Personality Mining

Ritu Chaturvedi

With the advent and increasing use of Internet technology, teaching methodologies other than traditional classroom teaching, such as web-based distance education have emerged and have gained popularity. According to a website of the Canadian Information Center for International Credentials (CICIC), Ontario's colleges and universities offer around 10,000 courses and over 800 online programs and distance learning opportunities. One of the major problems facing distance education courses is how teachers, who typically have no direct interaction with their students, adjust the pace of their teaching to that of their students and make proactive decisions to help students succeed in their course, without meeting them face to face. Personality mining is the study of mining student personality through observed and logged student data. Data mining techniques such as clustering and association rule mining are often used to correlate student characteristics such as personality, behavior and learning styles to their performance but the research is still in its early stages and can be improved. Various researchers have tried to solve this problem by using existing statistical and data mining techniques and by proposing improvements to such techniques. This survey is about their research.

Categories and Subject Descriptors: A.1 [**Introductory and Survey**]: ; I.2.7 [**Artificial Intelligence**]: Learning—*Knowledge acquisition*; H.2.4 [**Information Systems**]: Database Application—*Data Mining*

General Terms: Personality, e-learning

Additional Key Words and Phrases: data mining, machine learning, personality, e-learning

Contents

1. INTRODUCTION

This survey concerns research which attempts to solve the problem of improving learning outcomes in e-learning distance courses by studying the impact of students' personality and learning preferences on his/her learning process. In such e-learning distance courses, where students do not have a direct interaction with the teacher, it becomes challenging for the teacher to provide individualized assistance to students and to adjust the pace of their teaching appropriately (just like how a teacher in a traditional classroom adjusts) to effectively meet the course's learning outcomes. Research on correlating students' personality and learning preferences with their performance grades will empower the teachers teaching such distance courses towards meeting the course objectives.

Most of the relevant research papers were found by searching Google Scholar with the keywords "Personality Mining". Additional relevant papers were found while searching for journal and conference papers in the University of Windsor's leddy library collection of ACM journals, IEEE proceedings and LNCS lecture notes.

There were nine journal papers, thirteen conference papers and one Master's thesis that were identified as closely related to this survey. They are listed in the bibliography. Ten of these papers were selected to be annotated. The remainder of the survey is structured as follows : section 2.1 contains reviews of 8 papers, five of which used a data mining technique called association rule mining in their research to find effective personalized learning paths ([Chen 2005], [Du et al. 2005],[Jin et al. 2006], [Tian et al. 2007] and [Zheng et al. 2008]); two of them used clustering as the mining technique ([Jin et al. 2006] and [Zakrzewska 2009]); [Huang et al. 2008] in section 2.1.3 mention the use of data mining techniques in their experiments but do not specify the exact technique. Section 2.2 contains reviews of 2 papers [Fatahi et al. 2009]and [Arockiam et al. 2012] that use statistical and database programming techniques to propose models that study the impact of student personality and behavior on their learning styles and performances. Sections 3, 4, 5 and 6 contain the concluding comments, acknowledgments, annotations of ten selected papers and referees respectively.

In e-learning distance courses, since there is no direct interaction of students with teachers, automating the process of correlating the personality and behavior of students with their performances is crucial to achieving personalized learning strategies and simulating the role of a teacher in a traditional classroom setting. Although such distance courses seem to be very popular in United States and

Canada (e.g. according to a website of 'The Canadian Information Center for International Credentials (CICIC) ', Ontario's colleges and universities offer around 10,000 courses and over 800 online programs and distance learning opportunities), much of the research done in this area is by a research group from a University in China. This research group seems to be highly active in the area of personality mining, especially using the association rule mining technique. In fact, 11 of the 23 papers in the bibliography are either from China or Taiwan. Most of the papers reviewed in this survey use data mining techniques such as association rule mining and clustering in their research, although there are two papers in which the authors use traditional database programming and statistical techniques.

2. TECHNIQUES USED

2.1 Data Mining techniques

A review of the ten selected papers identifies association rule mining as the most commonly applied data mining technique in the research of personality mining. Clustering using k-means and hierarchical clustering is also used by some researchers to find groups of students that match with a certain learning strategy. Research papers that are presented in this section use these techniques to solve the considered problem.

2.1.1 *Association Rule Mining.* Du et al. [Du et al. 2005] propose a learner model in 2005 that consists of a personality model (PM) and a behavior model (BM), where personality model stores the learner's personality using Cattell's 16 personality factors and behavior model is represented by six different aspects of the sequence of actions that a learner takes while using the web page : behavior on courseware learning, tests / assignments, postings and interaction among students and teachers through the use of discussion boards, chat rooms etc. Each attribute of PM and BM is mapped to an integer value of 1, 2 or 3 representing high, middle and low. This transformed data is then mined using the apriori algorithm to analyse the relation between personality and behavior and generate association rules of the form Behavior => Personality. The authors claim that their algorithm is an improvement over the well-known apriori algorithm because it avoids the generation of redundant rules substantially, thereby reducing the complexity of the algorithm. The authors conducted experiments with 324 students of Xi'an Jiaotong University and collected their attributes such as the 16 Cattell's personality factors and 146,000 log records with attributes for the behavior model. Then they apply the proposed modified-Apriori algorithm to generate association rules between behavior and personality.

Tian et al. [Tian et al. 2007] propose an algorithm that builds a learner model using learner's personality, behavior and learning preferences and strategies. They collected data of 1013 students in an English language course in a University in China. This data consisted of student's personality such as warmth and emotional stability using Cattel's 16 personality factors, learning strategies such as mental (active/reflective), behavioral (interpersonal) and self-regulatory(motivation) using a learning strategy survey questionnaire and learning behavior such as the time spent on accessing course-ware materials using logged website data. Their algorithm then mines the relationship between personality and behavior using associative rule mining, personality and learning strategies using rough set theory

and learning strategies and learner's behavior(the authors give no details on how this matching is performed) to generate personalized learning strategies. Tian et al. [Tian et al. 2007] claim that their idea of mining relationships between personality and behavior and between personality and learning strategies was novel and effective. In particular, they mention one student who had a negative personality, lacked self-confidence, had strong inferiority complex and did not perform well in the course was seen to enhance his ability of answering questions correctly from 66.7% to 87.5% using auto-generated personalized learning strategies such as "participate more in group discussions".

In 2007, Huang et al. [Huang et al. 2007] propose an improved apriori algorithm that generates non-redundant association rules between behavior and personality. The algorithm combines learner's personality and behavior data and stores them in a database, generates a datacube using learning behavior, personality characters and time as the three dimensions, generates frequent k-itemsets for each dimension using apriori algorithm, generates association rules using these frequent itemsets and finally, it establishes a correlation between the itemsets in each rule generated retaining only those that meet a certain threshold. The authors claim that a total of 9766 association rules were generated from experiments conducted with 360 learners and only 6794 rules were retained using a correlation factor < 166.7.

Zheng et al. [Zheng et al. 2008] in 2008 claim that most of the previous studies on personality characterstics and learning strategies of learners are done in psychology where methods such as correlation and regression analysis and discriminator functions are used. These methods are subjective and time-consuming as they use human expert knowledge in their analysis to find the key attributes of learners that descibe their personality and learning styles. Perhaps, the authors were not aware of the earlier work done by researchers using data mining techniques. collect the learner's personality characterstics and store them in an information system. Their research first uses an algorithm based on rough-set theory to reduce the typically large set of personality attributes to a core set of key attributes. Rough-set theory is a mathematical model that works best for situations that deal with uncertain, imprecise and incomplete data. Then association rule mining is applied to generate rules such as "students with personality described by attributes set {Emotional Stability, Dominance, Social Boldness, Apprehension, Self-relaince, Experimental, Individuality} fit best with the learning strategy known as metacognitive strategy". The authors claim that the rough-set algorithm helped them reduce the attributes in the information system by at least 50%. They also claim to have generated meaningful rules that correlate the 5 learning strategies with groups of personality traits. Details on their experiments is given in section 5.

2.1.2 *Clustering.* Jin et al. [Jin et al. 2006] present an analysis of the various existing clustering methods to validate the feasibility of using k-means clustering algorithm as the most appropriate method for grouping learners. Then they build a learner model that consists of attributes that represent students' personality (e.g. warmth), motivation (e.g. curiosity), study style (e.g. visual/hands-on, serial/random), study concept (e.g. self-management), strategy (e.g. memorize/ cognize) and behavior on courseware learning (e.g. average time spent on each login). Finally, they apply k-means algorithm using Euclidean distance function to

group students with same personality into one cluster and match that to a learning strategy appropriate for that group. The authors also conducted experiments using another clustering method called BIRCH to compare and validate the performance and accuracy of the suggested k-means algorithm. They conducted experiments with 500 students from Xi'an Jiaotong University using Cattell's 16 Personality Factors questionnaire and filtered that to a dataset with 260 student records, eliminating those that took more than 55 minutes or less than 15 minutes to complete the questionnaire. The authors use two different methods to split their data into train and test datasets : one method split them randomly, whereas the other method used a 50-50 split. They claim to have the most accurate results in terms of performance of k-means for cluster sizes 4 and 5 with a 50-50 split model for training and test data, with cluster size of 5 outperforming all other cases.

Zakrzewska in 2009 [Zakrzewska 2009] made two general observations : 1. although a lot of research is being done on adaptive intelligent educational systems, there is still no agreement on what features such as student charactertics should be used or can be used in creating personlaized learning style in e-learning systems. 2. In many existing systems such as Arthur and 3DE, students were provided with different teaching strategies but the groups of students were made statically. The author emphasized the need to consider, not only the student's preferred learning style, but also his/her interface preferences. In lieu of this, she proposed a two-phase algorithm in which students are clustered , not only based on their learning styles but also on their usability preferences such as website layout and coloring scheme. In the first phase, groups of students, who are very similar and those who are outliers are formed using k-means clustering algorithm. The attributes used are student learning styles defined by Felder and Silverman (active/reflective, sensing/intuitive, visual/verbal/auditory and sequential/global). In the second phase, such groups are merged into bigger groups using hierarchical clustering algorithm until the maximum number of clusters is reached (the maximum number of clusters denotes the maximum number of teaching paths). Clusters are formed based on the threshold values given as input to the hierarchical clustering algorithm (such as a value t such that the similarity between the ith student's attributes and the cluster centroid $> t$, maximum number of clusters etc.). Any 1-element clusters indicate outliers such as outstanding students. The authors claim that their two-phase algorithm generates clusters that are very close in accuracy to the clusters generated by hierarchical clustering, although k-means performed better at dividing the students into almost equal clusters. Clusters from the learning styles perspectives showed some unexpected results since students with different learning styles were put together in a single cluster (e.g. sequential and global in the same cluster). Zakrzewska claims that her modified clustering algorithm yields better quality of clusters than the traditional algorithms and it does not to provide the exact number of clusters in advance and also that the best quality of clusters are obtained when both learning styles and website layout preferences such as color are used in the first phase of the algorithm. She also claims that the proposed algorithm successfully detects all outstanding students (outliers).

2.1.3 *Unspecified.* Huang et al. [Huang et al. 2008] propose an e-learning behavior model which they define as multi-dimensional and multi-hierarchical. The

dimensions include static (survey using questionnaires) and dynamic (using logged web data) ways of collecting personlaity information. The logged data has behavior information on how learners collect, process, publish and use materials on the e-learning website. The levels that the authors propose are primary (e.g. browsing the webpage), secondary (e.g. answering a quiz) and senior (e.g. collaborating with peers). The steps in the proposed algorithm are : 1. Collect learners' data (using questionnaires and logs). 2. Use a combination of statistical and mining methods to categorize the data collected into behavior, personality and e-learning resource characteristics. 3. Evaluate the characteristics found in step 2 as formative, immediate or stage evaluation. No details on how to evaluate are given. 4. Provide individualized guidance to the learner (done by a module called Intelligent feedback module). The authors did not specify the data mining technique they used to group student data. In fact, they did not conduct any experiments nor did they present any results. They claimed that the different dimensions and levels of hierarchy that they propose to impart personalized learning is effective but they do not back it up with any evidence.

2.1.4 *Summary .*

2.2 Other techniques

2.2.1 *Item Response theory.* Chen in 2005 [Chen 2005] propose an algorithm based on item response theory (IRT) which considers difficulty of course materials and learners' abilities to provide an appropriate learning path. The algorithm uses three different agents to accomplish this : interface agent (front-end), feedback and course recommendation agent (back-end). When a learner logins to the system for the first time, the interface agent collects learner's personal information and a set of key words that help locate course materials (called units) that the learner is interested in. A user has to register in the system, if he/she desires a personalized feedback. All this information is used to create the learner's profile. For first-time learners, the difficulty of the course unit chosen is set to medium. After the learner uses the course material, he/she is asked to answer two questions : 'How did you find the difficulty of the course material (response could be very hard, hard, moderate, easy, very easy)' and 'Do you understand the content of the course material ?' (response could be Yes or No). Then, the feedback agent collects these responses and re-evaluates the learner's abilities for that course unit using a maximum likelihood estimation function. The learner's profile is updated with these new abilities and the difficulty parameter of that course unit is ranked accordingly in the course database. The course recommendation agent is also informed of the learner's new abilities and therefore, it provides new recommendation for the learner. This process is repeated for other course units until the learner logs out of the system. A course material is modeled by the item characteristic function proposed by Rasch with a single difficulty parameter. The course material difficulty is adjusted (and ranked) according to a collaborative voting approach [Chen 2005]. The author conducted experiments on a course called 'Neural networks' that had three course units and 58 course materials. All the 210 learners logged into the system were Masters students. Chen [Chen 2005] claims that from a learner's perspective, the average degree of understanding of the recommended course materials is 0.825 and from

Year	Author(s)	Title of the paper	Technique	Major contribution
2005	Du, J.; Zheng, Q.; Li, H. and Yuan, W.	The research of mining association rules between personality and behavior of learner under web-based learning environment	Association rule mining	An improved apriori algorithm that substantially avoids the generation of redundant rules(16 personality and 6 behavioral attributes) , thereby reducing the complexity of the algorithm.
2006	Jin, D.; Qinghua, Z.; Jiao, D. and Zhiyong, G.	A Method for Learner Grouping Based on Personality Clustering	Clustering	An algorithm that validates the feasibility of using k-means clustering algorithm as the most appropriate method for grouping learners, before applying it to a learner model that consists of attributes that represent student's personality, study style,learning strategy and behavior on courseware learning to group students with same personality so that an appropriate learning strategycan be proposed.
2007	Huang, J.; Zhu, A. and Luo, Q.	Personality mining method in web based education system using data mining.	Association rule mining	An algorithm that uses a datacube to represent personality, behavior and time as its dimensions(d) to moderate the number of association rules generated by apriori algorithm (using roll-up, if there are no frequent itemsets ford and drill-down, if all d's itemsets are frequent).
2007	Tian Feng.; Zheng Z.; Gong Z.; Du, J. and Li, R	Personalized learning strategies in an intelligent e-learning environment	Association rule mining	An algorithm that blends two theories to mine relationships between personality(p), behavior(b) and learning strategies(ls) of students : rough-set theory to mine relationships between p and ls and association rule mining to mine relationships between p and b.
2008	Huang, K.; Li, F.; Zhao, M.; Wang, F. and Xu, X.	Design and Implement on E-learning Behavior Mine System	Unspecified	An algorithm that uses multi-dimensional (static - survey using questionnaires) and dynamic using logged web data) and multi-hierarchical (primary (e.g. browsing the webpage), secondary (e.g. answering a quiz) and senior (e.g. collaborating with peers)) data and a combination of statistical and mining techniques to group them.
2008	Zheng Q.; Wu X. and Li, H.	A rough set based approach to find learners key personality attributes in an e-learning environment	Association rule mining	An algorithm that first uses rough-set theory to reduce the number of key attributes that represent a student's personality and then uses association rule mining to find a correlation between this reduced set of attributes and a learning strategy.
2009	Zakrzewska, D.	Cluster Analysis in Personalized E-learning Systems	Clustering	An algorithm that first uses k-means to cluster students based on their learning styles and preferences (c), and then uses hierarchical clustering to merge clusters in c to bigger groups until the maximum number of clusters (=maximum number of teaching paths) is reached.

Table I. Personality mining using data mining techniques - major contributions

the perspective of the course material, the average proportion of the recommended course material that is understood by the learners is 0.837. He also observes that the average difficulty of the course materials that are recommended by their system is 1.815 from a learner's perspective, which indicates that most learners agree that the course material was moderately difficult.

2.2.2 *Database programming.* Personality differences between learners and the emotional state they are in, play an important role in defining successful learning strategies, particularly in e-learning courses, where there is no direct contact between the teacher and learner. According to Fatahi et al. [Fatahi et al. 2009], although much research has been done on predicting emotions in learners, personality as an independent parameter has not been explored. The authors propose an algorithm that consists of six steps : 1. Identify the learner's personality using Myers-Briggs Type Indicator (MBTI). For simplicity, they use only two of the four functions proposed in the MBTI model (Sensitivity/Intuition (S/N) and Extroversion / Introversion (E/I)) and generated four different personalities : ES, EN, IS, IN. 2. Choose a learning style : individual, competitive and collaborative. 3. Choose a virtual classmate agent(VCA) for the learner (this agent is used if the learning style is competitive or collaborative). The authors observe from their previous research [Ghasem-Aghaee et al. 2008] that the most effective VCA that can improve the learner's capability is one that exhibits a completely opposite behavior than the learner. For example, if the learner is an introvert, the VCA is an extrovert so that it can alert the learner to speed up in finishing a task (introverts take time to solve a problem and give answers, whereas an extrovert acts and answers without thinking much). 4. Present the course material to the learner. 5. Evaluate the emotions exhibited by the learner when learning the course material. The authors did not give details on how to evaluate emotions.6. Update the learning style based on the actions in steps $1 - 5$. For example, let a learner L be an introvert and pick a learning style of 'independent'. When an event happens (e.g. L is presented with a course quiz), L's emotion is evaluated as 'Satisfied' to a medium level and his answer to the quiz is correct, a different learning style such as 'competitive' is recommended for him (instead of working independently). The authors used data of 30 students in a series of English language exercises in an e-learning environment and used a programming environment with Visual C#, .Net and SQL Server. An existing Microsoft agent Merlin was used as the virtual classmate agent. They claim that the presence of a virtual classmate leads to an improvement in the overall learning experience of the learner and makes the learning environment more enjoyable.

2.2.3 *Statistical.* In a recent paper by Arockiam et al. [Arockiam et al. 2012], the authors propose a statistical learning model that studies the relationship between the learners cognitive skills such as retention and recollection and his / her personality traits such as extraversion, neuroticism and psychotocisim. The authors also present a review on how user interface design (UID) of the course webpage impacts the learners' retention and recall on course resources. The design they propose consists of three steps : 1. Developing a personality test to find traits using Eysenck Personality Inventory (EPI) method. Traits were categorized as Exraver-

sion, Neuroticism and Psychoticism.2. Developing a Retention and Recollection (R and R) questionnaire based on web-page content such as buttons, font and color scheme and 3. Statistical analysis of the results of steps 1 and 2. They conducted experiments involving 50 MCA students from a University in India. They started with 150 students and filtered it to 80 students by picking those who achieved more than 60% in their Undergraduate and Postgraduate degree programs, were of the same age group and were willing to participate in the survey. The students were subjected to three phases twice in a gap of 24 hours : Phase I: Personality test for 30 minutes, Phase II: Course web page browsing for 10 minutes, Phase III: Retention and Recollection test. Arockiam et al. [Arockiam et al. 2012] claim that extraverts have the best recollection and retention in e-learning, followed by those with a psychotistic personality and also that students who perform better immediately perform equally well even after 24 hours, irrespective of the web-page elements.

2.2.4 *Summary :* . Table II has the summary.

Year	Author(s)	Title of the paper	Technique	Major contribution
2005	Chen, C.	Personalized e-learning system using Item Response theory	Item Response Theory	An algorithm based on item response theory (IRT) which considers difficulty of course materials and learners' abilities to provide an appropriate learning path. A course material is modeled by the item characteristic function proposed by Rasch with a single difficulty parameter. The course material difficulty is adjusted according to a collaborative voting approach .
2009	Fatahi, S.; Kazemifard M. and Ghasem-Aghaee, N.	Design and Implementation of an E-Learning Model by Considering Learner's Personality and Emotions.	Database Programming	An algorithm that uses Visul C# and SQL Server to study learner's personlaity and emotions exhibited when learning the course material, and its relation to learner's satisfaction of the learning environment.
2012	Arockiam, L.; Charles Selvaraj, J. and Amala Devi, S.	An Impact of Personality Traits and Recollection and Retention Skill in E-Learning	Statistical	Statistical analysis of personality traits and user interface features to study their impact on student's retention and recall on course resources.

Table II. Personality Mining using other techniques - major contributions

3. CONCLUDING COMMENTS

This survey concerns a challenge that teachers teaching e-learning courses face when connecting with their students. Students succeed in courses that are taught with a certain learning strategy that suits their personality. In a classroom teaching model, successful teachers monitor the students in their class and adjust the pace of their teaching accordingly. However, in web-based e-learning courses, where there is no

direct interaction of students with teachers, it is a challenging task to automate this process of monitoring student personalities and tweaking learning strategies accordingly. Twenty three papers related to this topic of survey were identified and ten of them were reviewed in detail. In doing this, the following observations were made:

(1) Although distance courses seem to be very popular in the United States and Canada (e.g. according to a website of 'The Canadian Information Center for International Credentials (CICIC)', Ontario's colleges and universities offer around 10,000 courses and over 800 online programs and distance learning opportunities), much of the research done in this area is by a research group from a University in China. This research group seems to be highly active in the area of personality mining, especially using the association rule mining technique. In fact, 11 of the 21 papers in the bibliography are either from China or Taiwan.

(2) All the papers reviewed use a learner / student model but there seems to be no consensus on the number and representation of attributes (e.g. whether attributes should be represented numerically or as boolean).

(3) Most of the papers reviewed use advanced data mining techniques to find a correlation between student's personality, behavior and the course learning outcome. However, a very recent paper [Arockiam 2012] uses the traditional statistical technique to find the correlation between student's personality and understanding of the course concepts.

(4) The most commonly used data mining technique in personality mining is association rule mining. Some researchers have also used k-means and hierarchical clustering methods.

(5) Although research on personality mining of students enrolled in e-learning courses started almost a decade ago, most researchers have used the existing data mining algorithms such as apriori for association rule mining and k-means for clustering, instead of proposing new algorithms or suggesting modifications to these existing ones to cater to the scope and dimension of attributes specific to e-learning.

(6) It appears that many papers were published around the same time and were not aware of each other's work as demonstrated in Figure 1.

4. ACKNOWLEDGEMENT

I would like to thank Dr. Richard Frost for his unconditional support and guidance throughout this survey. I convey my deep regards to my supervisor Dr. C.Ezeife for encouraging me and directing me towards my goal. I would also like to thank my family and friends for providing me immense strength and support in completing this survey report.

5. ANNOTATIONS

5.1 Arockiam et al. 2012

AROCKIAM, L., CHARLES SELVARAJ, J., AND AMALA DEVI, S. 2012. An impact of personality traits and recollection & retention skill in e-learning. *Global Trends in Information Systems and Software Applications*, 357–366.

The problem which the researchers/authors addressed

Higher education has shown a drift from the classroom teaching model to web-based distance courses, also known as e-learning courses. Since there is no face-to-face interaction with the teacher in an e-learning environment, it is challenging for teachers to guage if the course structure suits each learners' capability and motivation.

Previous work by others referred to by the authors

The authors refer to previous work done by Ghasem [2008] and Du [2005].

Shortcomings of previous work

The authors did not list the shortcomings of any of the related work. But they did make a statement about Gasem-Aghaaee et al. [2008] that they mentioned the use of parameters such as culture as future work. Perhaps this indicates that the authors did not consider culture in this paper when studying the personality of e-learners.

The new algorithm.

The authors propose a learning model that studies the relationship between the learners cognitive skills such as retention and recollection and his / her persoanlity traits such as extraversion, neuroticism and psychotocisim. The authors also present a review on how user interface design (UID) of the course webpage impacts the learners' retention and recall on course resources. The design proposed consists of two steps :

1. Developing a personality test to find traits using Eysenck Personality Inventory (EPI) method. Traits were categorized as Exraversion, Neuroticism and Psychoticism.
2. Developing a Retention and Recollection (R and R) questionnaire based on web-page content such as buttons, font, color scheme etc.
3. Statistical analysis of the results of steps 1 and 2.

Experiments or analysis conducted

The authors conducted experiments involving 50 MCA students from a University in India. They started with 150 students and filtered it to 80 students by picking those who achieved more than 60% in their Undergraduate and Postgraduate degree programs, were of the same age group and were willing to participate in the survey. The students were subjected to three phases twice in a gap of 24 hours.

Phase I: Personality test for 30 minutes
Phase II: Course web page browsing for 10 minutes
Phase III: Retention and Recollection test

Personality Type	RandR - immediately
Extraversion	0.094239
Neuroticism	-0.19492
Psychoticism	-0.00521

Personality Type	RandR - after 24 hours
Extraversion	0.302963
Neuroticism	0.063758
Psychoticism	-0.21664

Table III. Arockiam et al. [2012] Table 2 Page 363 Arockiam et al. [2012] Table 3 Page 364

Results that the authors claim to have achieved

Claims made by the authors

The authors claim that :

1. Their contribution is a first step to explore a positive correlation between personality traits and recollection and retention skills in e-learning.
2. Extraverts have the best recollection and retention in e-learning, followed by those with a psychotistic personality.
3. Students who perform better immediately perform equally well even after 24 hours, irrespective of the web-page elements.

Citations to the paper by other researchers

There are no specific references to this paper by other researchers in this survey.

5.2 Fatahi, S. et al. 2009

FATAHI, S., KAZEMIFARD, M., AND GHASEM-AGHAEE, N. (2009). Design and Implementation of an E-Learning Model by Considering Learner's Personality and Emotions. , *Advances in Electrical Engineering and Computational Science, Springer Netherlands* , 423–434.

The problem which the researchers/authors addressed

Higher education has shown a drift from the classroom teaching model to web-based distance courses, also known as e-learning courses. Since there is no face-to-face interaction with the teacher in an e-learning environment, therefore, this results in the problem of automating the process of identifying a student's personality and learning style to detect if the pace of the course and its structure suits each learner's personality and learning style.

Previous work by others referred to by the authors

The authors refer to previous work done by Ghasem [2008] and Du [2005].

Shortcomings of previous work

According to the authors, although much research has been done on predicting emotions in learners, personality as an independent parameter has not been explored.

The new algorithm.

The authors propose an algorithm that consists of six steps :

1. Identify the learner's personality using Myers-Briggs Type Indicator (MBTI). For simplicity, they use only two of the four functions proposed in the MBTI model

(Sensitivity/Intuition (S/N) and Extroversion / Introversion (E/I)) and generated four different personalities : ES, EN, IS, IN.

2. Choose a learning style : individual, competitive and collaborative.

3. Choose a virtual classmate agent(VCA) for the learner (this agent is used if the learning style is competitive or collaborative). The authors observe from their previous research [Ghasem et al. 2008] that the most effective VCA that can improve the learner's capability is one that exhibits a completely opposite behavior than the learner. For example, if the learner is an introvert, the VCA is an extrovert so that it can alert the learner to speed up in finishing a task (introverts take time to solve a problem and give answers, whereas an extrovert acts and answers without thinking much).

4. Present the course material to the learner.

5. Evaluate the emotions exhibited by the learner when learning the course material. The authors did not give details on how to evaluate emotions.

6. Update the learning style based on the actions in steps $1 - 5$. For example, let a learner L be an introvert and pick a learning style of 'independent'. When an event happens (e.g. L is presented with a course quiz), L's emotion is evaluated as 'Satisfied' to a medium level and his answer to the quiz is correct, a different learning style such as 'competitive' is recommended for him (instead of working independently).

Experiments or analysis conducted

The authors conducted experiments with 30 students in a series of English language exercises in an e-learning environment. Visual C#, .Net and SQL Server were used for the experiments. An existing Microsoft agent Merlin was used as the virtual classmate agent.

Results that the authors claim to have achieved

The authors present two results : learner's satisfaction of the learning environment and the effect of the presence of a virtual classmate on learners learning style and capability.

Claims made by the authors

The authors claim that the presence of a virtual classmate leads to an improvement in the overall learning experience of the learner and makes the learning environment more enjoyable.

Citations to the paper by other researchers

Google scholar lists 6 citations that refere to this paper.

5.3 Zakrzewska, D. 2009

ZAKRZEWSKA, D. 2009. Cluster analysis in personalized e-learning systems. *Intelligent Systems for Knowledge Management (Springer Berlin Heidelberg)* , 252, 229–250.

The problem which the researchers/authors addressed

Higher education has shown a drift from classroom teaching model to web-based distance courses, also known as e-learning courses. Since there is no face-to-face interaction with the teacher in an e-learning environment, therefore, this results in the problem of automating the process of identifying a student's personality and learning style to detect if the teaching strategy and resources (such as website layout) are well-adjusted to the student's learning style.

Previous work by others referred to by the authors

The authors do not refer to any previous work.

Shortcomings of previous work

Although the authors have not stated the shortcoming of any particular author or previously done research, they make two general observations :

1. Although a lot of research is being done on adaptive intelligent educational systems, there is still no agreement on what features such as student charactertics should be used or can be used in creating personlaized learning style in e-learning systems.
2. In many existing systems such as Arthur and 3DE, students were provided with different teaching strategies but the groups of students were made statically.

The new algorithm.

The authors propose a two-phase algorithm in which students are clustered , not only based on their learning styles but also on their usability preferences such as website layout and coloring scheme. In the first phase, groups of students, who are very similar and those who are outliers are formed using k-means clustering algorithm. The attributes used are student learning styles defined by Felder and Silverman (active/reflective, sensing/intuitive, visual/verbal/auditory and sequential/global). In the second phase, such groups are merged into bigger groups using hierarchical clustering algorithm until the maximum number of clusters is reached (the maximum number of clusters denotes the maximum number of teaching paths). Clusters are formed based on the threshold values given as input to the hierarchical clustering algorithm (such as a value t such that the similarity between the ith student's attributes and the cluster centroid $>$ t , maximum number of clusters etc.). Any 1-element clusters indicate outliers such as outstanding students.

Experiments or analysis conducted

The authors conducted experiments with two student data sets using two different models. The first model used student learning styles as attributes and the second model used students' color choices given after examiming the website layout, in addition to the learning styles. A student dataset of 71 computer science students (from degree programs such as undergraduate, Masters, part-time and evening courses) was collected from a web-based collaboration system called Moodle. The second dataset of 73 students was artificially generated for comparison purposes. A machine learning tool called WEKA was also used to apply k-means and hierarchical clustering algorithm on these datasets to compare their performances with

that of the proposed two-phase algorithm .

Results that the authors claim to have achieved

The authors claim that their two-phase algorithm generates clusters that are very close in accuracy to the clusters generated by hierarchical clustering, although k-means divided the students into almost equal clusters. Clusters from the learning styles perspectives showed some unexpected results since students with different learning styles were put together in a single cluster (e.g. sequential and global in the same cluster). Detailed results are shown in sections 6.1 and 6.2 [Zakrzewska, D. 2009] on pages 242-246.

Claims made by the authors

The authors claim that

1. their modified clustering algorithm yields better quality of clusters than the traditional algorithms and it does not to provide the exact number of clusters in advance.
2. the best quality of clusters are obtained when both learning styles and website layout preferences such as color are used in the first phase of the algorithm.
3. their algorithm successfully detects all outstanding students (outliers).

Citations to the paper by other researchers

Google scholar lists 8 citations that refered to this paper.

5.4 Huang et al. 2008

HUANG, K., LI F., ZHAO, M., WANG, F., XU, X. 2008. Design and Implement On E-learning behavior Mine System. *Second International Conference on Genetic and Evolutionary Computing*, 406 – 409 .

The problem which the researchers/authors addressed

Although the authors do not state explicitly, it appears as though the problem addressed is to personalize learning materials in e-learning courses using personality and behavior of learners. Since there is no face-to-face interaction with the teacher in an e-learning environment, therefore, this results in the problem of a teacher not being able to study a student's personality and learning style, so that the course learning materials can be appropriately adjusted.

Previous work by others referred to by the authors

The authors do not refer to any previous work from my bibliography.

Shortcomings of previous work

The authors do not state any shortcomings of previous work.

The new algorithm.

The authors propose an e-learning behavior model which they define as multi-dimensional and multi-hierarchical. The dimensions include static (survey using questionnaires) and dynamic (using logged web data) ways of collecting personality information. The logged data has behavior information on how learners collect,

process, publish and use materials on the e-learning website. The levels that the authors propose are primary (e.g. browsing the webpage), secondary (e.g. answering a quiz) and senior (e.g. collaborating with peers). The steps in the proposed algorithm are :

1. Collect learners' data (using questionnaires and logs).
2. Use a combination of statistical and mining methods to categorize the data collected into behavior, personality and e-learning resource characteristics.
3. Evaluate the characteristics found in step 2 as formative, immediate or stage evaluation. No details on how to evaluate are given.
4. Provide individualized guidance to the learner (done by a module called Intelligent feedback module).

Experiments or analysis conducted

Huang et al. [2008] did not describe any experiments.

Results that the authors claim to have achieved

Huang et al. [2008] did not present any results.

Claims made by the authors

The authors claim that the different dimensions and levels of hierarchy that they propose to impart personalized learning is effective.

Citations to the paper by other researchers

There are no specific references to this paper by other researchers in this survey.

5.5 Zheng et al. 2008

ZHENG, Q., WU, X., LI, H. 2008. A rough set based approach to find learners' key personality attributes in an e-learning environment. *International Journal of Web-Based Learning and Teaching Technologies*, 3, 2, 29–56.

The problem which the researchers/authors addressed

Higher education has shown a drift from classroom teaching model to web-based distance courses, also known as e-learning courses and therefore, students succeed in courses that are taught with a certain learning strategy that suits their personality. Since there is no direct interaction of students with teachers, therefore, this results in the problem of a teacher not being able to study a student's personality and learning style so that they can adjust the course learning materials appropriately, unless this process is automated.

Previous work by others referred to by the authors

The authors do not refer to any previous work from the bibliography.

Shortcomings of previous work

According to the authors, most of the previous studies on personality characterstics and learning strategies of learners are done in psychology where methods such as correlation and regression analysis and discriminator functions are used. These methods are subjective and time-consuming as they use human expert knowledge

in their analysis to find the key attributes of learners that descibe their personality and learning styles.

The new algorithm.

The algorithm that the authors propose has four steps:

1. Collect the learner's personality characterstics and store them in an information system.
2. Use an algorithm based on rough-set theory to reduce the typically large set of personality attributes to a core set of key attributes. Rough-set theory is a mathematical model that works best for situations that deal with uncertain, imprecise and incomplete data.
3. Define a function that assigns weights to these key attributes.
4. Use association rule mining to generate rules such as "students with personality described by attributes (a1,a2,a3,a9,a10,a22) fit best with the learning strategy known as metacognitive strategy".

Experiments or analysis conducted

The authors collected a total of 28 attributes and 5 learning strategies of 157 students registered in a "Personalized English Learning" course of Xi'an Jiaotong University, China. They then applied the algorithm listed above to find the set of personality attributes that cater to each learning strategy. For example, learning strategy mcs(metacognitive strategy) caters to the personality set {Emotional Stability, Dominance, Social Boldness, Apprehension, Self-relaince, Experimental, Individuality}.

Results that the authors claim to have achieved

The authors claim that the rough-set algorithm helped them to reduce the attributes in the information system by at least 50%. They also claim to have generated meaningful rules that correlate the 5 learning strategies with groups of personality traits. Detailed results are given in tables 3 - 8 on pages 44-48 [Zheng et al. 2008].

Claims made by the authors

The authors claim that

1. Their algorithm reduces the volume of the information system by at least 50% in four out of the five learning strategies.
2. Learning style attribute "visual|auditory|experimental" has the deepest influence on the meta-cognitive strategy.
3. Personality trait "sensitivity" contributes most to all the 5 learning strategies.

Citations to the paper by other researchers

Google scholar lists 1 citation that refered to this paper.

5.6 Huang et al. 2007

HUANG, J., ZHU, A., AND LUO, Q. 2007. Personality mining method in web based education system using data mining. 2007. *International IEEE Conference on Grey Systems and Intelligent Services, 2007*, 155–158.

The problem which the researchers/authors addressed

The authors mention the need for personalizing web-based education to address the problem of automatically detecting student's personality and behavior while accessing the course materials, so that the course pace and structure can be well adjusted to them and thereby improve the overall learning experience.

Previous work by others referred to by the authors

The authors refer to none.

Shortcomings of previous work

The authors refer to none.

The new algorithm.

The authors propose an improved apriori algorithm that generates non-redundant association rules between behavior and personality. The algorithm works in five steps:

1. Learner's personality information is collected (there is no mention of how it is collected). Learner's behavior information is collected using the e-learning courseware website's logged data. Then the two are merged and stored in a database.
2. A datacube is generated with learning behavior, personality characters and time as the three dimensions.
3. Frequent k-itemsets are generated for each dimension using Apriori algorithm. If a dimension doesn't have any frequent k-itemsets, it implies that the dimension level is very low and therefore, needs to be rolled up to the next higher level. If a dimension has all its k-itemsets as frequent, it implies that the dimension level is high and therefore, needs to be drilled down to a lower level. This process is repeated until a desired level is established.
4. Non-redundant association rules are now generated using the following formula for each rule of the form s => (l-s) : support_count(l) / support_count(s) >= min_confidence.
5. A correlation between the itemsets in each rule generated in step 4 is evaluated. Only those that meet a certain threshold are retained.

Experiments or analysis conducted

The authors conducted experiments with 360 learners (300 used as data source and 60 used as test data).

Results that the authors claim to have achieved

The authors claim that 9766 association rules were generated in total. Using a correlation factor < 166.7, 6794 rules were retained.

Claims made by the authors

The authors claim that their algorithm has two advantages : shorter execution time attributed to the removal of redundant rules and high value of precision of the rules generated.

Citations to the paper by other researchers

Google scholar lists 1 citation that refer to this paper.

5.7 Tian et al. 2007

TIAN, F., QINGHUA, Z., ZHIYONG, G., DU, J., Li, R. 2007. Personalized Learning Strategies in an intelligent e-Learning Environment *11th International IEEE Conference on Computer Supported Cooperative Work in Design*, 973–978.

The problem which the researchers/authors addressed

Although the authors did not state explicitly, it appears as though the problem addressed is the challenge of recommending individualized appropriate learning strategies to learners in a web-based environment to improve the course's learning outcomes.

Previous work by others referred to by the authors

The authors refer to none.

Shortcomings of previous work

The authors do not list shortcomings of any previous research done - although, they do mention that the current research lacks well-defined and effective personality analysis methods.

The new algorithm.

The authors propose an algorithm that has four steps :

1. Build a learner model: A learner model stores personal learner information such as name and address, information such as preferences and relationships using the standard PAPI model, learner's personality characteristics such as warmth and emotional stability using Cattel's 16 personality factors, learning strategies such as mental (active/reflective) , behavioral (interpersonal) and self-regulatory(motivation) using a learning strategy survey questionnaire and learning behavior such as the time spent on accessing courseware materials using logged website data.
2. Mine the relationship between personality and behaviour using associative rule mining.
3. Mine the relationship between personality and learning strategies using rough set theory.
4. Match the learning strategies with the learner's personality and behavior to generate personalized learning strategies (the authors give no details on how this matching is performed).

Experiments or analysis conducted

The authors conducted experiments with 1013 students in an English language course in a University in China.

Results that the authors claim to have achieved

The authors identified how their experiments helped recommending personal strategy to learners. In particular, they mention one student who had a negative person-

ality, lacked self-confidence, had a strong inferiority complex and did not perform well in the course was seen to enhance his ability of answering questions correctly from 66.7% to 87.5% using auto-generated personalized learning strategies such as "participate more in group discussions".

Claims made by the authors

The authors claim that their idea of mining relationships between personality and behavior and between personality and learning strategies was novel and effective.

Citations to the paper by other researchers

Google scholar lists 1 citation that refer to this paper.

5.8 Jin et al. 2006

JIN, D., QINGHUA, Z., JIAO, D., AND ZHIYONG, G. 2006. A method for learner grouping based on personality clustering. *10th International IEEE Conference on Computer Supported Cooperative Work in Design*, 1–6.

The problem which the researchers/authors addressed

Although the authors did not state the problem addressed explicitly, it appears that the problem targeted is in context of analyzing student personalities and recommending appropriate learning strategies for students with different personalities in an e-learning environment to improve the course's learning outcomes and the student's overall learning experience.

Previous work by others referred to by the authors

The authors refer to work done on personality mining in e-learning by Du, J. et al. [2005].

Shortcomings of previous work

The authors do not list shortcomings of any previous work done in the area.

The new algorithm.

The authors first present an analysis of various existing clustering methods to validate the feasibility of using k-means clustering algorithm as the most appropriate method for grouping learners. Then they build a learner model that consists of attributes that represent students' personality (e.g. warmth), motivation (e.g. curiosity), study style (e.g. visual/hands-on, serial/random), study concept (e.g. self-management), strategy (e.g. memorize/ cognize) and behavior on courseware learning (e.g. average time spent on each login). Finally, they apply k-means algorithm to group students with same personality into one cluster so that a learning strategy appropriate for that group can be proposed.

Experiments or analysis conducted

The authors collected personality data of 500 students from Xi'an Jiaotong University using Cattell's 16 Personality Factors questionnaire and filtered that to a dataset with 260 student records (S), eliminating those that took more than 55 minutes or less than 15 minutes to complete the questionnaire. Then they applied

k-means algorithm to cluster S, with the number of clusters ranging between 3 and 9 and the distance function as Euclidean. The authors used two different methods to split their data into train and test datasets : one method split them randomly, whereas the other method used a 50-50 split. The authors also conducted experiments using another clustering method called BIRCH to compare and validate the performance and accuracy of the suggested k-means algorithm.

Results that the authors claim to have achieved

The authors highlighted the results of clusters obtained with cluster size of 4,5,6 and 8 with a 50-50 split model for training and test data. The authors claim to have the best results for cluster sizes 4 and 5.

Claims made by the authors

The authors claim that the performace of k-means was most accurate when cluster size was given as 4 and 5, with cluster size of 5 outperforming all other cases. The authors also claim that they propose a mathematical model to build their learner model but there is no evidence of such a mathematical model in their paper.

Citations to the paper by other researchers

Google scholar lists 1 citation that refer to this paper.

5.9 Chen 2005

CHEN, C. 2005. Personalized e-learning system using Item Response Theory *Computers and Education*, 44, 237–255.

The problem which the researchers/authors addressed

Students succeed in courses that are taught with a certain learning strategy that suits their personality. In a classroom teaching model, successful teachers monitor the students in their class and adjust the pace of their teaching accordingly. However, higher education has shown a drift from classroom teaching model to web-based distance courses, also known as e-learning courses, where there is no direct interaction of students with teachers. Therefore, the problem is to automate the process of studying student's personality and adjusting the course materials appropriately.

Previous work by others referred to by the authors

The authors do not refer to any previous work from my bibliography.

Shortcomings of previous work

The authors did not list the short-comings of any of the related work.

The new idea, algorithm, architecture, protocol, etc.

The authors propose an algorithm based on item response theory (IRT) which considers difficulty of course materials and learners' abilities to provide an appropriate learning path. The algorithm uses three different agents to accomplish this : interface agent (front-end), feedback and course recommendation agent (back-end). When a learner logins to the system for the first time, the interface agent collects

learner's personal information and a set of key words that help locate course materials (called units) that the learner is interested in. A user has to register in the system, if he/she desires a personalized feedback. All this information is used to create the learner's profile. For first-time learners, the difficulty of the course unit chosen is set to medium. After the learner uses the course material, he/she is asked to answer two questions : How did you find the difficulty of the course material (response could be very hard, hard, moderate, easy, very easy) and Do you understand the content of the course material ? (response could be Yes or No). Then, the feedback agent collects these responses and re-evaluates the learner's abilities for that course unit using a maximum likelihood estimation function. The learner's profile is updated with these new abilities and the difficulty parameter of that course unit is ranked accordingly in the course database. The course recommendation agent is also informed of the learner's new abilities and therefore, it provides new recommendation for the learner. This process is repeated for other course units until the learner logs out of the system. A course material is modeled by the item characteristic function proposed by Rasch with a single difficulty parameter. The course material difficulty is adjusted (and ranked) according to a collaborative voting approach (formula are provided in the paper).

Experiments or analysis conducted

The authors conducted experiments using a course called 'Neural networks' that had three course units and a 58 course materials. Experiments were conducted on the course unit called 'Perceptron' because it had a sufficiently high number of course materials in it (35) with various levels of difficulty. All the 210 learners logged into the system were Masters students.

Results that the authors claim to have achieved

The authors claim that from a learner's perspective, results show that the average degree of understanding of the recommended course materials is 0.825 and from the perspective of the course material, the average proportion of the recommended course material that is understood by the learners is 0.837. They also observe that the average difficulty of the course materials that are recommended by their system is 1.815 from a learner's perspective, which indicates that most learners agree that the course material was moderately difficult.

Claims made by the authors

The authors claim that the learners' comprehension of the recommended course material is high (indicated by results that the average degree of understanding of the recommended course materials is 0.825) and that most learners agree that the course material is moderately difficult (indicated by results that average difficulty of the course materials that are recommended by their system is 1.815 which is closer to two). The authors also claim that their proposed system can not only precisely provide personalized course material recommendation , given his / her learning ability, it can accelerate the learning efficiency and effectiveness.

Citations to the paper by other researchers

Google scholar lists 214 papers that refer to this paper.

5.10 Du et al. 2005

DU, J., ZHENG, Q., LI, H., AND YUAN, W. 2005. The research of mining association rules between personality and behavior of learner under web-based learning environment. *Advances in Web-Based Learning ICWL*, 406–417.

The problem which the researchers/authors addressed

Students succeed in courses that are taught with a certain learning strategy that suits their personality. In a classroom teaching model, successful teachers monitor the students in their class and adjust the pace of their teaching accordingly. However, higher education has shown a drift from classroom teaching model to web-based distance courses, also known as e-learning courses, where there is no direct interaction of students with teachers. Therefore, the problem is to automate the process of studying student's personality and adjusting the course materials appropriately.

Previous work by others referred to by the authors

None in the list. (There was one article that I had initially in my list of exact topics but it was in Chinese).

Shortcomings of previous work

The authors did not list the short-comings of any of the related work. In fact, there is no related work section.

The new algorithm.

The authors propose a Learner model that consists of a personality model (PM) and a behavior model (BM). Personality model represents the learners personality, motivation, style and strategy used. Each of these PM sub-categories are stored quantitatively as attributes e.g., personality is described by Cattell's 16 personality factors. The BM is represented by six different aspects of the sequence of actions that a learner takes while using the web page : behavior on courseware learning, tests / assignments, postings and interaction among students and teachers through the use of discussion boards, chat rooms etc. Each attribute of PM and BM is mapped to an integer value of 1,2 or 3 representing high, middle and low. This transformed data is then mined using the Apriori algorithm to analyse the relation between personality and behavior. The authors propose an improved Apriori algorithm that generates association rules of the form Behavior => Personality. Differences between traditional and proposed apriori algorithm are summarized using an example given in tables 3 and 4 on page 414 of section 3.2 [Du et al. 2005].

Experiments or analysis conducted

The authors developed a 'Personalized English Learning System' for 324 students of Xi'an Jiaotong University and collected their attributes such as the 16 Cattell's personality factors and 146,000 log records with attributes for the behavior model. Then they applied the proposed modified-Apriori algorithm to generate association rules between behavior and personality.

Table 6. Part of relationship between BM and PM

BM / Personality	B1	B2	B3	B4	B5	B6
A(Warmth)	PC			PC		PC
E(Dominance)		PC	PC	NC	PC	
G(Rule Consciousness)	PC					
L(Vigilance)	PC		NC		PC	
Q2(Self-Reliance)		PC	PC			
.....						

Fig. 2. Du et al.[2005] Table 6 Page 417

Results that the authors claim to have achieved

The authors state that some personality attributes show a positive correlation with the behavior attributes whereas others show a negative correlation. An example is shown in the table below. Here, BM implies behavior nodel and B1, B2 .. are behaviors extracted from logged data (e.g. B1 is time spent on courseware). PC implies positive correlation, whereas NC implies negative correlation.

Claims made by the authors

The authors claim that their algorithm is an improvement over the Apriori algorithm because it avoids the generation of redundant rules substantially, thereby reduces the complexity of the algorithm considerably.

Citations to the paper by other researchers

Google scholar lists 12 citations that refer to this paper.

6. REFERENCES

REFERENCES

Arockiam, L., Charles Selvaraj, J., and Amala Devi, S. 2012. An impact of personality traits and recollection and retention skill in e-learning. *Global Trends in Information Systems and Software Applications*, 357–366.

Bodewitz, M. 2004. Detecting the student personality. M.S. thesis, Faculty of Computer Science, University of Twente.

Chen, C. 2005. Personalized e-learning system using item response theory. *Computers and Education 44*, 237–255.

Chen, C. M., Hsieh, Y. L., and Hsu, S. H. 2007. Mining learner profile utilizing association rule for web-based learning diagnosis. *Expert Systems with Applications. 33,1*, 6–22.

Du, J., Zheng, Q., Li, H., and Yuan, W. 2005. The research of mining association rules between personality and behavior of learner under web-based learning environment. *Advances in Web-Based Learning–ICWL 2005 (Springer).*, 406–417.

El Bachari, E., Abelwahed, E. H., and Adnani., M. E. 2010. Design of an adaptive e-learning model based on learners personality. *Ubiquitous Computing and Communication Journal 5,3*, 27–36.

Faliagka, E., Tsakalidis, A., and Tzimas, G. 2012. An integrated e-recruitment system for automated personality mining and applicant ranking. *Internet Research 22,5*, 551 – 568.

Fatahi, S., Kazemifard, M., and Ghasem-Aghaee, N. 2009. Design and implementation of an e-learning model by considering learner's personality and emotions. *Advances in Electrical Engineering and Computational Science, Netherlands, Springer. 39*, 423–434.

Ghasem-Aghaee, N., Fatahi, S., and Ören, T. 2008. Agents with personality and emotional filters for an e-learning environment. In *Proceedings of the 2008 Spring simulation multiconference,San Diego, CA, USA*. Society for Computer Simulation International, 1–5.

Graf, S. and Kinshuk, K. 2006. Considering learning styles in learning management systems: Investigating the behavior of students in an online course. In First IEEE International Workshop on Semantic Media Adaptation and Personalization. *First International Workshop on Semantic Media Adaptation and Personalization.*, 25–30.

Huang, J., Zhu, A., and Luo, Q. 2007. Personality mining method in web based education system using data mining. In *International Conference on Grey Systems and Intelligent Services*. IEEE, 155–158.

Huang, K., Li, F., Zhao, M., Wang, F., and Xu, X. 2008. Design and implement on e-learning behavior mine system. In *Second International Conference on Genetic and Evolutionary Computing*. IEEE, 406–409.

Jin, D., Qinghua, Z., Jiao, D., and Zhiyong, G. 2006. A method for learner grouping based on personality clustering. In *10th International Conference on Computer Supported Cooperative Work in Design*. IEEE, 1–6.

Kim, E. B. and Schniederjans, M. J. 2004. The role of personality in web-based distance education courses. *Communications of the ACM. 47,3*, 95–98.

Markowska-Kaczmar, U., Kwasnicka, H., and Paradowski, M. 2010. Intelligent techniques in personalization of learning in e-learning systems. In *Computational Intelligence for Technology Enhanced Learning,*. 1–23.

Ozpolat, E. and Akar, G. B. 2009. Automatic detection of learning styles for an e-learning system. *Computers & Education 53,2*, 355–367.

Tian, F., Zheng, Z., Gong, Z., Du, J., and Li, R. 2007. Personalized learning strategies in an intelligent e-learning environment. In *11th IEEE International Conference on Computer Supported Cooperative Work in Design*. 973–978.

Xiaoyan, W. and Ping, D. 2011. A study of personalized learning system based on xml and web mining. *Cross Strait Quad-Regional Radio Science and Wireless Technology Conference (CSQRWC), IEEE. 2*, 1200–1203.

Yang, Q. and Wang, J. 2008. Research on learners personality mining based on improved decision tree algorithm. In *Second International Conference on Genetic and Evolutionary Computing*. IEEE, 398–401.

Zakrzewska, D. 2009. Cluster analysis in personalized e-learning systems. *Intelligent Systems for Knowledge Management (Springer Berlin Heidelberg) 252*, 229–250.

Zakrzewska, D. 2010. Student groups modeling by integrating cluster representation and association rules mining. *SOFSEM 2010: Theory and Practice of Computer Science, Springer Berlin Heidelberg.*, 743–754.

Zhang, K., Cui, L., Wang, H., and Sui, Q. 2007. An improvement of matrix-based clustering method for grouping learners in e-learning. In *11th IEEE international conference on Computer Supported Cooperative Work in Design*. 1010–1015.

Zheng, Q., Wu, X., and Li, H. 2008. A rough set based approach to find learners key personality attributes in an e-learning environment. *International Journal of Web-Based Learning and Teaching Technologies. 3,2*, 29–56.

	Authors	Year	Title	Papers cited by this paper	Papers citing this paper
1	AROCKIAM, L., CHARLES SELVARAJ, J., AND AMALA	2012	An impact of personality traits and recollection & retention skill in e-learning	Ghasem [2008], Du[2005]	None
2	FATAHI, S., KAZEMIFARD, M., & GHASEM-AGHAEE, N.	2009	Design and Implementation of an E-Learning Model by Considering Learner's Personality and Emotions.	Ghasem [2008], Du[2005]	El Bachary [2010]
3	ZAKRZEWSKA, D	2009	Cluster analysis in personalized e-learning systems.	None	Zakrzewska [2010]
4	HUANG, K., LI F., ZHAO, M., WANG, F., XU,	2008	Design and Implement On E-learning behavior Mine System.	None	None
5	ZHENG, Q., WU, X., LI, H.	2008	A rough set based approach to find learners' key personality attributes in an e-learning environment.	None	None
6	HUANG, J., ZHU, A., AND LUO, Q.	2007	Personality mining method in web based education system using data mining.	None	None
7	TIAN, F., QINGHUA, Z., ZHIYONG, G.,	2007	Personalized Learning Strategies in an intelligent e-Learning Environment	None	Ozpolat [2009]
8	JIN, D., QINGHUA, Z., JIAO, D., AND	2006	A method for learner grouping based on personality clustering.	Du [2005]	None
9	CHEN, C.	2005	Personalized e-learning system using Item Response Theory	Chen [2007]	
10	DU, J., ZHENG, Q., LI, H., AND YUAN, W. 2005.	2005	The research of mining association rules between personality and behavior of learner under web-based learning environment.	None	Ghasem [2008], Fatahi [2009], Jin [2006], Arockiam [2012], Zhang [2007]

Fig. 1. Papers reviewed