Mobile-based CBR System using Dempster-Shafer Modification Rule for Tourist Spots Recommendations in Buleleng Regency

I Made Agus Wirawan PTI, Undiksha Singaraja, Bali, Indonesia

ABSTRACT

Buleleng Regency has many unique cultures and beauty of natural sceneries which are always the charm and attraction for tourists. It requires the effort to increase the number of visits of the tourists by providing the information of tour object in accordance with the tourists' interest and also easily for accessible. This study aimed at designing a mobile-based CBR system using Dempster-Shafer modification rule to provide the recommendation of tourist spots in Buleleng Regency. The process of recommendation based on the data tourist trips in the past as represented into the case based. The data contained in the case-based consisted of a traveler profile and tourist spots that are visited. Traveler profile included gender, country, age, occupation, income per month, and the frequency of visits. The results of this study indicated that the mobile-based CBR system using the Dempster-Shafer modification rule can be applied in providing recommendations of tourist spots for tourists who visited Buleleng Regency. Based on tests performed using K-Fold Cross Validation, were showed that the accuracy average of recommendations of tourist spots was 1) 18% for fully accordance, 2) 62% for partial accordance and 3) 18% of error rate.

General Terms

Case Based Reasoning, Recommendation System.

Keywords

Mobile Phone, Case Based Reasoning, Recommendation System, Dempster-Shafer, Evidence Ullage, Tendency Factor.

1. INTRODUCTION

Buleleng Regency is one of the regencies in the northern part of Bali that has a variety of cultural uniqueness and beauty of the natural sceneries that have always been an attraction for tourists, either foreign tourists or domestic tourists. To increase the number of visits of the tourists, various efforts have been made by the Government of Buleleng through the Department of Culture and Tourism. One of them was by providing the tour object information through media website. However, this information was not based on the tourists' interest, therefore tourists were difficult to determine the tour object that will be visited. In addition, access information through the media website requires resources (internet connection) were quite large, therefore the information presented was very difficult to access in particular areas that have a low internet connection. Along with the rapid development of technology, need for easily accessible information 'wherever' and 'whenever' it one of the main factors in tourism promotion efforts [1].

Based on the description of the problem and the research that has been done before by [2, 3, 4, 5, 6, 7, 8, 9, 10] it is

Sri Hartati JIKE, UGM Yogyakarta, Indonesia

necessary for development of a mobile-based CBR system using the Dempster-Shafer modification rule to recommend tourist spots that exist in Buleleng Regency. The focus of this study were 1) to design a mobile-based CBR system using the Dempster Shafer modification rule for recommendations of tourist spots in Buleleng Regency and 2) to determine the accuracy of the resulting percentage of mobile-based CBR system using the Dempster-Shafer modification rule for recommendations of tourist spots in Buleleng Regency.

2. DEMPSTER-SHAFER THEORY

Dempster-Shafer method (DST) was first introduced by Dempster. He did the experiment with the uncertainty model with using range probability rather than a single probability. Then in 1976, the Dempster-Shafer theory published in a book called Mathematical Theory of Evidence [11]. Evidence theory is more flexible than the theory of probability [12].

2.1 Frame of Discernmnet (FOD)

In Dempster-Shafer theory is also known Frame Of Discernment or also called environment is denoted by Θ .

$$\Theta = \{ \theta_{l_1}, \theta_{2, \dots, n}, \theta_n \}$$
(1)

Where:

 Θ = FOD or *environment*.

 $\theta_{l_{n}}$, $\theta_{2\dots}$, θ_{n} = Element in the environment.

Environment is another term for the universe of discourse of set objects that are mutually exclusive. All subsets of the environment can produce as 2^N patterns formed from Θ until \emptyset , where N was the number of elements Θ .

2.2 Mass Function and Ignorance

DST is called degree of belief are assumed to be mass (represented by m) of an object or also called Basic Probability Assignment (BPA). Basic belief assignment can be obtained from the probability prior on start node [13]. Each mass can be expressed as a function that maps of the power set in interval between 0 and 1 [11].

m:
$$P(\Theta) \longrightarrow [0, 1]$$

Empty mass is defined as:

 $m(\emptyset) = 0$

While the number of overall mass for each subset X of the power set is 1.

$$\sum_{X \in P(\Theta)} m(X) = 1 \tag{2}$$

2.3 Combining Evidence

Combining Evidence is essential in the process of summarizing information from various sources. Where different sources (e.g. experts) give a different assessment for a set of events, therefore we need a combination of assessment of each source in compiling the information. To solve this problem is used Dempster's Combination Rule [11].

$$(m_1 \oplus m_2)(Z) = \frac{\sum_{X \cap Y = Z} m_1(X) m_2(Y)}{1 - \sum_{X \cap Y = \emptyset} m_1(X) m_2(Y)}$$
(3)

Where:

 $(m_1 \oplus m_2)(Z) =$ Mass function of evidence (Z), which is a result of a combination of evidence (X) and evidence (Y).

$(m_1)(X)$	= Mass function of evidence (X).
(m ₂)(Y)	= Mass function of evidence (Y).
k	= Mass of evidential conflict.

 \oplus = Operator direct sum.

One of the difficulties in the implementation of Dempster Shafer Theory is the normalization process and may lead to an opposite result than are expected [11]. In [10], Xu Ling-yu stated there are two factors that cause errors in the Dempster Shafer theory:

- a. Inadequate definition of compatibility evidence.
- b. Resource conflict between evidence.

Here some weaknesses of using Dempster Shafer Combination Rule, as follow:

i. Problem 1.

Assumed there is a set of hypotheses $\Theta = \{A, B, C\}$ with value:

 $m_1 \rightarrow m_1(\{A\})=0.99, m_1(\{B\})=0.01, m_1(\{C\})=0, m_1(\Theta)=0$

 $m_2 \rightarrow m_2(\{A\})=0, m_2(\{B\})=0.01, m_2(\{C\})=0.99, m_2(\Theta)=0$

By using the Dempster Shafer Combination Rule, can be obtained $m_3({A}) = 0$, $m_3({B}) = 1$, $m_3({C}) = 0$, $m(\Theta) = 0$. From these results indicate that the hypothesized value of B 100% correct. Obviously these conclusions is wrong, because the evidence given by both (m_1 and m_2) for hypothesis B is smaller (almost zero) than evidence A for $m_1({A}) = 0.99$ and evidence C for $m_2({C}) = 0.99$. However, because the value of the mass of evidence of mass $m_3({A})$ and $m_3({C})$ is 0. Such errors often arise in the real world.

ii. Problem 2.

Assumed there is a set of hypotheses Θ = {A, B, C} with value:

 $m_1 \rightarrow m_1(\{A\})=0.5, m_1(\{B,C\})=0.5, m_1(\Theta)=0$

$$m_2 \rightarrow m_2(\{A,B\})=0.5, m_2(\{C\})=0.5, m_2(\Theta)=0$$

By using the Dempster Shafer Combination Rule, it can be obtained: $m_3 (\{A\}) = 0.333$, $m_3 (\{B\}) = 0.333$, $m_3 (\{C\}) = 0.333$, $m (\Theta) = 0$. This suggested that the three propositions A, B, and C have the same degree of belief. However, this

conclusion was not correct, because there was no evidence that fully supported the hypothesis B. It can be seen from the DST rules, which a subset of evidence $m_1 (\{A\}) m_2 (\{A, B\})$ and $m_1 (\{A\}) m_2 (\{A\}) considered equal, but the actual evidence of <math>m_2 (\{A, B\})$ does not fully support A, but also support B. In a study conducted by [10] and [14], proposed a concept called ullage and evidence tendency factor as an extension of Xu Ling-yu's rule.

2.3.1 Evindece Ullage

Evidence ullage (Δ) is defined as a ratio of the mass (m) is lost when the combination of evidence m (A_i) with evidence m (B_j). Evindece ullage can be used to handle errors in combining between conflicts evidence and the compatible evidence. Value of Δ is determined by the subdivision of the evidence that shown in Table 1.

Table	1:E	vidence	ullage	combination
raute	1. L	viucnee	unage	combination

	Subdivisins	Δ
Compatible	Completely compatible	$\Delta_1 = 0$
Evidences	evidences	
for	$A_i = A, B_j = A$	
$A_i \cap B_j = A$	Inclusive compatible	Δ_2
-	evidences	$(0 < \Delta_2 < \Delta_3)$
	$A_i = A, B_i \supset A, or$	
	$B_i = A, A_i \supset A$	
	Intersectant compatible	$\Delta_3 = 0$
	evidences	$(\Delta_2 \leq \Delta_3 \leq 1)$
	$A_i \supset A, B_i \supset A, A_i \cap B_i = A$	
Conflict	Interference conflict	$\Delta_4 = 0$
Evidences	evidences	$(\Delta_4 < 1)$
for	$A_i = A, A_i \cap B_j = \emptyset$ or	
$A_i \cap B_j = \emptyset$	$B_i = A, A_i \cap B_i = \emptyset$	
	Negative conflict	$\Delta_5 = 1$
	evidences	
	$A_i \cap A = \emptyset, B_i \cap A = \emptyset,$	
	$A_i \cap B_i = \emptyset$	

Common forms of the evidence ullage formula, is defined by Eqs. (4)

$$\Delta(Z, m_1, m_2, X, Y) = 1 - \frac{|X \cap Z| + |Y \cap Z|}{2|X \cup Y|}$$
(4)

Where || is denoted as the number of the cardinality of the set of evidence.

2.3.2 Evidence Tendency Factor.

Evidence Tendency Factor (η) is the approval degree of A when $m_1(A_i)$ and $m_2(B_j)$ are combined. Assume evidence resources are 100% reliable, the evidence tendency factor η can be calculated by Eqs.(5):

$$\eta(Z, m_1, m_2, X, Y) = \begin{cases} 1 & X \cap Y = Z \\ \\ \frac{m_1(X)}{m_1(X) + m_2(Y)} & X = Z \land X \cap Y = \emptyset \\ \\ \frac{m_2(Y)}{m_1(X) + m_2(Y)} & Y = Z \land X \cap Y = \emptyset \end{cases}$$
(5)

2.3.3 Combination Rule.

In the study conducted by [10] and [14], Combination Rule used is an extension of Xu Ling-yu's rule, is defined by Eqs. (6)

$$(\boldsymbol{m}_{1} \oplus \boldsymbol{m}_{1})(\mathbf{Z}) = \frac{\sum \boldsymbol{m}_{1}(\mathbf{X})\boldsymbol{m}_{2}(\mathbf{Y})(1-\Delta)\boldsymbol{\eta}}{1-\sum \boldsymbol{m}_{1}(\mathbf{X})\boldsymbol{m}_{2}(\mathbf{Y})\Delta}$$
(6)

Where, $m_1(A_i) m_2(B_j) (1-\Delta)$ is the effective fraction of $m_1(A_i) m_2(B_j)$, and $m_1(A_i) m_2(B_j) (1-\Delta) \eta$ is the true value of the mass support A. It can be considered when $(1-\Delta) \eta$ is

worth 0 for $A_i \cap B_j = \emptyset$ as well as 1 to $Ai \cap B_j \neq \emptyset$. Xu Lingyu's rule still follows the shape of the traditional Dempster's rule.

3. RESULTS and DISCUSSION

A Dempster-Shafer modification rule is used to combine any evidence (traveler profile data). Here's an example of the process of recommendation on mobile-based CBR system using six traveler profile data, which is displayed one by one. Here's an example of a tourist trip data presented in Table 2.

ID	Gender	Age	State	Occupation	Income per month	Frequency of visits	Tour Object
1	Male	20 until 29 Years Old	Australia	Student	-	Twice	1003,1005,1015
2	Male	30 until 39 Years Old	Australia	Private Employees	\$1000 until \$5000	Once	1003,1005,1017,1018
3	Female	30 until 39 Years Old	Australia	Entrepreneur/ Business	\$1000 until \$5000	Twice	1003,1005,1015
4	Female	Upper 50 Years Old	Jepang	Pensioner	\$1000 until \$5000	Once	1004,1019,1020,1022
5	Female	30 until 39 Years Old	Hongkong	Entrepreneur/ Business	Upper \$50000	Twice	1001,1011,1019,1022

Tabel 2: Examples of tourist visitation data in Buleleng Regency

If there is a data traveler profile with name \rightarrow David, Gender \rightarrow Male, Age \rightarrow 30 until 39 Years Old, State \rightarrow France, Occupation \rightarrow Entrepreneur / Business, Income per month \rightarrow \$ 1,000 until \$ 5,000 and Frequency of visits \rightarrow Once. Based on the recommendations performed in the table 2, it can be concluded that the recommended tour object by mobile-based CBR system is Melanting Waterfall (1003), and Air Sanih (1005) with the assurance 0.63156 degrees.



Form

In the figure 1 is the splash form for mobile application. In the figure 2 displays the data of tourist spots and confidence level of the recommendation. In this interface, there are four processes used to revise the recommendations, such as the Save, Change, Add, and Delete.

Based on the tests has been performed using K-Fold Cross Validation (K=3) with 126 data of travel, it can be generated the average percentage of accuracy was 18% fully accordance,

62% partial accordance and error rate was 18%. In the first testing there were 2 data (5%) that had low levels of confidence.

4. CONCLUSION and SUGGESTION

Conclusions obtained in the study of mobile-based CBR system using the Dempster Shafer modification rule were:

- Based on the finding that has been described previously, mobile-based CBR systems using Dempster Shafer modification rule could be applied to provide recommendations for tourists to visit the tour object in Buleleng Regency.
- Based on the results of tests performed, mobile-based CBR systems using Dempster Shafer modification rule produced 1) 18% for fully accordance, 2) 62% for partial accordance and 3) 18% of error rate.

Based on the results obtained in this study, it can be suggested several things:

- 1. For further research, it is necessary to apply a method that can improve the process of computational of the Dempster Shafer modification rule.
- 2. It need to be applied a method that is able to handle many conflicting evidence when the combination rule
- 3. A mobile-based CBR system can be developed to determining the path to be traversed, or be integrated with a web service that will facilitate tourist arrivals.

5. ACKNOWLEDGMENTS

Authors would like to thank for the Department of Culture and Tourism in Buleleng Regency that has allowed and supported the researchers because this research can be carried out smoothly.

6. REFERENCES

- Goh, D.H., Ang. R.P., Lee, C.S., and Lee, C.K. 2010. Determining Services for the Mobile Tourist. J.Computer Information Systems, 51(1), 31 – 40.
- [2] Wellem, T., 2009, Semantic Web Sebagai Solusi Masalah Dalam E-Tourism Di Indonesia, *Pros.Seminar Nasional Aplikasi Teknologi Informasi 2009* (SNATI 2009).
- [3] Hafsah., Kaswidjanti, W. dan Cili, T.R., 2010, Aplikasi Berbasis Web Pemilihan Obyek Pariwisata Di Yogyakarta Menggunakan Metode Tahani, Pros. Seminar Nasional Informatika., ISSN: 1979-2328.
- [4] Niknafs, A.A., Shiri, M.E. dan Javidi, M.M., 2003, A Case-Based Reasoning Approach in E-Tourism: Tour Itinerary Planning, Proceedings of the 14th International Workshop on Database and Expert Systems Applications, J. IEEE., 1529-4188/03.
- [5] Chi, T.H., La, H.H., Chu, Y.H. dan Lin, W.C., 2009, A Mobile Tourism Application Model based on Collective Interactive Genetic Algorithms, *Fourth Int. Conference* on Computer Sciences and Convergence Information Technology, J. IEEE., DOI 10.1109/ICCIT.2009.280.
- [6] Miitina, T., 2005, Perbedaan Proporsi Segrmen Pasar Geografis, Demografis, Psikografis dan perilaku terhadap Keputusan Membeli Produk Wisata Budaya di Kalimantan Timur, J. Aplikasi manajemen., Vol 3, No 2.
- [7] Gavalas, D. dan Kenteris, M., 2011, A web-based pervasive recommendation system for mobile tourist guides, *j. springer.*, 15:759–770.

- [8] Wickramarathne, T.L., 2008, A Belief Theoretic Approach for Automated Collaborative Filtering, *Thesis*, Master of Science Faculty University of Miami, Miami
- [9] Petrovic, S., Mishra, N. dan Sundar, S., 2011, A novel case based reasoning approach to radiotherapy planning, *Int. J. Archive Expert Systems with Applications.*, Vol 38, Issue 9.
- [10] Ding, Y.J., Wang, S.L., Zhao, X.D., Lv, M. and Yang, X.Q. 2009. A Modified Dempster-Shafer Combination Rule Based on Evidence Ullage. Int. Conference on Artificial Intelligence and Computational Intelligence. J. IEEE, 10.1109/AICI.2009.117.
- [11] Giarratano, J. and Riley, G. 2005. Expert System Principles and Programming, Fourth Edition. Course Technology a divison of Thomson Learning.Inc. Canada.
- [12] Chowdhury, S. 2012. Fusing probability density function into Dempster–Shafer theory of evidence for the evaluation of water treatment plant. Int.J.Springer, 10.1007/s10661-012-2840-5.
- [13] Simon, C. and Weber, P. 2006. Bayesian Networks Implementation of the Dempster Shafer Theory to Model Reliability Uncertainty. Proc.the First Int. Conference on Availability, Reliability and Security (ARES). J. IEEE, 0-7695-2567-9/06.
- [14] Chang, H.L., Min, S.C., Zhong, Y.L. and Hee, Y.Y. 2010. Reasoning with Imprecise Context Using Improved Dempster-Shafer Theory. ACM Int.Conference on Web Intelligence and Intelligent Agent Technology. J. IEEE, 10.1109/WI-IAT.2010.190.