

A Multi- Study Program Recommender System Using Content-Based Filtering and Analytical Hierarchy Process (AHP) Methods

Linda Marlinda¹, Priadhana Edi Kreshana² and Dana Indra Sensuse³

¹ Department of Information Management, AMIK Bina Sarana Informatika, Kramat Raya 25, 10450, Jakarta

² Faculty of Computer Science, University of Indonesia, UI Depok
Jakarta, 16424, Depok

³ Faculty of Computer Science, University of Indonesia, UI Depok
Jakarta, 16424, Depok

Abstract

Deciding a field of study from several provided fields is one of the most essential moments in a new student life because it is related to the future he/she wants to be. Since there are many criteria that must be considered, then through this recommender system, students are able to know what fields are the most appropriate for them. This system is built based on Analytical Hierarchical Process (AHP) method and Content Based filtering. When a student fills out a questionnaire, he must be consistent with his/her answer to obtain the best output based on his/her will and characteristics. After conducting some experiments of 100 respondents, 87 respondents succeeded filling out the questionnaire completely. Among provided field of studies, 22.99% of them chosen are medicine, 16.09% are accounting, and 5.7% are not chosen at all.

Keywords: *Recommender System, Analytical Hierarchy Process (AHP) Method, Content Based Filtering, Multi Study Program.*

1. Introduction

According to the Private Higher Education Coordinating Board (Kopertis) Region III resolution concerning regulation and codification of field of study in college/university, there are 524 fields of study opened by all colleges/universities in Jakarta [1]. In this study, the authors investigated 38 main fields of study with 169 sub-study fields.

After graduating from senior high school, usually a lot of students encounter difficulties in deciding desired field of study when they enter higher education. This

phenomenon happens because many fields of study opened by the colleges/universities, whether existing or new fields, without enough information about it. If the they facilitate abundant information to the prospective students, there is possibility that the students will be confident enough to decide field of study that they want to take based on the expected outcome they will achieve for the future.

An analysis of multi-criteria recommender system that is built based on Analytical Hierarchical Process (AHP) and Content Based Filtering (CBF) techniques is expected to be able to provide recommendations that a student needed. Information filtering process in the decision recommender system aims to generate a recommendation for a student as a guidance to choose what field of study he/she desired. This process is conducted by classifying the criteria, sub criteria, and the alternatives from the questionnaire provided. The process of filling out the questionnaire by the student must be consistent to achieve the best expected outcome from the system.

2. Methodology

This section briefly explains Recommender Systems, Content Based Filtering (CBF), and Analytical Hierarchical Process (AHP), as well as investigates some researches which have been done related to the implementation of the recommender systems, then finally methodology used in the study.

2.1 Recommender Systems

According to [2] a recommender system is a system which recommends a solution of explicit problems based on hierarchical structure. In the most general formulation, recommendation problems are reduced by rating to items which have not been recognized by a user.

A recommender system itself can be defined as a system which generates individual recommendation as an output, or it is also defined as a system which helps user to find desired items by giving recommendation based on one of the recommended content items (content-based filtering) or similar user ratings on recommended items (collaborative filtering). In another area, for example, an information filtering recommender system uses techniques of problems analysis to help customers find products they want to buy by producing similarity of prediction score or a list of recommended products for specific customers [3].

The recommender system is also usually classified into the following categories [2] :

- a. Content based recommendation (Content-based Filtering): The user will be recommended items similar to the user preferred in the past.
- b. Collaborative recommendations (Collaborative Filtering): The user will be recommended items similar to the user preference and interest in the past;
- c. Hybrid approach (hybrid collaboration): These methods combine collaborative and content-based methods.

Collaborative Recommendation (Collaborative Filtering) method is able to predict user interest to a new item based on other users which have the same tendency. Instead of content analysis, a collaborative filtering system relies absolutely on interest ratings of participated community members. This recommender system has two categories as below.

1. Cognitive system. It characterizes the content of information sources and information needs of potential users of the item, and then uses the representations to intelligently match the item to the user.
2. Sociological filtering system. It works based on personal relationships and individual organizations within a community.

2 Content-based Filtering (CBF)

The Content-Based Filtering Recommender System provides recommendations by comparing content representation contained by an item [2]. The content-based recommender system is restricted by following limitations.

- a. Limited Content Analysis

Content-based technique is limited by features that are explicitly associated with the objects that recommend this system.

- b. When the system can only recommend items that score is highly against a user profile, the user is restricted to be recommended items similar to those already rated.
- c. New users Problem.

Users must have enough information to assess the item before the recommender content-based system can really understand user's preferences and presents a reliable recommendation to the user. In the content-based recommender system, item or object is defined by its features [3]. The implementation of the item features can be anything, as long as the features can describe the item well. The examples of the features are keywords, which are usually used to describe article, webpage, etc.

The content-based filtering approach recommends item to the user based on item description which has been evaluated. In other words, the system will recommend an item because the item similar to the item preferred by the user sometime in the past [4].

2.3 Multi Criteria Decision Making (MCDM)

Multi Criteria Decision Making (MCDM) is a decision-making technique based on several existing alternatives or a theory which explains about decision making process by considering many criteria. In order to model recommendation problems as MCDM, one must follow four general steps of modeling methodology to make a decision for the problem [5]:

- a. Decision object. It defines the object on which decisions must be made and the reasons of the decision recommendation.
- b. Family of criteria. Namely, it defines identification and modeling of a set of criteria that influence decision, and a complete and non-redundant recommendation.
- c. Global preference model. It defines aggregator function for marginal preference on each criterion to global preference from decision maker for every item.
- d. Decision support process. Namely, it defines study of various categories and types of recommender systems that can be used to support decision makers' recommendations, in accordance to the results of previous steps.

Implementation of the Multi Criteria Decision Making (MCDM) method in a recommender system has yet to be explored systematically. System recommender is capable to explain some system contributions which involve several MCDM methods. This system has many domain applications. On the other hand, a comprehensive

analysis which has been done will facilitate understanding and system development that can identify dimension which distinguish, explain, and categorize multi-criteria system recommender, based on existing taxonomy and categorization that is used in analysis and classification of online user decision sample.

MCDM is a theory of decision making that considers a limited set of alternative options against many criteria. The problem in MCDM can be formulated as follows:

Suppose there are M criteria and N alternatives. We must choose some or a set of alternatives which fulfill criteria as maximum as possible [6]. MCDM problem can be modeled in decision matrix.

Table 1: Decision Matrix [6]

Alternatif	Criteria				
	C ₁ W ₁	C ₂ W ₂	C ₃ W ₃	...	C _N W _N
A ₁	a ₁₁	a ₁₂	a ₁₃	...	a _{1N}
A ₂	a ₂₁	a ₂₂	a ₂₃	...	a _{2N}
A ₃	a ₃₁	a ₃₂	a ₃₃	...	a _{3N}
...
A _M	a _{M1}	a _{M2}	a _{M3}	...	a _{MN}

Decision matrix is a matrix of size M x N where the element a_{ij} indicates the performance of alternative A_i when evaluated against the criteria C_j (for i = 1,2,3, ..., M and j = 1,2,3, ..., N).

2.4 Analytic Hierarchy Process (AHP)

Analytic Hierarchy Process (AHP) is an MCDM method introduced by [7]. AHP uses hierarchical structure, from objectives, criteria, sub criteria, to alternatives. Therefore there is a possibility to compare among alternatives using hierarchical criteria. In the implementation, AHP uses *pair wise comparison* to measure degree of importance among criteria, also relative performance among alternatives on each criterion. AHP uses a scale in determining the relative importance [8].

In everyday life, decision making is usually done by using intuition. However, this method has many shortage, thus a new systematic method was developed. It is called decision analysis, which contains intelligence, perception, and philosophy. After using intelligence, perception, and philosophy to create the model, determine possible value, establish value in the expected result and explore preference towards time and preference towards risk, then to reach at a decision, logic is required.

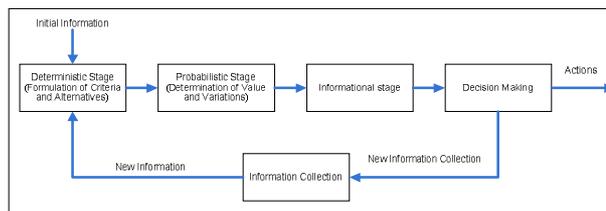


Fig. 1 A Cycle of Rational Decision Analysis [9]

AHP calculation can be described by using the following formula [8]:

$$A_{AHP}^i = \sum_{j=1}^N a_{ij} w_j, i = 1,2,3, \dots, N \tag{1}$$

The study begins by observing implementation of the study field recommender system, which aims to minimize possible mistakes caused by the failure in selecting a field of study. This research uses descriptive analysis method by presenting survey and interview results in the form of questionnaire manually.

Next, collecting secondary data on site is conducted toward prospective students so that accurate information of the study field decision can be obtained. In addition to that, identification of the system is done with considering supporting variables through filling out the questionnaire into the system directly.

Then results of the interviews and the questionnaires become data to be processed by using AHP (Analytical Hierarchical Process) approach to obtain strategic steps that must be done when deciding a study field. The decision has to be followed up immediately by some actions or it can be reviewed if there is new information that affects the outcome reducing uncertainty, so that it will obtain a new decision.

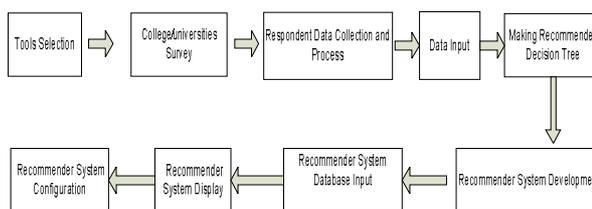


Fig. 2 Steps Conducted in the Study

Data collection as shown at Fig. 2 was done by using two steps. In the first step, a questionnaire is filled out manually, and in the second step, filling out through the system that is implemented directly to the respondents.

The questionnaire data is processed by using AHP approach with matrix manipulation as a comparison analysis. In order to determine strategic step priority in

implementing the main study field decision, then we propose fourteen criteria, and thirty eight alternatives of the main study field. The decision is based on AHP model developed in the recommender system. AHP structure proposed in this research consists of:

- a. Criteria containing all testing subjects in national exams for the degree of senior high school, vocational high school, or Madrasah Aliyah.
- b. Alternatives that focus on the main course consisting of 38 main courses.

In next step, a user must rate several alternatives of the sub study fields according to some views towards criteria that must be chosen. The criteria can be divided into:

- a. Criteria which focus on internal and external part of new applicants, and career.
- b. Sub criteria that consist of the needs on reliable graduate or student, interest, talent, department origin, teacher, tuition, scholarship, student capacity, campus popularity, lecturer quality and quantity, campus location, availability of alumni forum, jobs / career facilitation, parents decision, network with industries, classmate influence, accreditation of study program, study facilities (location, hotspot area, library, class facilities (AC, LCD, computer), SW and HW lab, canteen, sports field, mosque, and content of career)
- c. Alternatives containing decision of the sub study field.

3. Implementation and Results

Implementation of the Recommender System is done by using Content-based Filtering (CBF). Information filtering process aims to generate a decision recommendation expected by the new student. This process is done by classifying criteria, sub criteria and alternative in the existing questionnaire. Rating the items by respondents must be consistent to obtain the expected output.

3.1 ERD (Entity Relationship Diagram)

ERD is a modeling technique in relational database which based on perception in real world. This world is always consists of a set of objects that is connected to each other. An object is called entity and the connection with other object is called relationship. An entity has unique characteristic and has attributes to distinguish with other entities. The ERD of the recommender system is shown at Fig 3, while the class diagram is shown at Fig 4.

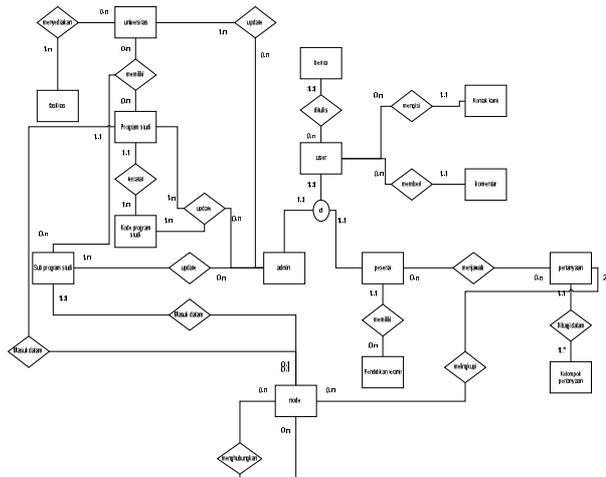


Fig. 3 ERD of Study Program Recommender System

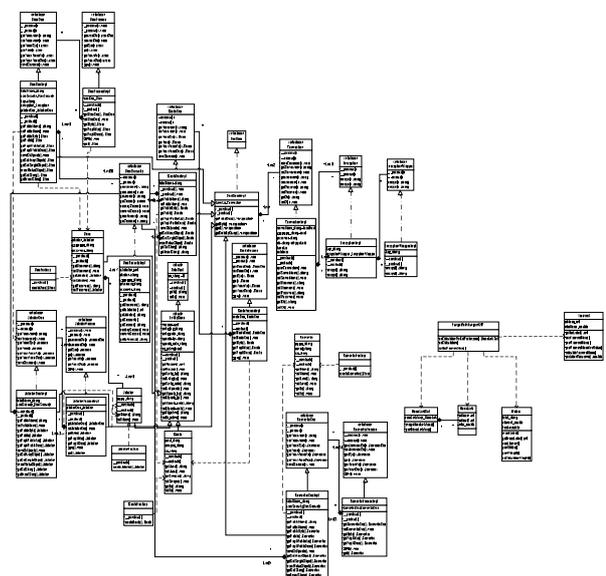


Fig. 4 Class Diagram of Study Program Recommender System

The user interface of the recommender system is designed as user friendly as possible so that a user does not need to take a training for using the system. The login screen of the system and the online questionnaire page are shown as below.



Fig. 5 Login Screen



Fig. 6 Questionnaire of the Consultation Page (AHP)

3.2 Consultation Page

This page serves as a questionnaire for deciding a study field through the system. To consult what a study field that will be recommended by the system, there are three steps:

- First step is answering the question based on the most significant subject in deciding a study field. There are 14 subjects available. They are Physics, Chemistry, Biology, Math, Bahasa Indonesia, English, Foreign language, geography, Science of Hadith, Sociology, Theory of vocational, Science of interpretation, History / anthropology, and Economics, in which they are compared each other. The subjects provided are those that are tested in SMA, SMK, and Aliyah National Testing.
- Second step is answering the question is based on subjects that have influences in deciding the main field of study which is divided into 169 question groups.
- Having succeeded in determining the main field of study, next step is deciding sub field of study from the selected main field of study.

Based on the questionnaires answered by respondents, it can be concluded that computer science is the most influence main study field with percentage of 2,86%; and for the sub study field, a respondent tends to decide

computer system with percentage of 23,74% from the same main study field in Computer Science. The respondent is able to see which colleges/universities that provide a such study field in Kopertis Region III, and percentage of the study field chosen, as shown at the figure below,



Fig. 7 Percentage of the Selected Main Alternative Field Studies

4. Conclusion

The multi-criteria recommender system with AHP and Content-based Filtering has been developed to support potential students or high school graduates in selecting a field of study. The multi-criteria decision making implemented in the system is able to solve complex problems, and able to handle interdependence of elements in the system and not imposing linear thought.

The recommender system provides information and recommendation that is most appropriate for the student directly so that he / she can be supported in deciding the best field for future study. This recommender system can intelligently generate decision quickly as soon as the respondents fill out a study field, which is impossible to do using manual application (i.e. expert choice), where the respondent cannot obtain the result in real time. The disadvantage of current recommender system is that this system does not show respondent inconsistencies as the expert choice does for every criteria, sub criteria, and alternatives.

Acknowledgments

We are truly grateful for the tremendous support from DIKTI that gave fund for this research, and the e-Government Laboratory, University of Indonesia that patiently gave guidance during this research.

References

- [1] Directorate General of Higher Education, Socialization Guidance, Jakarta, Ministry of Education. DGHE, 2009.
- [2] Adomavicius, G., et al. "Toward the Next Generation of Recommender Systems: A Survey of the State of the Art and Possible Extensions", IEEE Transactions on Knowledge and Data Engineering. 2005.
- [3] Burke, R. "Hybrid Recommender Systems: Survey and Experiments". User Adapt. Inter. Vol. 12 pp. 331-370, 2002.
- [4] Montainier, M., et al, "A Taxonomy of Recommender Agents on the Internet. Artificial Intelligence", Kluwer Academic Publisher. Netherlands, Review 19: 285-330, 2003.
- [5] Manouselis, N., et al. "Analysis and Classification of Multi-Criteria Recommender System", World Wide Web: Internet and Web Information Systems, Special Issue on Multi-channel Adaptive Information System on the World Wide Web. Springerlink, 2007.
- [6] Fulop, J. "Introduction to Decision Making Methods", Laboratory of Operations Research and Decision Systems, Computer and Automation Institute, Hungarian Academy of Sciences, 2007.
- [7] Saaty, T., L. "The Analytic Hierarchy Process", British Library Cataloguing in Publication Data, USA, 1988.
- [8] Triantaphyllou, E., et al. "A Computational Evaluation of The Original and Revised Analytic Hierarchy Process", Computer ind. Engng Vol. 26, No.3, pp.609-618. Elsevier Science Ltd, 1994.
- [9] Marimin. "Teknik dan Aplikasi Pengambilan Keputusan Kriteria Majemuk". PT. Gramedia Widiasarana, Indonesia. Jakarta, 2005.

Linda Marlinda earns a bachelor's degree from the University of Information Engineering YARSI Jakarta in 1995 and a master's degree in Computer Science from STMIK Nusa Mandiri Jakarta in 2010. She is a lecturer and researcher at AMIK BSI Jakarta. Her research areas are expert systems, database systems and information system design.

Priadhana Edi Kresnha achieved his bachelor degree from School of Computer Science, University of Indonesia in 2007 and master degree from the same university in 2010. Now he is pursuing his doctoral degree in Computational Intelligence field. His research areas include robotics, cryptography, computer graphics, computer vision, and their applications to the entire domain of Computer Science.

Dana Indra Sensuse earns his bachelor degree in Soil Science (Bogor Agricultural University, Indonesia, 1985), Master of Library and Information Studies (Dalhousie University, Halifax, Canada, 1994), and Ph.D in Information Studies (Toronto University, Canada, 2004), He is a lecturer at University of Indonesia, and serves as Head of e-Government Lab at University of Indonesia.