Mining Users' Behaviors in Intelligent Educational Games Prime Climb a Case Study

Alireza Davoodi
Department of Computer Science
The University of British Columbia
2366 Main Mall, Vancouver, BC, V6T1Z4, Canada
+1 (604) 8225108
davoodi@cs.ubc.ca

Samad Kardan
Department of Computer Science
The University of British Columbia
2366 Main Mall, Vancouver, BC, V6T1Z4, Canada
+1 (604) 8225108
skardan@cs.ubc.ca

Cristina Conati
Department of Computer Science
The University of British Columbia
2366 Main Mall, Vancouver, BC, V6T1Z4, Canada
+1 (604) 8224632
conati@cs.ubc.ca

ABSTRACT
This paper presents work on applying clustering and association rule mining techniques to mine users' behavior in interacting with an intelligent educational game, Prime Climb. Through such behavior discovery, frequent patterns of interactions which characterize different groups of students with similar interaction styles are identified. The relation between the extracted patterns and the average domain knowledge of students in each group is investigated. The results show that the students with significantly higher prior knowledge about the domain behave differently from those with lower prior knowledge as they play the game and that pattern could be identified early during the interactions.

Keywords
Intelligent Educational Games, Behavior Discovery, Association Rule Mining

1. INTRODUCTION
Many Adaptive educational systems apply data mining techniques to answer the need for understanding and supporting varying learning styles, capabilities and preferences in students[1, 2, 3]. Along this line of research, we concentrate on understanding how students interact with Prime Climb (PC) as an adaptive educational game and whether there is a connection between behavioral patterns and attributes (for instance higher knowledgeable vs. lower knowledgeable students) in the students. Developing an interactive environment in which more number of students can learn the desired skills requires a pedagogical agent which maintains more accurate understanding of individual differences between users and provides more tailored interventions. For instance, if a pedagogical agent is capable of identifying whether a group of students have higher domain knowledge than the other group, it can be possible to leverage such information to construct a more accurate user model and intervention mechanism.

Behavioral discovery has been vastly used in educational systems but there is limited application in educational games like PC in which educational concepts are embedded in the game with minimum technical notation to maximize game aspects (i.e. engagement) of the system. In PC, students follow an exploratory mechanism to explore and understand the methods and practice them. This paper describes the first step toward leveraging students' behavioral patterns into building more effective adaptive edu-game. The ultimate goal is devising mechanisms for making abstract high level meaning from raw interaction data and leveraging such understanding for real-time identification of characterizing interaction styles to enhance user modeling and intervention mechanism in an edu-game like Prime Climb.

2. Prime Climb Intelligent Edu-game
Prime Climb (PC) is an intelligent educational game for students in grades 5 and 6 to practice number factorization skills. In PC, the player and his/her partner climb a series of 11 mountains of numbers by pairing up the numbers which do not share a common factor [4]. There are two main interactions of a player with PC:
***Making Movements*:** A player makes one or more movements at each time, by clicking on numbered hexagons on the mountains.
***Using Magnifying Glass Tool*:** The magnifying glass (MG) tool is always available for the user to benefit from. The MG is used to show the factor tree of a number on the mountains; it is located in the top right corner of the game.

3. Data Collection/User Representation
For behavior discovery, we used the student's interaction data with the first 9 mountains of 43 students who completed at least 9 levels (mountains) of Prime Climb. Each user is represented by a vector of features. Some of the features are shown in Table 1. Each feature is a measure computed based on user's interactions with one or more mountains. In this paper we provide the results for two feature sets.

Mountains-Generic-Movement(1-9) features: Contains features calculated based on the users' movements behavior on mountains 1 to 9. A "*mountain-generic*" feature is a feature which is calculated across all mountains not individual mountains.
Mountains-Generic+Specific-MG+Movement(1-4) features: Contains features calculated base on user's movement and MG-usage behaviors on mountains 1 to 4. A "*mountains-specific*" feature is measured based on data from an individual mountain.

Table 1: Some Features used for behavior discovery

Movement Features
Sum/Mean/STD number of correct/wrong moves across mountains
Sum/Mean/STD of time on [correct/wrong] moves across mountains
Mean/STD length of sequence of correct/wrong moves
Mean/STD time spend per sequence of correct/wrong moves
Magnifying Glass (MG) Features
Sum/Mean/STD of MG Usage
Mean/STD number of [correct/wrong] movements per each MG usage

4. Clustering and Association Rule Mining
Prior to performing clustering, feature selection mechanism is applied to filter out irrelevant features [5]. Then, the optimal number of clusters is determined as the lowest number suggested by C-index, Calinski and Harabasz[4] and Silhouette [6] measures of clustering validity. Next, the GA *K*-means (*K*-means for short) clustering algorithm [1], which is a modified version of GA *K*-means [7], is applied to cluster the users into an optimal number of clusters.

Table 2: Extracted Rules for Mountains-Generic-Movement(1-9)

Rules for Cluster 1[HPK]: (Size: 10/43 = 23.26%)
Mean-Time-on-Movements(1-9) = Higher, [6/6=100%]
Mean-Time-Spent-On-Correct-Movements-On-Mountains(1-9) = Higher, ([5/5=100%])
Rules for Cluster 2[LPK]: (Size: 33/43 = 76.74%)
Mean-Time-On-Movements(1-9) = Lower, [33/37=89.19%]
○ STD-Time-On-Wrong-Correct-Moves(1-9) = Lower, [33/35=94.29%]
Mean-Time-On-Consecutive-Wrong-Movements(1-9) = Lower, [31/35=88.57%]
○ STD-Time-On-Movements(1-9) = Lower, [31/33=93.94%]
○ STD-Time-On-Correct-Movements(1-9) = Lower, [31/33=93.94]

The Hotspot algorithm is used to extract the rules for each discovered cluster. The clusters are then compared for statistical difference on a measure called *cluster's prior knowledge*:

$$\boldsymbol{Cluster's\ prior\ knowledge} = \frac{\sum_{student \in cluster} \boldsymbol{pre\_test(student)}}{\boldsymbol{Cluster's\ size}}$$

$pre\_test(student)$ is the student's score on a pre-test taken before playing PC. The max, average and standard deviation of the scores across the students are 15, 11.7 and 3.29 respectively.

Behavior Discovery on Mountains-Generic-Movement(1-9) Set: Feature selection mechanism selected 18 features out of original 30 features. The optimal number of clusters was found to be 2. The result of a t-test showed that there is a statistically significant difference between the prior knowledge of cluster 1 of students (*higher prior knowledge (HPK) group*) (*M*=13.0 , *SD*=2.0) and cluster 2 of students (*lower prior knowledge (LPK) group*) (*M*=11.3 , *SD*=3.45), *p*=.03 and *cohen-d*=.53. Table 2 shows the rules extracted for each cluster using the Hotspot algorithm. Each bulleted item in Tables 2 and 3 shows an extracted rule. "Higher" and "Lower" are the *bin*s. We considered two bins in this study. The bin shows whether the value of the feature is located in the higher or lower portion of the feature values across students. The cut-off point for splitting a range of feature's values to 2 ranges of lower and upper ranges is calculated specifically for the feature in each extracted rule by the Hotspot algorithm. In front of each rule is a fraction whose numerator and denominator respectively shows the number of students in the cluster and total students on which the rule applies.

The extracted rules show that the students belonging to the HPK cluster, spent more time on movements and correct movements across 9 mountains. This could indicate that the HPK students were more involved in the game and spent more time before making a movement. In contrast, the group of LPK students spent lower time on making movements as well as wrong movements. This could be an indication of less involvement in the game by the LPK group. The other patterns show a lower standard deviation on time spent on making movements and correct movements for LPK group. This indicates that this group of students showed a consistent pattern of lack of engagement in the game.

Behavior Discovery on Mountains-Generic+Specific-MG+Movements(1-4) Set: This feature set only employs interaction data from the first 4 mountains. Such feature set is mainly valuable for constructing an online classifier to classify students to different classes based on their interaction with the game during the gameplay. Table 3 shows the discovered clusters and extracted rules. The result of the t-test shows a statistically significant difference between cluster1's *prior knowledge* (*M*=13.28, *SD*=1.58) and cluster2's *prior knowledge* (*M*=11.39 , *SD*=3.4), *p*=.02, *cohen-d*=.60. Also, around 16% of students belong to HPK cluster and 84% belong to the LPK group.

Table 3: Extracted Rules for Mountains-Generic+Specific-MG+Movements(1-4)

Cluster 1[HPK]: (Size: 7/43 = 16.28)
Mean-Time-On-Movements(4) = Higher, (100% [5/5])
Mean-Time-On-Correct-Movements(3) = Higher, (100% [3/3])
Cluster 2[LPK]: (Size: 36/43 = 83.72%)
Mean-Time-On-Correct-Movements(1-4) = Lower, (100% [35/35])
Mean-Time-On-Movements(1-4)=Lower, (100% [34/34])

This result is very similar to the results when data from all 9 mountains is included. Similar patterns can be seen when more interaction data from upper mountains is included in patterns analysis.

5. CONCLUSION/FUTURE WORK

This paper discusses behavior discovery in PC. To this end, different sets of features were defined. The features were extracted from interaction of students with PC in the form of making movements from one numbered hexagon to another numbered hexagon and usages of the MG tool. In order to identify frequent patterns of interaction in groups of students, firstly a feature selection mechanism was applied to select more relevant features from set of all features. Then a K-Means clustering was applied to cluster the students into optimal number of clusters. Once clusters were built, the Hotspot algorithm of Association Rule Mining is applied on the clusters to extract frequent interaction patterns. Finally the clusters were compared to each other on their cluster's *prior knowledge*. When interaction data from all 9 mountains is included in behavior discovery, it was found that the students with higher prior knowledge were more engaged in the game and spent more time on making movements. On the contrary, the students with lower prior knowledge, spent less time on making movements, indicating that they were less involved in the game. Behavior discovery also was conducted on truncated sets of features in which only a fraction of interaction data was included. The results showed that using the interaction data from the first four mountains resulted in groups of students that are statistically different on their prior knowledge.

As for future work, an online classifier will be built which identifies frequent patterns of interaction in the students and classify them into different groups in real time and leverages such information to build a more personalized user model and adaptive intervention mechanism in PC.

6. REFERENCES

[1] Kardan, S., and C. Conati., 2011, A Framework for Capturing Distinguishing User Interaction Behaviours in Novel Interfaces. EDM 2011, Netherlands, 159-168

[2] Judi, M., Baldwin, J., 2012, Identifying Successful Learners from Interaction Behaviour, EDM 2012, Greece, 160-163

[3] Lopez, M. I., Romero, C., Ventura, S., Luna, J.M., 2012, Classification via clustering for predicting final marks starting from the student, participation in Forums, EDM 2012, Greece, 148-152

[4] Manske, M. , Conati, C., 2005, Modelling Learning in an Educational Game. AIED 2005, 411-418

[5] Guyon, I., & Elisseeff, A. 2003. An introduction to variable and feature selection. *J. of Machine Learning Research*, *3*, 1157–1182.

[6] Milligan, G. W., & Cooper, M. C. 1985. An examination of procedures for determining the number of clusters in a data set. *Psychometrika*, *50*(2), 159–179.

[7] Kim, K. Ahn, H. 2008. A recommender system using GA k-means clustering in an online shopping market. Expert Syst. Appl. 34, 2, 1200-1209.