ISSN 2730-2601
RICE Journal of Creative Entrepreneurship and Management, Vol. 1, No. 3, pp. 62-73,
September-December 2020

doi: 10.14456/rjcm.2020.17
Received 3.12.20/ Revised 13.12.20/ Accepted 23.12.20

Collaborative Filtering for Project Document Assessment of Missing Score Estimation

Areerat Trongratsameethong
Ataya Tipprason
Department of Computer Science
Faculty of Science, Chiang Mai University
Chiang Mai, Thailand
Email: areerat.t@cmu.ac.th, ataya.eve@gmail.com

Abstract

Measuring project document quality is usually handled by an assessment committee. Sometimes the committee may not cover all assessment issues. This can result in discrepancy of project document quality and evaluation, and thus may affect recommendations given in the project document. The researchers therefore would like to propose *the collaborative filtering technique* to predict scores left unevaluated by the committee. The User-based Collaborative Filtering (UBCF) was used in this research. It consists of two main steps: similarity calculation and prediction. Experiments were conducted. There were 20 computer project documents assessed by the assessment committees and their assessment scores were collected. Some of the assessment scores were removed randomly and the UBCF was used to predict the removed scores. The experimental results showed that accuracy of the UBCF used in the experiments varied from 78 to 84 percent.

Keywords*: Collaborative filtering, project document assessment, recommendation system, assessment score prediction.*

1. Introduction

Project document recommendation is usually done for university students' senior projects. Recommendations on project documents rely on assessment scores given by committees concerned. Criteria for evaluating project documents include: content correctness, format correctness, understanding of the project contents, and completeness. There are many sub issues to be considered under particular criteria. The sub issues depend on contents put in each chapter of the project report. Typically, a computer project document is divided into six chapters: abstract, introduction, literature review, system analysis, system design, and conclusion and discussion. It has occurred quite often that some committee members may overlook or leave out some sub issues in their assessment, resulting in discrepancy of project document quality and evaluation. This certainly affects recommendations of project documents as feedback to the students on the project.

In coping with such discrepancy, the researchers are looking for a possible solution, particularly *Collaborative Filtering* as a technique used in recommendation systems. This technique has been used in many areas, such as online shopping, social media, education, to name but the major ones. The key concept of collaborative filtering is user or item similarity calculation (Dou, Yang & Deng, 2016; Xiaoyuan & Taghi, 2009). It consists of two main steps: similarity calculation and prediction. The first step is to find users or items that have high similarity with the user or item of interest as the neighbors. The characteristics or behaviors of the neighbors, such as ratings, comments, likes, and so on will later be evaluated and used for the prediction step.

In this paper, the researchers proposed the Collaborative Filtering technique to predict the missing scores left unevaluated by the assessment team member(s). The experiments were conducted and the accuracy of the used collaborative filtering was measured.

2. Literature Review

Collaborative filtering has been used in various fields of study, such as content similarity in genome data prediction (Ying-Wei, Xin & Yong-Ge, 2012), movie recommendation system (Purnomo & Endah, 2019), and course recommender (Chen, Liu & Shang, 2020). There are two main types of collaborative filtering: User-based Collaborative Filtering (UBCF) and Item-based Collaborative Filtering (IBCF). Both types belong to the memory-based collaborative filtering.

UBCF takes the users with the same rating item as a user set. Later, it predicts the user's rating to another item according to others' rating in the same user set. The key point of the UBCF algorithm is to find the neighbors with the greatest similarities with the interested user. After getting the similarity of the interested user to others, the similar neighbors of interested users according to the similarity are chosen. Finally, the rating of interested users to specific items is predicted using the rating history of similar neighbors and get the recommender results.

IBCF compares the similarity of different items, and later predicts the rating to a similar item of a user according to its current rating of items. Like the UBCF algorithm, the rating to different items of the same user needs to be collected.

In this research, the UBCF is selected to predict the missing scores by identifying the neighbors with the greatest similarities with the interested user.

3. Research Objective

The research objective of this study was to apply the collaborative filtering approach to estimate missing scores of computer project document assessment.

4. Research Methodology

Research methodology of this research work consisted of three main parts:

(1) Gathering computer project document assessment scores,

(2) Predicting computer project document assessment missing scores, and

(3) Measuring accuracy of prediction scores.

4.1 Gathering Computer Project Document Assessment Scores

In this research, computer project documents were assessed by an assessment team. Each document was assessed by five committee members: a project advisor, a course coordinator, and other three specialists. The assessment scores for each document were collected with an assessment form. There are four main assessment issues used on the form: content correctness, format correctness, understanding of the project contents, and completeness. The documents were assessed in six chapters: abstract, introduction of the project contents, literature review, system analysis, system design, followed by conclusion and discussion. Each chapter uses different sub issues for assessment depending on its characteristics as summarized below:

1) Abstract
    a. Content Correctness
        - Sub Issue 1: Conciseness
        - Sub Issue 2: Conveying important points
        - Sub Issue 3: No misleading information
        - Sub Issue 4: No other irrelevant concepts
        - Sub Issue 5: Containing clear research results
        - Sub Issue 6: Key words conveying the project's purpose
    b. Format Correctness
        - Sub Issue 1: Margin
        - Sub Issue 2: Font and size
        - Sub Issue 3: Line spacing
        - Sub Issue 4: Page number
        - Sub Issue 5: Length of contents
        - Sub Issue 6: Length of keywords
    c. Understanding of the project contents
        - Sub Issue 1: Giving an overview of the work
        - Sub Issue 2: No words or sentences considered misleading in the document
        - Sub Issue 3: Summarizing the needs of users
        - Sub Issue 4: Showing ability to understand, learn and expand knowledge from the project
        - Sub Issue 5: Readability of contents
    d. Completeness
        - Sub Issue 1: Showing a structure of introduction, content and summary
        - Sub Issue 2: Reporting objective, methodology, and conclusion
        - Sub Issue 3: Unity and coherence shown in the abstract
        - Sub Issue 4: Showing complete components: title, author, advisor, academic year, contents, and keywords
2) Introduction
    a. Content Correctness
        - Sub Issue 1: Preciseness and conciseness
        - Sub Issue 2: Showing evidence of information and reasons
        - Sub Issue 3: Showing the feasibility of the project
        - Sub Issue 4: Presenting theories and information in support of the project
        - Sub Issue 5: Objectives being consistent with the project title
        - Sub Issue 6: Objectives being measurable and evaluable
        - Sub Issue 7: Project contributions clearly stated
        - Sub Issue 8: Project scope and plan clearly defined

b. Format Correctness
   - Sub Issue 1: Paragraphs properly divided
   - Sub Issue 2: Numbering used in objectives
   - Sub Issue 3: Margins properly set
   - Sub Issue 4: Character font and size properly used
   - Sub Issue 5: Paragraph and numbering properly used
   - Sub Issue 6: Page numbers properly sequenced

c. Understanding
   - Sub Issue 1: Consistency in the project objectives
   - Sub Issue 2: Writing consistent in style with logical organization
   - Sub Issue 3: Clear description of the project

d. Completeness
   - Sub Issue 1: Containing background, reason and need of the project
   - Sub Issue 2: Presenting theoretical concepts, current situations, problems, and the importance of the problem
   - Sub Issue 3: Including the results with yields and benefits.

3) Literature Review

   a. Content Correctness
      - Sub Issue 1: Concepts well compiled and summarized
      - Sub Issue 2: Contents of the articles well re-paraphrased
      - Sub Issue 3: No copying of the referenced articles

   b. Format Correctness
      - Sub Issue 1: Paragraphs properly divided
      - Sub Issue 2: Numbering properly sequenced
      - Sub Issue 3: Margins properly set
      - Sub Issue 4: Character font and size properly used
      - Sub Issue 5: Paragraph and numbering properly used
      - Sub Issue 6: Page numbers properly sequenced

   c. Understanding
      - Sub Issue 1: Contents synthesized according to the studied issues
      - Sub Issue 2: Contents well re-paraphrased
      - Sub Issue 3: Contents consistent with the research project.

   d. Completeness
      - Sub Issue 1: Concepts or theories related to research project
      - Sub Issue 2: Research results shown in solving the problem in the research project
      - Sub Issue 3: Giving a theory of relevant knowledge and programs used in the research project
      - Sub Issue 4: Having at least two parts: theories and related works

4) System Analysis

   a. Content Correctness
      - Sub Issue 1: Correctness of relationship between Context diagram (DFD0) and Data Flow Diagram level 1 (DFD1)

- Sub Issue 2: Correctness of relationship between data stored in DFD1 and Entity Relationship (ER) diagram
- Sub Issue 3: Correctness of relationship between Entities in ER diagram and Attributes in Data Dictionary
- Sub Issue 4: Tables in database corresponding with ER diagram
- Sub Issue 5: Correctness of DFDs
- Sub Issue 6: Example data in data dictionary from the real data

b. Format Correctness
- Sub Issue 1: Table Format
- Sub Issue 2: Page numbering
- Sub Issue 3: Picture arrangement
- Sub Issue 4: Font
- Sub Issue 5: Margin
- Sub Issue 6: Format of DFD
- Sub Issue 7: Format of ER diagram
- Sub Issue 8: Format of Data Dictionary

c. Understanding
- Sub Issue 1: Diagrams being clear
- Sub Issue 2: Lines in diagrams not overlapped
- Sub Issue 3: Lines in diagrams being straight
- Sub Issue 4: Symbols being used if unconnected with the desired symbols
- Sub Issue 5: Relationships in diagram not ambiguous

d. Completeness
- Sub Issue 1: Completeness of dataflow diagram
- Sub Issue 2: Completeness of Database Design described in ER diagram
- Sub Issue 3: Stating issues of system analysis

5) System Design

a. Content Correctness
- Sub Issue 1: Consistency in user interfaces
- Sub Issue 2: Design properly divided: first middle and last sections
- Sub Issue 3: System designed for all level of users
- Sub Issue 4: Inputs sufficiently designed for the needs
- Sub Issue 5: Outputs/Reports designed for all levels of users
- Sub Issue 6: Software and hardware selected as suitable for the system

b. Format Correctness
- Sub Issue 1: Sequence of pictures
- Sub Issue 2: Margin
- Sub Issue 3: Paragraph and Line spacing
- Sub Issue 4: Picture size
- Sub Issue 5: Font and size

c. Understanding
   - Sub Issue 1: Pictures clear and easy to understand
   - Sub Issue 2: Order of pictures with a clear sequence
   - Sub Issue 3: The design used to develop user interfaces easily
   - Sub Issue 4: Clear description of the pictures

d. Completeness
   - Sub Issue 1: Completeness of pictures
   - Sub Issue 2: Completeness of screen layout of user interfaces
   - Sub Issue 3: All software and hardware specified in the system

6) Conclusion and Discussion

   a. Content Correctness
      - Sub Issue 1: Statistics properly used
      - Sub Issue 2: Research results connected to concepts and theories described in Chapter 2
      - Sub Issue 3: Conclusion consistent with the objectives
      - Sub Issue 4: Containing analysis, synthesis and conclusion of actual research results

   b. Format Correctness
      - Sub Issue 1: Tables properly used
      - Sub Issue 2: Table format
      - Sub Issue 3: Paragraph and Line spacing
      - Sub Issue 4: Page number
      - Sub Issue 5: Margin
      - Sub Issue 6: Font and size

   c. Understanding
      - Sub Issue 1: Research results easy to understand
      - Sub Issue 2: Research results showing unity and clarity
      - Sub Issue 3: The results separately discussed, issue by issue

   d. Completeness
      - Sub Issue 1: Complete documentation
      - Sub Issue 2: Complete discussion on all issues

Example scores of the abstract are displayed in Table 1. The average scores of the main issues were later computed as examples shown in Table 2. Finally, an average score of all chapters, and the computer project document assessment score were computed. Assessment scores for sub issues were defined as follows:

1) Very Good: Score value of 5. The document achieves 80 percent of quality for the assessment issue.
2) Good: Score value of 4. The document achieves 70 percent of quality for the assessment issue.
3) Fair: Score value of 3. The document achieves 60 percent of quality for the assessment issue.
4) Poor: Score value of 2. The document achieves 50 percent of quality for the assessment issue.

5) Very Poor: Score value of 1. The document achieves less than 50 percent of quality for the assessment issue.

Table 1: Examples of Sub Issue Assessment Scores: Abstract

Sub Issues	Committee Member 1	Committee Member 2	Committee Member 3	Committee Member 4	Committee Member 5
Sub Issue 1	5	4	2	5	4
Sub Issue 2	4	2	5	4	5
Sub Issue 3	5	4	5	5	5
Sub Issue 4	4	4	4	4	5
Sub Issue 5	5	4	5	4	5
Sub Issue 6	4	4	5	4	4

Table 2: Examples of Main Issue Assessment Scores: All Chapters

Issues	Abstract	Introduction	Literature Review	System Analysis	System Design	Conclusion and Discussion
Content Correctness	4.30	2.47	3.84	4.44	3.18	4.21
Format correctness	3.54	4.97	3.85	3.64	4.73	3.51
Understanding	3.01	2.58	4.76	3.99	3.85	4.54
Completeness	4.29	3.88	4.61	2.56	4.04	3.23

The computer project document assessment score of the document in Table 2 is 3.83.

4.2 Measuring Accuracy of Prediction Scores

Sometimes committee members may not assess all sub issues of the project document. The missing scores were estimated using UBCF algorithm. It consists of two main steps:

1) Calculating the similarity scores rated by committees using Pearson correlation-based similarity (Xiaoyuan & Taghi, 2009) as expressed in Equation 1.
2) Predicting the missing score(s) of sub issue(s) using weighted sum as displayed in Equation 2.

$$w_{u,v} = \frac{\sum_{i\in I}(r_{u,i}-\bar{r}_u)(r_{v,i}-\bar{r}_v)}{\sqrt{\sum_{i\in I}(r_{u,i}-\bar{r}_u)^2 \sum_{i\in I}(r_{v,i}-\bar{r}_v)^2}}\ , \qquad (1)$$

where $w_{\mathrm{u,v}}$ similarity between committee *u* and *v*
i sub/main issue *i* rated by both committee members *u* and *v*
I *all* sub/main issues rated by both committee members *u* and *v*

$r_{u,i}$ score of sub/main issue *i* rated by committee member *u*
$r_{v,j}$ score of sub/main issue *i* rated by committee members *v*
$\bar{r}_u$ average score of the co-rated sub/main issues rated by committee member *u*
$\bar{r}_v$ average score of the co-rated sub/main issues rated by committee member *v*

$$P_{u,i} = \frac{\sum_{n \in N} r_{u,n} w_{i,n}}{\sum_{n \in N} |w_{i,n}|}, \tag{2}$$

where *Pu,i* prediction score of sub/main issue *i* of committee member *u*
u committee member *u*
i is a sub/main issue *i*
n other sub/main issue *n* rated by committee member *u*
N all other sub/main issue rated by committee member *u*
$r_{u,n}$ score of sub/main issue *n* rated by committee member *u*
$w_{i,n}$ similarity between committee members *u* and other for sub/main issue *n*

Only the similarities ($w_{u,v}$) having value greater than zero are considered because the negative value means the score rated by two committee members are not similar or not correlated with each other.

Table 1 assumes that the sub issue 1 of the committee 1 is missing. The given example shows how to compute the missing score.

$$\bar{r}_1 = \frac{4+5+4+5+4}{5} = 4.4$$

$$\bar{r}_2 = \frac{2+4+4+4+4}{5} = 3.6$$

$$w_{1,2} = \frac{\left(\left(4\text{-}4.4\right)*\left(2\text{-}3.6\right)\right)+\left(\left(5\text{-}4.4\right)*\left(4\text{-}3.6\right)\right)+\left(\left(4\text{-}4.4\right)*\left(4\text{-}3.6\right)\right)+\left(\left(5\text{-}4.4\right)*\left(4\text{-}3.6\right)\right)*\left(\left(4\text{-}4.4\right)*\left(4\text{-}3.6\right)\right)}{\sqrt{\left(\left(4\text{-}4.4\right)^2+\left(5\text{-}4.4\right)^2+\left(4\text{-}4.4\right)^2+\left(5\text{-}4.4\right)^2+\left(4\text{-}4.4\right)^2\right)*\left(\left(2\text{-}3.6\right)^2+\left(4\text{-}3.6\right)^2+\left(4\text{-}3.6\right)^2+\left(4\text{-}3.6\right)^2+\left(4\text{-}3.6\right)^2\right.}\left.\right)}$$

$$= 0.8/1.959592$$

$$= 0.408248$$

For similarity scores between committee members 1 and the rest are:

$w_{1,3} = 0.408$

$w_{1,4} = 0.612$

$w_{1,5} = 0.408$

The missing score of sub issue 1 of committee member 1 is later predicted using Equation 2 as follows:

$$P_{1,1} = \frac{(4*0.408)+(2*0.408)+(5*0.612)+\ (4*0.408)}{0.408+0.408+0.612+0.408} = 3.89$$

The prediction score is 3.89; the actual score value is 5.

If all scores rated by the committee members (except the missing score) are equal, the $w_{u,v}$ will be equal to zero. In such a case, the missing score is set as well as the others as an example shown in Table 3.

Table 3: Examples of Scores Causing Zero Similarity Value

Sub Issues	Committee Member 1	Committee Member 2	Committee Member 3	Committee Member 4	Committee Member 5
Sub Issue 1	5	5	4	2	4
Sub Issue 2	4	4	4	2	5
Sub Issue 3	4	4		3	2
Sub Issue 4	5	4	4	4	3
Sub Issue 5	4	4	4	2	3
Sub Issue 6	5	5	4	4	2

$$\bar{r}_3 = \frac{4+4+4+4+4}{5} = 4$$

$$\bar{r}_1 = \frac{5+4+5+4+5}{5} = 4.6$$

$$w_{3,1} = \frac{((4\text{-}4)*(5\text{-}4.6))+((4\text{-}4)*(4\text{-}4.6))+((4\text{-}4)*(5\text{-}4.6))+((4\text{-}4)*(4\text{-}4.6))*((4\text{-}4)*(5\text{-}4.6))}{\sqrt{((4\text{-}4)^2+(4\text{-}4)^2+(4\text{-}4)^2+(4\text{-}4)^2+(4\text{-}4)^2)*((5\text{-}4.6)^2+(4\text{-}4.6)^2+(5\text{-}4.6)^2+(4\text{-}4.6)^2+(5\text{-}4.6)^2)}} = 0$$

In this case, the missing score in Table 3 was set to 4.

If a committee member does not rate all sub issues of a main issue. The average scores of all sub issues of that main issues are used. The average scores of neighbors or other committee members are used in the collaborative filtering shown as examples in Table 4.

Table 4: Examples of Missing Scores of All Sub Issues of Content Correctness in Introduction

Issues	Content Correctness					Format Correctness	Understanding	Completeness
	C1	C2	C3	C4	C5	…	…	…
Abstract	3.77	3.52	4.93	3.08	3.71	…	…	…
Introduction	3.64		3.70	3.79	4.26	…	…	…

Issues	**Content Correctness**					**Format Correctness**	**Understanding**	**Completeness**
	C1	**C2**	**C3**	**C4**	**C5**			
Literature Review	3.77	4.18	4.61	4.51	4.70	…	…	…
System Analysis	3.76	3.88	4.09	3.50	4.65	…	…	…
System Design	3.86	3.98	4.19	3.42	4.20	…	…	…
Conclusion and Discussion	3.37	4.41	3.07	4.11	3.19	…	…	…

Note: C1, … C5 refer to Committee Member 1, … Committee Member 5.

In Table 4, the similarities computed by Equation 1 are:

$w_{2,1}$ = -0.649102303

$w_{2,3}$ = -0.772721764

$w_{2,4}$ = 0.835633748

$w_{2,5}$ = -0.17040505

The missing score of Committee Member 2 for the introduction chapter calculated by Equation 2 is:

$$P_{1,1} = \frac{(3.79 * 0.84)}{0.84} = 3.79$$

4.3 Predicting Computer Project Document Assessment Missing Scores

The accuracy of the prediction scores were measured by two equations: Equation 3 and Equation 4. Mean Absolute Percentage Error (MAPE) (Stephanie, 2017) is used in Equation 3 for measuring percent error, and percent accuracy is used in Equation 4.

$$M = \frac{100}{n} \sum_{t=1}^{n} \frac{|A_t - F_t|}{At}, \tag{3}$$

where M a mean absolute percentage error

A_t an actual score value

Ft a prediction score value

n number of prediction score values

$$\text{Accuracy} = 100 - M \tag{4}$$

The prediction value of Sub Issue 1 of Committee Member 1 in Table 1 is 3.89, and the actual value is 5. The number of missing values (n) for this case is 1. The accuracy is computed as follows:

$$M = 100 * \left(\frac{|3.89-5|}{5}\right) = 22.2$$

Accuracy = 100 – 22.2 = 77.8%

5. Experimental Setup and Experimental Results

The assessment form was produced by Google form. There were 20 computer project documents assessed by the committee. There were 5 committee members assessing each document. The scores after rating by the committee members were divided into two groups:

1) Scores at the sub issue level: Each data set looks like Table 1. The assessment scores of sub issues of these data sets were removed in values 1 to 5 randomly. There were 80 data sets generated for this level.
2) Scores at the main issue level: For each data set, average scores of each main issue for each committee member were calculated. Each data set looks like Table 4. The average assessment scores of main issues of these data sets were also removed in values 1 to 5 randomly. There were 80 data sets generated for this level.

The missing scores were predicted and their accuracies were computed. The experimental results are displayed in Table 5 and 6.

Table 5: Experimental Results of Sub Issue Data Sets

Number of Removed	*M*	Accuracy
1	22.10	77.90
2	21.87	78.13
3	21.78	78.22
4	21.74	78.26
5	21.77	78.23
Average	21.85	78.15

Table 6: Experimental Results of Main Issue Data Sets

Number of Removed	*M*	Accuracy
1	15.63	84.37
2	15.75	84.25
3	18.07	81.93
4	15.21	84.79
5	15.09	84.91
Average	15.95	84.05

6. Conclusion and Discussion

The collaborative filtering approach was applied to predict missing scores of computer project document assessment. The Google form was used to gather assessment scores from committee members. The assessment scores rated by committee members were divided into two groups: assessment scores at the sub issue level and assessment scores at the main issue level. The experiment data were set by removing assessment scores of these two groups in values 1 to 5 randomly. There were 80 data sets for each group. Missing scores or removed scores of these experiment data were predicted. The experimental results revealed that the accuracy of the data at the sub issue level and the main issue level were 78 percent and 84 percent in average, respectively.

The assessment scores at the main issue level are from an average of all sub issues under a particular main issue. Therefore, the assessment scores at the main issue level do not cause much fluctuation. As reported in this study, the results in accuracy of predictive scores at the main issue level appeared better than those at the sub issue level. The obtained findings suggest that it should be possible, in addition to quality assessment, to apply the collaborative filtering approach to predict a set of products and services for users who have the same rating behaviors in business contexts as seen fit by decision-makers concerned.

7. The Authors

Areerat Trongratsameethong is a lecturer of the Computer Science Department, Faculty of Science, Chiang Mai University. She has her Ph.D. in Computer Science from Mahidol University, Bangkok, Thailand. She worked for business companies and software houses for many years while gaining experience in software development, system analysis and design, and software requirement analysis. Her current expertise and teaching focus on fundamentals of programming, fundamentals of database system, object-oriented design, and ontology design and development.

Ataya Tipprason is a developer at Lampang Provincial Public Health Office. She has a Master of Science degree in Computer Science from Chiang Mai University, Chiang Mai, Thailand. Her research interest lies in the areas of database system and object-oriented design.

8. References

Chen, Z., Liu, X. & Shang, L. (2020). Improved course recommendation algorithm based on collaborative filtering. *Proceedings of 2020 International Conference on Big Data and Informatization Education (ICBDIE),* Zhangjiajie, China, 466-469.

Dou, Y., Yang, H. & Deng, X. (2016). A survey of collaborative filtering algorithms for social recommender systems. *Proceedings of The 12th International Conference on Semantics, Knowledge and Grids (SKG) 2016,* Beijing, 40-46.

Purnomo, J. E. & Endah, S. N. (2019). "Rating prediction on movie recommendation system: Collaborative Filtering Algorithm (CFA) vs. Dissymetrical Percentage Collaborative Filtering Algorithm (DSPCFA). *Proceedings of The 3rd International Conference on Informatics and Computational Sciences (ICICoS) 2019,* Semarang, Indonesia, 1-6.

Stephanie, G. (2017). Mean absolute percentage error (MAPE). *Statistics How To [Internet]* (Online). https://www.statisticshowto.com/mean-absolute-percentage-error-mape/, December 1, 2020.

Xiaoyuan, S. & Taghi, M. K. A. (2009). Survey of collaborative filtering techniques. *Hindawi Advances in Artificial Intelligence 2009(421425),* 1-19.

Ying-Wei, C., Xin, X. & Yong-Ge, S. (2012). A collaborative filtering recommendation algorithm based on contents' genome. *Proceedings of IET International Conference on Information Science and Control Engineering 2012 (ICISCE 2012),* Shenzhen, 1-4.