Weakest Integrity Traits Identification of Teachers using Association Rule Mining

Mohamad Farhan Mohamad Mohsin, Faudziah Ahmad, Aniza Mohd Din, Ku Ruhana Ku Mahamud, Roshidi Din
School of Computing, UUM College of Arts & Sciences
Universiti Utara Malaysia
Kedah, Malaysia
{farhan, fudz, anizamd, ruhana, roshidi}@uum.edu.my

Abstract— **The government has arranged many programs for teacher development however the training is organized to fit yearly calendar without considering the right teacher for the right training. The selection of teacher to attend training is done randomly, by rotation and not based on their work performance. This paper investigate the weakest integrity trait of teacher using association rule technique with the aim can assists the school management to organize training related to teachers integrity performance and avoid sending a wrong teacher for a training. A dataset of Trainees Integrity Dataset (TID) representing 1500 secondary school teachers in Langkawi Island, Malaysia in the year 2009 were pre-processed and mined using apriori. The knowledge from the mining was analyzed based on demographic and integrity trait of teacher. The finding indicates that adaptability and stability are the weakest integrity trait among teachers. Besides that, the analysis also unable to prove that demographic factor such as the age and gender of teachers reflect their low integrity performance. The finding can be a guideline for school management to propose a suitable training program for teacher to improve integrity mainly at the adaptability and stability trait.**

Keywords- association mining, data mining, teacher integrity, big five model

I. INTRODUCTION

Teacher is worldwide known as an important person in helping the country develop a valuable citizen. They hold high responsibility to educate student as preparation for them to lead the country. This task is difficult therefore, teachers must have high integrity, and they must establish integrity within the school to accomplish the goal.

To maintain the high level of integrity, the government through the Ministry of Education (MoE) has arranged many programs for teacher development through training, workshop, and seminar. Currently, the training program is randomly organized to fit yearly calendar without considering the right teacher for the right training. The selection of trainee to attend training is done randomly, sometimes by rotation and most of the selection is not based on their work performance. Up until now, there is no intelligent model to assist school management to determine the integrity level of teacher and assign them to right training program. This situation leads to difficulty for school management to identify problematic teacher and can mistakenly sent wrong teacher to attend training. If the requirement of training program is mismatched with trainee, there is high possibility the objective of the training cannot be achieved.

Teachers who do not possess such negative attitudes would not be suitable for an integrity program. If this category of teachers is selected by random chance, this results in a mismatch of supply and demand of needs and cause failure in integrity training programs to achieve its goals. According to some studies, selecting the wrong person can cost businesses between 30% and 200% of a person's annual salary. If a role worth $70,000 a year is filled with the wrong candidate, it could end up costing between $21,000 and $140,000 – a loss most businesses can't afford to make [2]. Even though there is a keen action form the government to raise the teacher integrity, there appear to be a number of reports that some teachers unable to perform their tasks efficiently, have problem with students and some of them involves in criminal [9,11].

Integrity is viewed as the quality of having an intuitive sense of honesty and truthfulness of a human in completing task. It is normally a product of culture with the relation of good trait and discipline of human in all situations [12]. As the result of high integrity, human will be more accountable and responsible towards his job and struggle to reach objective. It is a key success to everyone and many organizations have benefited enormous performance increases when their employees own good integrity behavior. As stated in [1] nothing can work without the present of integrity.

Related to integrity, [17] stressed out that the integrity among teacher is critically important and sometimes the solution towards integrity issues is difficult to answered. Finding teacher with high integrity characteristic is not an easy since there is no agreement as to what makes a teacher has high quality attitude in school. Such criteria like the experiences that the teacher has, the education level, and the performances achieved by their students can be used to measure teacher integrity. Besides that, the integrity test such as the Big Five Model [3] and personality test [16] also can be a guideline to measure the level of teacher's integrity

In school, teachers must uphold the highest integrity and be a good role models to their students at all times. Their responsibility towards students are not limited during school hours but endlessly continued after the class end. Teachers are the nearest person who can be example for students besides their parent. Consider as an idol for student, teacher need to be fulfill with a good package of moral quality. For the reason to attract students to behave nicely, teachers must have a quality mind and good attitude to impress student. They need to be familiar with the students as well as the school community to be able to successfully educate students. Good quality teachers must be able to target and plan how to manage students becoming good quality students [10, 13]. Besides that, they

also need to be a very understanding person and show their respect of the student's behavior [14]. Stated by [17], an effective teacher is capable of inspiring greater learning gains in their students when compared with teacher with low teaching quality. Those traits have a strong relationship with teacher's integrity. With the high capacity of integrity, teacher will be motivated to educate student and helps the country to develop human capital.

Since the MoE faces problem in discovering problematic teacher related to integrity, a model to identify the weakest integrity trait of teacher is proposed in this paper. The aim of this model is to assist the school management to discover integrity trait that are mostly lack among teachers and proposes a suitable training program for them. To achieve that, data mining algorithm called apriori is used to mined Teacher Integrity Dataset (TID). The dataset consists of 1500 information about demographic and integrity score of secondary school teacher in Langkawi Island, Malaysia in the year 2009.

This paper is organized as follows. Section II outlines the basic notion of AR. The model development of the study is discussed in section III. The experiment and result will be presented in section IV and final sections conclude this work.

II. ASSOCIATION RULE

In this section, the basic of association rule mining is discussed. Association rule mining or AR mining is the identification of frequent items that occur in a database of transaction. Each item (i_j) in a transaction is an important feature that contributed to the computation of item set and generation of rules. Basically, let $I = \{i_1, i_2, \ldots, i_m\}$ be a set of item and D be a set of transactions, where each transaction T is a set of items such as that $T \subseteq I$. An AR is an implication of form $X \rightarrow Y$, where $X \subset I$, $Y \subset I$, and $X \cap Y = \varnothing$. The rule $X \rightarrow Y$ has support s in the transaction D if s% of transactions id D contain $X \cup Y$. The rule $X \rightarrow Y$ holds in the transaction with confidence c if c% of transaction in D that contain X also contain Y. AR mining's processes begin with searching for frequent item set with user-specified minimum support and later rules are contrasted by binding the frequent item with its values and class. Strong rules are defined as rules that have confidence more than the minimum confidence threshold.

.

III. METHODOLOGY

This chapter describes how this work was conducted. The methodology is divided into two stages that are preparing data for AR Miner and pattern extraction. Fig. 1 illustratively depicts the methodology of this work.

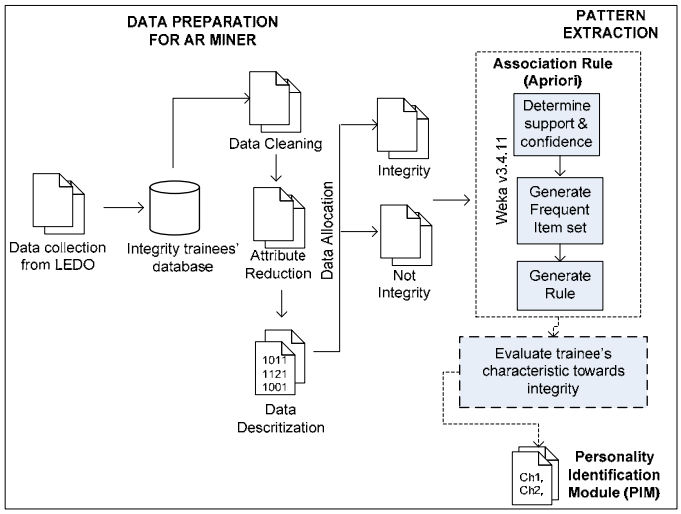

Figure 1. Methodology of the study

A. Preparing data for AR Miner

In this phase, two main activities were involved that were data collection and data preprocessing. Firstly, the experiment data for this study was collected from organization. A set of filled questionnaires of secondary teachers in Pulau Langkawi, Kedah, Malaysia was taken as a case study. The questionnaire was designed by Langkawi Education District Office (LEDO) and it was distribute to teachers while they attending training or seminar in the year 2009. There were 1500 respondents and 26 attributes including a decision class. The decision class holds two values "Integrity" if the accumulative score of integrity test is more than 80% and "Not Integrity" if the score is lesser than 80%. Out of 1500, 36.17% respondents were male and 63.87% were female teachers. The filled questionnaires were input into excel and named as Trainees Integrity Dataset (TID).

The TID represents two information's that are demographic background and five key traits on integrity of teachers. There integrity elements are inherited from the Big Five Model (FFM) [3] that are need for stability, extraversion, originality, accommodation, and consolidation. Each element is represented with four set of questions. Theoretically, the Big Five model identifies traits and structure of human personality and it act as a guideline to measure integrity of a human [3]. The model has been widely used in various areas such as motivating a human. FFM has been found to produce consistent results over the past years and is the basis of characterizing personality [4, 5, 6]. Measuring the integrity of a training candidate can help authority such as employer and training provider to identify the levels of integrity of a person. Table I lists the integrity traits as stated in [3].

TABLE I. INTEGRITY TRAITS BASED ON FIVE FACTOR MODEL [3]

Traits	Details
Openness (Originality, Imagination, openness to experience)	The openness of a trainee towards new experience. Human with openness is said to own more intellect, imaginative, curious, appreciative of art, and sensitive to beauty.
Conscientiousness (Will to achieve, Consolidation)	A tendency to show self-discipline, act dutifully, and aim for achievement against measures or outside expectations.
Extraversion (Positive Emotionality, Sociability)	Related to positive emotions and surgency. It has tendency to seek out stimulation, has engagement with the external world, always enjoy being with people, and often equipped with full of energy as well as an action-oriented individuals
Adaptability (Agreeableness, Accommodation)	Is a social harmony trait and has tendency to be compassionate and cooperative rather than suspicious and antagonistic towards others.
Need For Stability (Negative Emotionality, Neuroticism)	Called as emotional instability that has relation to how a trainee respond to stress, experience negative emotions

The second step of preparing data is data preprocessing. In this activity, the problem in TID were identified and resolved. During preprocessing task, all dataset were pre-processed where all unknown numeric attributes were replaced with mean value while max value for character attributes. Since the AR algorithm in WEKA only accepts nominal type data, the dataset were then discretized using boolean reasoning technique [15]. To increase the mining speed and accuracy, only the important attribute was given priority for mining and those attributes which were not related to the study were ignored.

During the selection process, 20 attributes which represent integrity test score were reclassified into five new groups based on the FFM model [3]. The new groups were stability, extraversion, originality, accommodation, and consolidation. Each attribute represent an integrity question with the maximum score was 5. In this process, the sum score of each attribute based on the type of integrity were aggregated and the total score was 20. For example, the attributes A1, A2, A3, and A4 are the set of integrity question for Extraversion. Let say the score is given as A1=1, A2=4, A3=4, A4=1 then the total accumulative score for Extraversion is 10. Based on the accumulative score, the value is reclassified into two classes, either "15-20" or "0-14". The "15-20" group was considered as high score and "0-14" as lower. Fig. 2 summarizes the process of attribute selection and reduction.

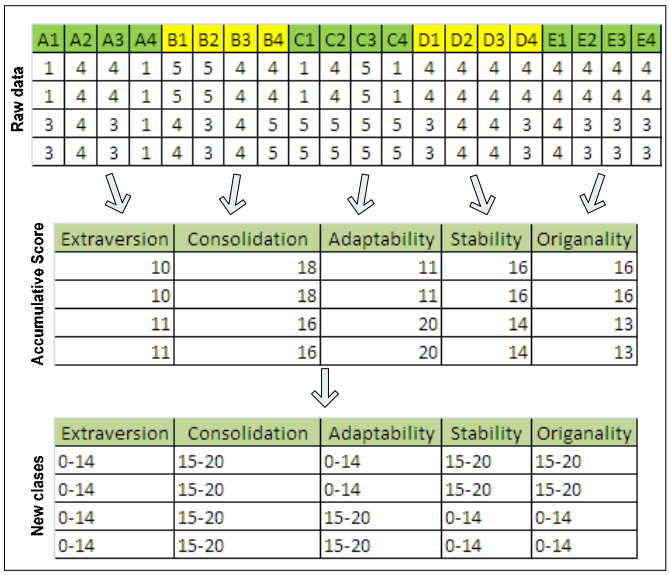

Figure 2. Attribute selection and reduction

In final step of data preprocessing, TID was separated into "Integrity" and "Not Integrity" group and only "Not Integrity" group was presented to AR. The reason is to investigate the weakest trait of integrity that reflects teacher performance. Besides can concentrates into one group, this step can eliminate longer mining time as well as can reduce the number of knowledge generated.

After pre-processed, there were 450 records out of 1500 and 10 attributes. The attributes were gender, age, position, grade, department, stability, extraversion, originality, accommodation, and consolidation. Table II shows the statistic of "Not Integrity" group based on gender.

TABLE II. THE STATISTIC OF NOT INTEGRITY GROUP BASED ON GENDER READY FOR MINING

Gender		
Male	**Female**	**Total**
151	303	454

B. *Pattern Extraction*

In pattern extraction, AR algorithm called Apriori in WEKA data analysis tool was chosen as pattern extraction tool [7]. The clean TID dataset was presented to Apriori algorithm and during mining, the length of frequent item set, support, and confidence value of each itemset was recorded. In this study,

the minimum support value was set to 10% while confidence value was limited to upper than 95%. The maximum number of rule was set to 1000. The knowledge extracted from apriori was represented in term of frequent teacher integrity trait matching and the rules were further analyzed. The next section will discuss the finding of this study.

IV. ANALYSIS AND FINDING

This section reports the finding of this work. After mined with Apriori, there were 1000 rules generated and the algorithm also had generated interesting integrity trait with the maximum combination length is 7. Table 3 shows the number of interesting trait combination (***f_Char***) towards frequent item set length. L_n is representing the length and *n* is the number of attribute combination. For example, if L_4, there is four attributes combination. From the table III, the highest *f_Char* was generated in L_4 and it was followed by L_3. The number of *f_Char* is getting lesser when length increases.

TABLE III. THE THE NUMBER OF F_CHAR ACCORDING TO L_N

	L_n						
	L_1	L_2	L_3	L_4	L_5	L_6	L_7
f_Char	21	155	432	533	299	72	5

The results were further analyzed. We continued our analysis on the *f_Char* by focusing at L_4 until L_7. The L_1-L_3 group were ignored the trait is too compact and short. This exploration was divided into two parts. Firstly was to investigate the demographic trait. Secondly was to investigate the integrity criteria trait. Table IV and V depict the sample of *f_Char* when the number of integrity trait combination is 4. *S%* in both tables are the support value.

TABLE IV. THE THE SAMPLE OF F_CHAR WHEN L=4

f_Char (L_4)	Length	S%
Adaptability, **Stability**, Originality	3	7
Extraversion, **Adaptability**, **Stability**	3	4
Extraversion, **Adaptability**, **Originality**	3	4
Consolidation, **Adaptability**, **Stability**	3	6
Consolidation, **Stability**, **Originality**	3	4
Extraversion, Adaptability, **Stability**	3	4
Extraversion, Consolidation, **Stability**	3	3
Extraversion, Consolidation, **Adaptability**	3	3
Extraversion, Consolidation, **Adaptability**, **Stability**	4	1
Extraversion, **Adaptability**, **Stability**, Originality	4	1
Extraversion, Consolidation, **Stability**, Originality	4	1
Extraversion, Consolidation, **Adaptability**, **Stability**	4	1

TABLE V. THE SAMPLE OF F_CHAR WHEN L=5

f_Char (L_5)	Length	S%
Adaptability, **Stability**, Originality	3	4
Extraversion, **Adaptability** , Stability	3	
Consolidation, **Adaptability**, **Stability**	3	6
Consolidation, Stability, Originality	3	
Extraversion, **Adaptability**, **Stability**	3	2
Extraversion, Consolidation, **Stability**	3	2
Extraversion, Consolidation, **Adaptability**	3	2
Extraversion, **Adaptability**, Consolidation, Originality	4	3
Extraversion, Consolidation, **Adaptability**, **Stability**, Originality	5	1

Under demographic analysis, we found that the type of gender is not an important trait to determine the integrity level of teachers. Even though the gender attribute had appeared 55.08% in L_4, both sexes have similar probability to behave with low integrity. Moreover, the study also found that the age group within "26-30" was the highest contributor in all *f_Char* and interestingly, most of them were female. Then, the salary grade and position were found not as a strong trait to determine integrity.

Under integrity traits, we concentrated our analysis at the weakest score of integrity. The *f_Char* were scanned and all five integrity traits with the weakest score "0-15" were marked shown in figure 4 with the font bold. The result of the scanning process is depicted in Table V.

TABLE VI. THE THE RESULT OF SCANNING PROCESS TOWARDS INTEGRITY SCORE "0-15" IN F_CHAR

f_Char	E	C	A	S	O
L_4	73	54	179	165	61
L_5	26	18	145	106	24
L_6	2	2	43	40	0
L_7	0	0	4	0	0
Total	101	74	371	311	85

E- Extraversion **C**- Consolidation **A**- Adaptability **S**-Stability **O**-Originality

From the table, adaptability is the weakest integrity trait among teachers. In all *n*, adaptability has the highest score which indicates most teacher are lacking at the adaptability aspect. Then, it was closely followed by stability which also indicates the less integrity among teachers. The score of other trait – extraversion, consolidation, and originality is considered low compared to adaptability and stability.

Then, we further investigated the combination of more than one trait. The observation reveals that the adaptability and stability traits were frequently occurred together in all *f_Char*. For example, in L_4, there were 47 combinations of adaptability and stability. This indicates the teacher who lack at adaptability will also have problem related to stability. Beside them, the combination of "stability and originality" and "adaptability and originality" was also high. This explanation is summarized in

Table VI. It indicates the frequency of the weakest trait for two traits combination in L_4 and L_5.

TABLE VII. THE FREQUENCY OF THE WEAKEST TRAIT FOR TWO TRAITS COMBINATION IN L_4 AND L_5

	E	C	A	S	O
L_4					
E		-	16	-	-
C	12		-	18	11
A	-	-		47	23
S	19	-	22		25
O	15	-	-	-	
L_5					
E		5	9	8	5
C	15		-	-	-
A	6	-		34	9
S	11	-	-		15
O	-	-	-	-	
L_6					
E		-	1	-	-
C			-	-	-
A	-	-		10	1
S	-	-			1
O	-	-	-	-	

E- Extraversion **C**- Consolidation **A**- Adaptability **S**-Stability **O**-Originality

For longer combination, association rule generated several interesting traits matching however the number of combination were small. The table VII summarizes the interesting trait generated in L_4 with the number of combination are three and four. From the table VII, adaptability and stability exits in all combinations. Besides that, extraversion is looked frequently paired with consolidation traits.

TABLE VIII. THE INTERESTING TRAIT GENERATED IN L_4 WITH THE NUMBER OF COMBINATION ARE THREE AND FOUR

f_Char (L_4)	Length	S%
Adaptability, Stability, Originality	3	7
Extraversion, Adaptability, Stability	3	4
Extraversion, Adaptability, Originality	3	4
Consolidation, Adaptability, Stability	3	6
Consolidation, Stability, Originality	3	4
Extraversion, Adaptability, Stability	3	4
Extraversion, Consolidation, Stability	3	3
Extraversion, Consolidation, Adaptability	3	3
Extraversion, Consolidation, Adaptability, Stability	4	1
Extraversion, Adaptability, Stability, Originality	4	1
Extraversion, Consolidation, Stability, Originality	4	1
Extraversion, Consolidation, Adaptability, Stability	4	1

In general, the results indicate adaptability and stability are the weakest trait among teacher due the frequency of both traits exits in *f_char*. The other three characteristics are also exits in *f_char* but due to their occurrence are low, we decided to ignore them as weak integrity traits. According to literature; adaptability refers to the tendency of teacher to be compassionate and cooperative rather than suspicious and antagonistic towards others. While stability refers to the ability of teacher respond to stress, high work load and dealing with negative emotions. Since adaptability and stability are the weakest trait among teacher, the school management may organize suitable training to improve both traits.

V. CONCLUSION

The paper discovers the weakest integrity trait of teachers using association rule technique. The aim of this study is to assist the school management in identifying integrity trait that are mostly lack among teachers and proposes a suitable training program for them. Moreover, the model also can help the school management to shortlisted list of candidate for training. Through the analysis, adaptability and stability are found as the weakest trait among the teacher. The school management may use this guideline to propose a suitable training program for teacher to improve integrity mainly at the adaptability and stability. This study is aimed to extend the literature on developing successful training programs, intelligent trainee selection model as well as shortlisted the list of candidate for training.

REFERENCES

[1] K. Christensen, The Magazine of the Rotman School of Management, Fall 2009, pp. 16-20.

[2] M. Dowding. *Northern Rivers Business Magazine*: Autumn 2011.

[3] P.J. Howard, and J. M. Howard. "The Owner's Manual for Personality at Work". Austin, Texas: Bard Press, 2001.

[4] L.M. Saulsman and A.C., Page. "The five-factor model and personality disorder empirical literature: A meta-analytic review". Clinical Psychology Review 23 (8): 1055–1085, 2004.

[5] M.R. Barrick, and M.K. Mount. "The Big Five Personality Dimensions and Job Performance: A Meta-Analysis". Personnel Psychology 44: 1–26, 1991.

[6] M.K. Mount and M.R. Barrick. "Five reasons why the "Big Five" article has been frequently cited". Personnel Psychology 51: 849–857, 1998.

[7] I.H., Wittenand and E., Frank. "Data Mining: Practical machine learning tools and techniques". Morgan Kaufmann: San Francisco, 2005

[9] Azam Ahmad. School teacher suspected rape 16 years old student. The Star, 2010. (online). www.thestar.com (5th July 2010)

[10] Baker. J. Nobody's Perfect, But You Have To Be: The Power of Personal Integrity in Effective Preaching, New York: Mc Graw-Hill, 1990

[11] Kamarul Hassan. Guru Sekolah disyaki Merogol Pelajar Tingkatan 2. Utusan Malaysia, 2010. (Online) www.utusan.com.my (5th July 2010)

[12] Popper, K. R. Objective Knowledge in higher educational systems processes. Proceeding of 5th international conference, University of Sydney, Australia. pp. 60-69, 1983.

[13] Nillsen, R. The concept of Integrity in Teaching and Learning, Symposium on Promoting Academic Integrity, Newcastle NSW, 2004.

[14] Thornton, M. Working with Integrity, Code Of Business Product, BhpBiliton, 2004. (0nline) www.bhpbilliton.com (1st August 2010)

[15] Nguyen, H.S., "Descretization problem for rough set methods". In Proc of First Int. Conf. on Rough Set and Current Trend in Computing, 1998, pp. 545-552.

[16] Daniel, N. "A test of character". The Guardian (London), 2009.

[17] Starling, J.W. The effects of teacher stability on third grade student achievement as measured by the North Carolina End-of-Grade tests in reading and mathematics. Gardner-Webb University. Phd Thesis, 2009.