Field-aware Knowledge Tracing Machine by Modelling Students' Dynamic Learning Procedure and Item Difficulty

Wenbin Gan*, Yuan Sun*, Shiwei Ye†, Ye Fan†, and Yi Sun†
*National Institute of Informatics,SOKENDAI,Tokyo, Japan,
†School of Computer and Control Engineering,University of Chinese Academy of Sciences,Beijing, China

***Abstract*—Knowledge tracing is essential for adaptive learning to obtain students' current knowledge states to provide adaptive service. However, the learning process is evolving constantly since students learn and forget over time. Moreover, the difficulty of learning materials can also have huge influence on their performances. This work shows an ongoing project on knowledge tracing by proposing a Field-Aware Knowledge Tracing Machine (FA-KTM) to integrate students' dynamic learning procedure (learning and forgetting) and adaptive item difficulty. We propose methods to model item difficulty and the dynamic learning procedure and present the framework to integrate them together. Preliminary analyses on a dataset show the promising results.**

***Index Terms*—Knowledge Tracing, Field-aware Factorization Machines, Item Difficulty, Learning, Forgetting**

I. Introduction

Knowledge tracing (KT), aiming at tracing students' knowledge states and skill acquisition levels over time, is essential for personalized learning and is also a fundamental part in intelligent tutoring systems [2]–[4], as it can predict how students will perform on future interactions based on previously observed performances on problems, thus providing adaptive remedial learning materials based on their individual needs.

Cognitive diagnostic models and data mining techniques have been widely used in characterizing students' implicit knowledge proficiency from a static perspective. However, the learning process of students is not static but evolves over time [6] since students learn and forget over time, which makes tracing students' knowledge inherently difficult. Educational psychologists have long converged that student's dynamic learning procedure has huge influence on their performances [2], [6], as well as in the prediction accuracy of knowledge tracing. Two classical educational theories (the Learning curve theory and the Ebbinghaus forgetting curve theory) can well explain this dynamic changes [6]. Students' knowledge proficiency can be enhanced with trials and can decline over time due to forgetting. Moreover, Minn [2] has proved that problem difficult can also have huge influence on students' learning outcomes. Various methods have been proposed to trace students' knowledge dynamically, such as Item Response Theory (IRT), deep knowledge tracing (DKT) [4], [7], Knowledge Tracing Machine (KTM) [3] and DAS3H model [5].

This work is supported by JSPS KAKENHI Grant Number 18K11597.

As initial discovery, we found that existing work on knowledge tracing do not fully consider the learning and forgetting in the dynamic learning procedure, or just consider part of them. Moreover, most of the studies assume that the difficulty of a specific item is constant, which is impractical in the real learning process, as students may attempt the same problems for many times, and they may gradually master the knowledge contained in the problem, hence the difficulty of the item should be adjusted based on the student's current knowledge level.

II. Field-aware Knowledge Tracing Machine

This paper proposed a Field-Aware Knowledge Tracing Machine (shorted as FA-KTM) by integrating students' dynamic learning procedure (learning and forgetting) and real-time item difficulty to perform KT.

A. Item Difficulty

Based on the above analysis, item difficulty should be determined based on students' current learning process instead of being constant. Here we define it as following:

$$d(i,j,t) = \delta_j + \sum_{k \in KC(j)} \beta_k + \theta_m \left[\frac{\sum_k \Psi_{i,k,t}}{|KC(j)|} \right] + \theta_n \Psi_{i,j,t} \quad (1)$$

where $d(i,j,t)$ is the difficulty level of item j to student i at time t, δ_j and β_k are the inherent difficulty level of item j and knowledge component (KC) k, respectively. The last two terms of Eq. (1) represent the adjustive terms to accommodate the difficulty level of the item to a student's current knowledge level. Inspired by [2], they are defined in Eq. (2).

$$\Psi_{i,v,t|v=\{j,k\}} = \begin{cases} \left[\frac{|\{x_{i,v}==0\}|_{0:t}}{|N_{i,v}|_{0:t}} * (c-1) \right], & if\, |N_{i,v}|_{0:t} \geq 5 \\ c, & else \end{cases} \quad (2)$$

where $\Psi_{i,j,t}$ and $\Psi_{i,k,t}$ are quantified into $c+1$ levels of difficulty (ranging from 0 to c). $N_{i,v}$ is the set of problems or skills the student i has attempted before time t, and $x_{i,v}$ is the outcome of the attempt from student i to problem j or skill k. An outcome of 0 is a failure. If a student attempts a problem or skill less than 5 times, the difficulty level will be set as c meaning the highest level of difficulty.

B. Modelling Student Learning and Forgetting

Learning and forgetting are two widely accepted procedures in educational psychology that could influence the learning outcomes. The more exercises a student does, the bigger gain

2375-9259/19/$31.00 ©2019 IEEE
DOI 10.1109/ICDMW.2019.00151

of knowledge proficiency he will get. Moreover, the longer the lag time from the previous interaction, the more probability he will forget some knowledge. Based on these two assumptions, we define the learning as:

$$l(i,j,t) = \Phi_{i,j,t} + \sum_{k \in KC(j))} \Phi_{i,k,t} \tag{3}$$

$$\Phi_{i,v,t|v=\{j,k\}} = \theta_{v,3w+1} log(1+W_{i,v,0:t}) + \theta_{v,3w+2} log(1+ F_{i,v,0:t}) - \theta_{v,3w+3} log(1+A_{i,v,0:t}) \tag{4}$$

where learning $l(i,j,t)$ is composed of the acquisition from attempting both the same items and also different items containing the same set of skills. $W_{i,v,0:t}$ and $F_{i,v,0:t}$ denote the amount of times that skill or item v has been correctly and incorrectly recalled among $A_{i,v,0:t}$ times of attempts in time t by student i. Being correct or incorrect in some items both contribute to the acquisition of knowledge.

Early studies about forgetting revealed that the retention rate decreases exponentially as time passes by [6], hence we formulate the forgetting behaviour as:

$$f(i,j,t) = \theta_{j,j} e^{-\Delta_{j,j}} + \theta_{k,k} \sum_{k \in KC(j))} e^{-\Delta_{k,k}} + \theta_{j,j-1} e^{-\Delta_{j,j-1}} \tag{5}$$

where $f(i,j,t)$ can be represented as the memory strength, which is composed of the lag time between an interaction and the previous interaction with the same item $\Delta_{j,j}$, with the same skill $\Delta_{k,k}$, and between adjacent interaction in the sequence $\Delta_{j,j-1}$. For some problems, they are related or similar with each other, or the skills contained are related with each other, the lag time between these interactions in the sequence may affect the performance of the question. Incorporating the sequence time gap into the model may capture this effect [4].

C. Proposed FA-KTM Model

Based on the item difficulty and the modelling of learning and forgetting, this paper proposes a Field-aware Knowledge Tracing Machine leveraging the recent field-aware factorization machine (FFM) framework [1] to enrich the proposed model by embedding the features in d dimensions and modelling pairwise interactions between those features. The FFM framework learns a parameter per field of features in order to draw different importance on different field of features.

For an embedding dimension of $d = 0$, our model reads:

$$P(Y_{i,j,t} = 1) = \sigma(\alpha_i - d(i,j,t) + l(i,j,t) + f(i,j,t)) \tag{6}$$

by incorporating Eq. (1), (3) and (5). Thus, the probability of student i correctly attempting item j at time t depends on his ability α_i, the real-time difficulty level of the item j and his learning and forgetting during this period of time.

For higher embedding dimensions $d > 0$, all features are embedded in d dimensions and their interactions are modelled in a pairwise manner among different fields. The fields in our model are the different factors considered. The quadratic term of our model is

$$\phi_{FA-KTM} = \sum_{i=1}^{n} \sum_{j=i+1}^{n} \langle w_{i,f_j}, w_{j,f_i} \rangle x_i x_j \tag{7}$$

where x_i and x_j are the i^{th} and j^{th} features of the input vector, and f_i and f_j are the fields of i^{th} and j^{th} features, respectively.

III. WORK IN PROGRESS

We have conducted some preliminary analyses on a public dataset named KDD Algebra I 2005-2006. [1] As shown in Figure 1, we visualize the difficulty level of the first 40 problems attempted by the first 30 students in skill level using Eq.(2), and also the corresponding response outcome. We can see that students answering difficult problems are more likely to get the wrong answer for their attempts than the easy ones, which shows the effective of the proposed adaptive item difficulty. Moreover, we also plot the lag time of 200 problems attempted by a student in Figure 2, we can see that long lag time interval from the same items may lead to the failure for the same practices in the later attempts.

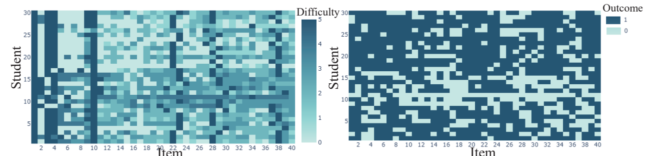

Fig. 1. Problem difficulty (left) and response outcome (right) over first 40 problems attempted by 30 students.

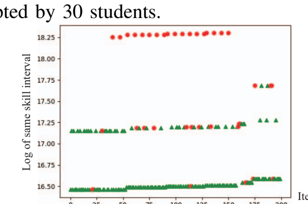

Fig. 2. The effect of lag time interval on the responses of a student over 200 attempted problems.

Our ongoing work consists of the following potential directions: Firstly, assessing students' knowledge levels and conducting the knowledge tracing over time: we will integrate more side information into our model and leverage other advanced framework to enrich our model to improve its performance and further test it on several public datasets, and compare with other models. Secondly, conduct educational application: we plan to integrate the KT function into some real tutoring systems and provide students with proper adaptive remedial materials based on their individual needs, and conduct case study to prove its effectiveness on students' learning.

REFERENCES

[1] Juan, Y., et al, "Field-aware factorization machines for CTR prediction," the 10th ACM Conf. on Recommender Systems, pp. 43-50, 2016.
[2] Minn, S., et al., "Improving knowledge tracing model by integrating problem difficulty," IEEE ICDM Workshops, pp. 1505-1506, 2018.
[3] Vie, J. J., Kashima, H., "Knowledge Tracing Machines: Factorization Machines for Knowledge Tracing," AAAI, Vol. 33, pp. 750-757, 2019.
[4] Nagatani, K., et al., Ohkuma, T., "Augmenting Knowledge Tracing by Considering Forgetting Behavior," WWW, pp. 3101-3107, 2019.
[5] Choffin, B., et al., "DAS3H: Modeling Student Learning and Forgetting for Optimally Scheduling Distributed Practice of Skills," EDM, 2019.
[6] Chen, Y., et al., "Tracking knowledge proficiency of students with educational priors," the 2017 ACM Conf. on Information and Knowledge Management, pp. 989-998, 2017.
[7] Minn, Sein, et al. "Deep knowledge tracing and dynamic student classification for knowledge tracing," 2018 IEEE ICDM, pp.1182-1187.

[1] http://pslcdatashop.web.cmu.edu/KDDCup/downloads.jsp