. The results of CFPreference and CFScore are similar, but not as good as PR and SR.

5 Summary

This paper proposes a Forgetting-punished MOOC Recommendation (FMR) on prerequisite. FMR recommends for qualified and unqualified learning objects that are diagnosed by location. Prerequisite learning objects are recommended for the unqualified locations, and subsequent learning objects are recommended for the qualified locations. The feature of learning score is punished for forgetting to model the reality better. It is different from normal MOOC recommendation on learning objects that are not learned before. Experiment verifies the improvement on accuracy by prerequisite recommendation (PR) and subsequent recommendation (SR). The prerequisite may be more than one learning object. The "and" relation between prerequisite learning objects deserves further research.

Acknowledgment. The work is funded by computer science and technology subject of Shanghai Polytechnic University with No. xxkzd1604.

References

1. Breslow, L., Pritchard, D.E., De Boer, J., Stump, G.S., Ho, A.D., Seaton, D.T.: Studying learning in the worldwide classroom: research into edX's first MOOC. Res. Pract. Assess. **8**, 13–25 (2013)
2. Kizilcec, R.F., Piech, C., Schneider, E.: Deconstructing disengagement: analyzing learner subpopulations in massive open online courses. In: Proceedings of the Third International Conference on Learning Analytics and Knowledge, pp. 170–179. ACM (2013)
3. Pappano, L.: The year of the MOOC. New York Times, New York (2012)
4. Polyzou, A., Karypis, G.: Grade prediction with course and student specific models. In: Bailey, J., Khan, L., Washio, T., Dobbie, G., Huang, J.Z., Wang, R. (eds.) PAKDD 2016. LNCS (LNAI), vol. 9651, pp. 89–101. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-31753-3_8
5. Yang, Y., Liu, H., Carbonell, J., Ma, W.: Concept graph learning from educational data. In: Proceedings of the Eighth ACM International Conference on Web Search and Data Mining, pp. 159–168. ACM (2015)
6. Liu, J., Jiang, L., Wu, Z., Zheng, Q., Qian, Y.: Mining learning dependency between knowledge units from text. VLDB J. Int. J. Very Large Data Bases **20**(3), 335–345 (2011)
7. Huang, X., Yang, K., Lawrence, V.B.: An efficient data mining approach to concept map generation for adaptive learning. In: Perner, P. (ed.) ICDM 2015. LNCS (LNAI), vol. 9165, pp. 247–260. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-20910-4_18
8. Scheines, R., Silver, E., Goldin, I.M.: Discovering prerequisite relationships among knowledge components. In: EDM, pp. 355–356 (2014)

9. Vuong, A., Nixon, T., Towle, B.: A method for finding prerequisites within a curriculum. In: EDM, pp. 211–216 (2011)
10. Liang, C., Wu, Z., Huang, W., Giles, C.L.: Measuring prerequisite relations among concepts. In: Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, pp. 1668–1674 (2015)
11. Talukdar, P.P., Cohen, W.W.: Crowd sourced comprehension: predicting prerequisite structure in Wikipedia. In: Proceedings of the Seventh Workshop on Building Educational Applications Using NLP, pp. 307–315. Association for Computational Linguistics (2012)
12. Wang, S., et al.: Using prerequisites to extract concept maps from textbooks. In: Proceedings of the 25th ACM International on Conference on Information and Knowledge Management, pp. 317–326. ACM (2016)
13. Agrawal, R., Golshan, B., Terzi, E.: Grouping students in educational settings. In: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1017–1026. ACM (2014)
14. Lu, Z., Pan, S.J., Li, Y., Jiang, J., Yang, Q.: Collaborative evolution for user profiling in recommender systems. In: IJCAI, pp. 3804–3810 (2016)
15. Sun, Y., Han, J., Yan, X., Yu, P.S., Wu, T.: Pathsim: meta path-based top-k similarity search in heterogeneous information networks. Proc. VLDB Endow. **4**(11), 992–1003 (2011)
16. Chen, Y., Zhao, X., Gan, J., Ren, J., Hu, Y.: Content-based top-N recommendation using heterogeneous relations. In: Cheema, M.A., Zhang, W., Chang, L. (eds.) ADC 2016. LNCS, vol. 9877, pp. 308–320. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46922-5_24
17. Yu, H., O'Riedl, M.: A sequential recommendation approach for interactive personalized story generation. In: Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems, vol. 1, pp. 71–78. International Foundation for Autonomous Agents and Multiagent Systems (2012)
18. Huang, Y.-M., Huang, T.-C., Wang, K.-T., Hwang, W.-Y.: A Markov-based recommendation model for exploring the transfer of learning on the web. J. Educ. Technol. Soc. **12**(2), 144 (2009)
19. Mi, F., Faltings, B.: Adaptive sequential recommendation using context trees. In: IJCAI, pp. 4018–4019 (2016)
20. Lee, Y., Cho, J.: An intelligent course recommendation system. SmartCR **1**(1), 69–84 (2011)
21. Education Growth Advisors: Learning to adapt: understanding the adaptive learning supplier landscape. PLN/Bill and Melinda Gates Foundation (2013)
22. Ebbinghaus, H.: Memory: a contribution to experimental psychology. Ann. Neurosci. **20**(4), 155 (2013)
23. Schacter, D.L.: The seven sins of memory: insights from psychology and cognitive neuroscience. Am. Psychol. **54**(3), 182 (1999)
24. Averell, L., Heathcote, A.: The form of the forgetting curve and the fate of memories. J. Math. Psychol. **55**(1), 25–35 (2011)
25. Breese, J.S., Heckerman, D., Kadie, C.: Empirical analysis of predictive algorithms for collaborative filtering. In: Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence, pp. 43–52. Morgan Kaufmann Publishers Inc. (1998)
26. Clark, D.: Adaptive MOOCs. CogBooks adaptive learning. Copyright CogBooks (2013)