

Mining the Change of Fuzzy Quantitative Association Rules for Summative Assessment

Chih-Hong Huang¹ Tony Cheng-Kui Huang¹ Shih-Sheng Chen^{2,*}

¹Department of Business Administration

National Chung Cheng University

168, University Rd., Min-Hsiung, Chia-Yi, Taiwan.

bmahck@ccu.edu.tw

²Department of Information Management

National Chin-Yi University of Technology

No.57, Sec. 2, Zhongshan Rd.

Taiping Dist., Taichung, 41170, Taiwan.

*Corresponding author: sschen@ncut.edu.tw

Abstract

A learning management system (LMS) prevails and it accumulates an amount of data about the progress of student learning, demographic and students' background over different sessions. Educators are very concerned about the shifts of unknown relationships among a thousand variables about students in a LMS for adjusting their teaching strategies and pedagogies. However, educators are not satisfied with the traditional reports, which are explored with limited relative variables on the issue of summative assessment and learning achievement in static sessions. The information about the changes of unknown relationships among many variables cannot be produced with statistical methods in traditional reports. Our study proposes a mining change of fuzzy quan-

titative association rules model to reveal the information. This model can discover the six types of changes rules from unknown relationships among many variables with nominal or numerical attributes. Experiments are carried out to evaluate the proposed model. We empirical demonstrate how the model helps educators understand the changing characteristics of students and to modify their teaching practices.

Keywords: data mining, change mining, fuzzy association rules

1. Introduction

Change is an important issue in education. To our knowledge, no educational study investigates the changes of relationships among the characteristics of students and

learning achievements over different sessions in learning management system (LMS). No reports are discussed how to capture the changes of relationships of many unknown variables with different attribute types about students over different sessions.

Traditionally, the reports on the characteristics of students and learning achievements are using statistical methods in a static session. Astin (1993) [1] proposed the input-environment-output (I-E-O) model to explore the correlations variables in the input, environment, and output dimensions. Nowadays, e-learning or web-learning programs are incorporated in a LMS and the more data such as students' behavior are recorded with LSM. However, exploring the relationships on the hundred or thousand variables in LMS about characteristics of students and learning achievements, statistical methods such regression analysis technique are inadequate. Regression analysis technique cannot be explored in relations of a thousand hidden variables. Fortunately, association rules mining is suited and it can discover the unknown relationships among various variables in LMS.

Association rules mining can apply to find which items are frequent bought in a transaction for sales strategies [2]. Our study applies the association rule technique to find the relationships of many different variables among characteristics of students and learning achievements. However, there may have two potential problems existing in this technique. First, the attributes of data using in the association rule mining are nominal types. Quantitative data such as scores of students cannot be processed in mining classical association rules. Second, the technique cannot find the changes of associations over different sessions. According to the two arguments, we propose a model to detect the changes of fuzzy association

rules for addressing the two problems. Then, our model can be applied to reveal the changes of relationships of different students' characteristics and learning achievements over different sessions.

2. Related works

For exploring the data in dynamic world, the change mining approach is important. Change Mining is to monitor the process of change, explore how models have changed and predict possible changes in time-associated data [3]. In this study, we focus our attention on how fuzzy quantitative association rule mining is incorporated in change mining and is applied in educational field.

Many researchers concentrate on association rule mining technique in education fields. Romero and Ventura [4] surveyed the application of association rule mining in different type of web-based educational systems. Originally, the attributes of data (*i.e.* items), handled with association rules mining, are nominal. The data with quantitative attributes are not processed with traditional association rules mining. Srikant and Agrawal [5] proposed approach to handle quantitative data and discover the rules (called *quantitative association rule*) from them. Delgado *et al.* [6] proposed a general model for fuzzy association rules and discussed applications.

To the best of our knowledge, there are no researches to address the change mining in fuzzy quantitative association rules. Therefore, our work fills the research gap by proposing a change mining model for this type of knowledge. We redefine our change rules to detect fuzzy quantitative association rules. Our approach can help detect the shifts of variables in I-E-O model affecting the student achievements over sessions.

3. Problem statement and definition

Based on the Chen and Huang's study [7] in mining fuzzy quantitative sequential patterns, we improve the idea and apply it in association rule domain.

We apply FTDA algorithm, proposed by Hong *et al.* [8], to generate our fuzzy quantitative association rules from student data. To evaluate the degree of difference between r_i^t and r_j^{t+1} , we use a similarity computation method to measure the closeness between these rules. We develop a Similarity Computation Index (SCI) formula to calculate the similarity degree between two rules at different time points.

4. Mining the changes of fuzzy quantitative association rules

We present a “FuzzChgMining” model to mine the change of fuzzy quantitative association rules in student behaviors. In procedure 1, we set thresholds of support α and confidence λ , for FTDA algorithm, and the algorithm generates RS_i^t and RS_j^{t+1} from datasets at t and $t+1$ periods, respectively. In procedure 2, we calculate SCI with β . In procedure 3, we set a RMT parameter to classify three different types of changed rules, emerging type, unexpected type, and added or perished type. In procedure 4, we evaluate the change degree for three change types. When the support changing ratio, θ , satisfies significant threshold, ψ , we put the rule into final result list. We define three types of change rules, and show them in Definitions 1-3 (see [9] more details).

Definition 1. (Emerging Types) The fuzzy quantitative association rule r_j^{t+1} is defined as an emerging rule with regard to rule r_j^t . If the conditions, $SD_{ijL}^{item}=SD_{ijL}^{qlt}=SD_{ijR}^{item}=SD_{ijR}^{qlt}=1$, hold. And, the $sup(r_j^{t+1})$ and the $sup(r_j^t)$ are significantly different.

Definition 2. (Unexpected Types) The fuzzy quantitative association rule r_j^{t+1} is defined as an unexpected rule with respect to r_j^t . If the conditions, $SD_{ijL}^{item}=SD_{ijL}^{qlt}=SD_{ijR}^{item}, SD_{ijR}^{qlt}<1$, and $SCI_{ij}>RMT$ hold. And, the $sup(r_j^{t+1})$ and $sup(r_j^t)$ are significantly different.

Definition 3. (Added/Perished Type) r_j^{t+1} is defined as an added rule with respect to all rules discovered in RS^t if $MaxSCI_j^{t+1} \leq RMT$ hold. Conversely, r_j^t is defined as a perished rule with respect to all rules discovered in RS^{t+1} if $MaxSCI_j^t \leq RMT$ hold.

A huge number of 6 types change rules are onerous for teachers making decisions, but not all of change rules are worth to pay teacher's attention. So, we have to calculate the significance for every rule in these types, and set a threshold and an index help teachers make decision more efficient. The threshold, ψ , is to filter trivial change rules out and only keeps significant ones and the index θ is to measure the significance between r_i^t and r_j^{t+1} . For calculating the significance of emerging rules, we define support changing ratio. For calculating the significance of unexpected rules, we define support changing ratio. For calculating the significance of an added rule and a perished rule, we define support changing ratio of added rules.

5. Experimental results

We use a real-life datasets to evaluate the effectiveness of the FuzzChgMining model. The dataset is a Bachelor student grade dataset, collected over three semesters of the department of business administration of the National Chung Cheng University (NCCU) in Taiwan which the semester time-period is from 2009 (TP1), 2010 (TP2) to 2011 (TP3). For mining fuzzy behavior-interval quantitative association rules, linguistic-terms and fuzzy membership functions are introduced to

represent a behavior-interval. In this experiment, we used 7 linguistic-terms to represent the quantitative linguistic-term (*qlt*) of behavior-intervals.

We compared the intersection between Fig. 1 and Fig. 2, we focused on emerging decrease rules with $\beta = 0.4$, $\psi = 0.01$, $\alpha = 0.01$, RMT=0.75 from TP1 to TP2, and from TP2 to TP3. As our filtered, the numbers of emerging decrease rules from TP1 to TP2 is 30, and from TP2 to TP3 is 8. As our matched, we discovered only seven rules are intersection rules, and shown in Tab. 1. The rule 5 in Tab. 1 is with the highest support in TP2-TP3, and with the highest tendency. Other type change rules also can be explored by the same way.

6. Conclusion

Fuzzy quantitative association rule mining is a useful method for discovering student behaviors through time-periods. It is beneficial to educators for quick and easy decision making. By updating educator's knowledge over time trends, educators make appropriate teaching strategies and practices. Educators adopt our model to make better summative assessments, detect students' problems and understand shifts of students' characteristics, and then adjust the enrollment policies or pedagogies.

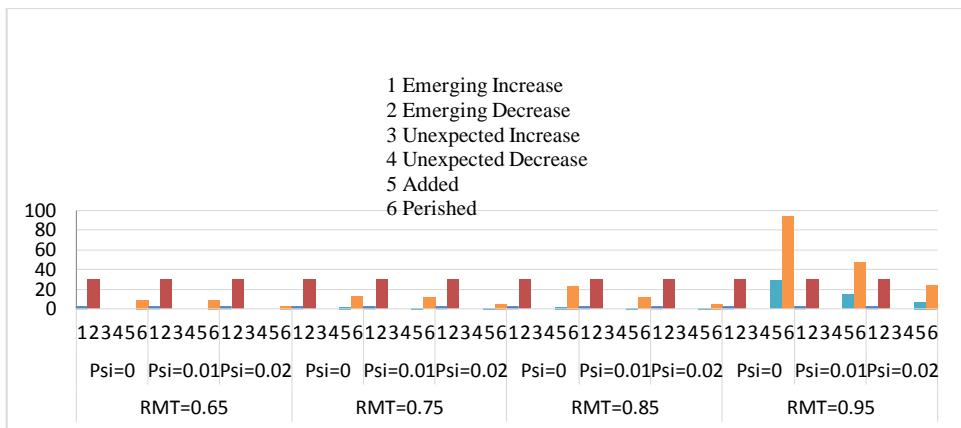


Fig. 1: The number of change rules for six change types with B0.4-A0.01 (TP1-TP2).

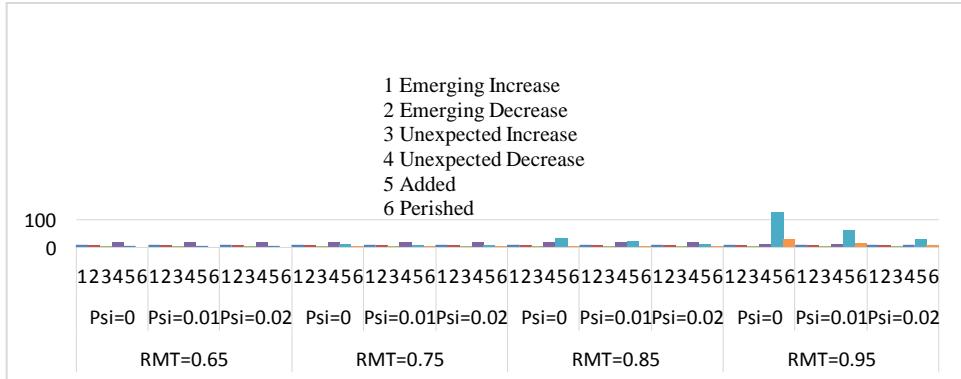


Fig. 2: The number of change rules for six change types with B0.4-A0.01 (TP2-TP3).

Tab. 1: IntersectionEmerging Decrease rulesbetweenTP1-TP2 and TP2-TP3 with $\beta=0.4$, $\psi=0.01$, $\alpha=0.01$ and RMT=0.75

| Fuzzy Quantitative Association Rule | TP1-TP2 | | TP2-TP3 | | Tendency |
|--|---------|----------|---------|----------|----------|
| | Support | θ | Support | θ | |
| r_1 : (Attendance, Excellent) \wedge (MidTerm, Worst) \rightarrow (Semester, Low) | 0.098 | 0.424 | 0.055 | 0.439 | 0.561 |
| r_2 : (Attendance, Excellent) \wedge (FinalReport, High) \rightarrow (Semester, Low) | 0.038 | 0.24 | 0.016 | 0.579 | 0.421 |
| r_3 : (Attendance, Excellent) \wedge (FinalReport, VeryHigh) \rightarrow (Semester, Low) | 0.076 | 0.397 | 0.057 | 0.25 | 0.750 |
| r_4 : (MidTerm, Worse) \wedge (FinalReport, High) \rightarrow (Semester, Low) | 0.044 | 0.228 | 0.028 | 0.364 | 0.636 |
| r_5 : (MidTerm, Worse) \wedge (FinalReport, VeryHigh) \rightarrow (Semester, Low) | 0.079 | 0.521 | 0.064 | 0.19 | 0.810 |
| r_6 : (MidTerm, Worse) \wedge (FinalTerm, Low) \rightarrow (Semester, Low) | 0.053 | 0.411 | 0.012 | 0.774 | 0.226 |
| r_7 : (MidTerm, Worse) \wedge (FinalTerm, High) \rightarrow (Semester, Low) | 0.022 | 0.542 | 0.012 | 0.455 | 0.545 |

7. References

- [1] A. W., Astin, Assessment for Excellence: The Philosophy and Practice of Assessment and Evaluation in Higher Education, 1993.
- [2] R., Agrawal, and R., Srikant, “Fast Algorithms for Mining Association Rules in Large Databases,” *Proc. of the 20th International Conference on Very Large Data Bases*, pp. 487-499, 1994.
- [3] M., Böttcher, “Contrast and Change Mining,” *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 1(3), pp. 215-230, 2011.
- [4] C., Romero, and S., Ventura, “Educational Data Mining: a Survey from 1995 to 2005,” *Expert Systems with Applications*, 1(33), pp. 135-146, 2007.
- [5] R. Srikant and R. Agrawal, “Mining Quantitative Association Rules in Large Relational Tables,” *Pro. of the 1996 ACM SIGMOD international conference on Management of data*, pp. 1-12.
- [6] M. Delgado, N. Marín, D. Sánchez, and M.-A. Vila, “Fuzzy Association Rules: General Model and Applications,” *IEEE Transactions on Fuzzy Systems*, 11(2), pp. 214-225, 2003.
- [7] Y.-L., Chen and C.-K., Huang, Tony, “Discovering Fuzzy Time-Interval Sequential Patterns in Sequence Databases,” *IEEE Transactions on Sys-*tems, Man and Cybernetics-Part B, 35(5), pp. 959-972, 2005.
- [8] T.-P., Hong, C.-S., Kuo, and Chi, S.-C., “Mining Association Rules from quantitative data,” *Intelligent Data Analysis*, 3(5), pp. 363-376, 1999.
- [9] C.-H., Huang, Tony, “Mining the Change of Customer Behavior in Fuzzy Time-interval Sequential Patterns,” *Applied Soft Computing*, 12, pp. 1068-1086, 2012.