

Personalization Travel Support Engine Using Reinforcement Learning

Pisit Sukonmanee
Anongnart Srivihok
Arunee Intrapairot
g4464007@ku.ac.th

Abstract- *In the tourist industry, there are tremendous websites which providing information about tourisms. Tourists who seek traveling information always spend a lot of time in searching for the relevant information. This paper proposes a prototype of Personalization Travel Support Engine by using reinforcement learning theory. The engine can learn user's interests in the provided web site for travel information. The goal of this engine is to adaptively increase reward for interesting or relevant trip information and penalty for un interesting or non relevant trip information. The reinforcement learning is the algorithm to learn individual user's profiles and analyze user's interaction to the presented best trip information. The results from this study reveal that it is possible to develop Personalization travel support engine. As well, the information from this study provides useful information regarding the areas of personalization of web support system and e-Tourism.*

Keywords: Personalization, Reinforcement Learning, Travel, Thailand.

1. Introduction

At present information technology (IT) plays important role in working environment, many organizations use IT as a tool in making their business run smoother and competing faster in the market. In tourist industry, IT such as the Internet and WWW also has the major role in business process. On line business is more competitive than traditional one since there are plenty of low cost online stores offering products and services on the Internet. Further, customer royalty for on line business is low comparing to traditional market so that it is a difficult problem for a company to attract new and keep customers in e-Commerce. Traditional marketing is not always successful on the Internet, and thus more specific on-line system such as one-to-one marketing should be helpful. In order to be more competitive on the Internet marketing, it is compulsory to offer customers with products or services which are better suiting for each customer. (Changchien et al. 2004)

During the past few years online massive marketing by using a push technology and informative websites always containing a great deal of information have been introduced to users. Each day users always get plenty of spam mails for advertising purposes. These are not only increase lot of traffics on the Internet but also create some irritations to email receivers. Further there are the mountains of informative web sites providing tremendous information for users. The existing search engines do not allow users to find the relevant information easily. Due to these challenging, web personalization and one to one marketing has been introduced to the ecommerce business, including tourist sector, retail, banking and finance, and entertainments.

In this study Personalization Travel Support Engine is introduced to assist and manage traveling information for user. It is another means to offer the information that matches the users' requirements. This system applies the Reinforcement Learning to help analyzing the customer behavior and studying customer interest.

2. Related works

Lieberman [2] developed a program to record customer behaviors while accessing the website that was created. This program analyzed, estimated and tracked the user's browsing behaviors and interests from the hyperlink in an HTML and offered the suitable information to web users. The benefit of this method is that the user can find the information without any efforts to retrieve it, however, this program involved much time and efforts in machine learning.

Later, Joachims et al. [6] developed Web Watcher Program that could analyze the web user's interaction with the website. A Reinforcement Learning program was adopted together with the purpose to offer the most suitable information to user by showing link in HTML.

The WAIR system [9] explained about information filtering techniques, by using reinforcement learning program. System learnt the user's interests by observing his or her behaviors during the interaction with the system and the user personal information

that he or she was provided. Compared with other developed techniques, it was found that Reinforcement learning technique was the most efficient in retrieval information.

The study of Vassiliou et al [1] stated that the personalization system was divided into two categories, which are:

Content Based Filtering, which was based on the information retrieval and the case-based reasoning research. Its function was dependent on the relationship of the series of information and the past record of users' behavior and to compare those things to match the real unique interest of the user.

Collaborative Filtering, which is based on the similar behavior of users. This method needs the users to answer the personal data, which they like or dislike.

Yuan introduced the comparison shopping system [4] which supported the personalization system. Comparison shopping feature keeps the record of users, analyzes users' behavior, manage the record and gives the reward to the products based on those records. This method is called Temporal Difference Reinforcement Learning, which is one of the effective Reinforcement Learning process.

3. Related Theory

The characteristic of reinforcement learning [3] is a trial-and-error feature. The reward will be given when the answer to a question is correct, while the penalty will be awarded when there is an error. There are three elements involve in reinforcement learning, those are

- Agent that monitors the environment state
1. Environment state that records the information which impacts an action
 2. Action which the agent will perform to direct to the state

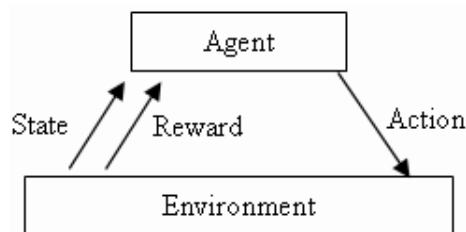


Figure 1. An agent interacting with its environment [7]

To allow learning agent be successful, one can set the goal with the elements of Reward function, which indicates the value of action in each state and the Value function indicating the total reward in each state.

Q Learning [3] is the reinforcement learning algorithm adopted to find the optimal policy. The

optimal policy that the agent can find is not the best result but is the most appropriate result at that time. This algorithm will give the state and select to go to that state which it analyzes as the related information. The information is in an alternative pattern, being in the form of $g(s_i, a_{ij})$, while g value is not greater than 1 when selecting to go to any state. Function $Q(s, a)$ is the function that yields the best reward. The function is in the form of an equation as follows:

$$\hat{Q}(s_t, a_t) \leftarrow \alpha \left[r + \gamma \max_{a_{t+1}} \hat{Q}(s_{t+1}, a_{t+1}) \right]$$

Whereas \hat{Q} denotes the value of taking action a in state s_t .

S_t denotes a set of states that agent can be in t , while t has a value of 1, 2, 3,...

$\max_{a_{t+1}} \hat{Q}(s_{t+1}, a_{t+1})$ the maximum reward that is attainable in the state following the current one.

a denotes the action that an agent chooses to perform

r denotes a reward function, mapping the action

α denotes the learning rate, set between 0 and 1.

γ denotes discount rate, set between 0 and 1.

4. Design of Personalization Travel Support Engine

This goal-oriented approach is to exploring personal interest by maximizing the reward to the item which user concerns and awarding the penalty to the items that user does not concern.

Environment (state): Travel List which users can select

Agent: Agent records data from user behavior on clicking and reading on the web sites. Then it analyzes users' requirement, and gives reward and penalty

Action: Filtering the travel list according to the agent's analysis.

Reward: Reward a value for the state that a user selects to perform.

Then engine offers a trip information to determine the user's interest and records the interactions and behaviors from the last exploration including clicking characteristics in browsing travel information.

4.1. Personalization Travel Support Engine Structure

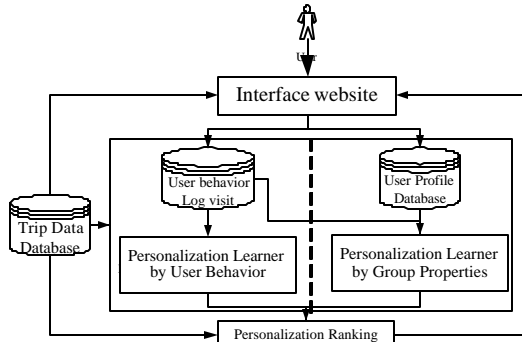


Figure 2. Personalization Travel Support Engine Structure

The Personalization Travel Support Engine Structure is composed of the followings

Interface Web Site. This is the part that the users see when viewing any website. It records the information that the web users always visit, analyzes the user behavior from such visit and offers the trip information that matches the user's unique requirement.

Personalization Ranking. Its function is to rank the trip information for the web users. The work process of this element is based on the initial weight of learning and the user's concern. Personalization Ranking is the part that will select trip list from the similar sequence pattern of the personal data.

Personalization Learner is the process of learning and analyzing of website usage behavior to understand user's interest. Clicking and reading habit will be evaluated to maximize the reward or award penalty and studied in creating ranking.

User Profile Database. This is the database of web user, which is to be used in the stage of travel management. Depending on the user's behavior, the database will help mapping the trip list to the user's requirement. Profile database is categorized into two types:

- User's properties data
- User's behavior

4.2. Personalization Learner

To perceive individual user's interest, one has to study user's behavior by means of the information from Interface Web Site that records two categories of data.

1. Web user profile includes user name, age, and sex
2. Traveling Information includes identification number, duration, categories, trip lowest price, trip highest price and destination country.

There are two learning approaches using in this study: personalization learner by group properties and by user behavior.

Personalization Learner by Group Properties is learning from all users in one group to find the group

interests of travel information by using given data on user age and gender.

Personalization Learner by User Behavior: records the data will be analyzed and the travel information to find the unique interest of each web user. Reinforcement learning algorithm, called Q Learning is adopted at this stage.

Q Learning helps maximizing a reward to the item on the list which is clicked and awarding penalty to the item that is not clicked, as shown in Eq. (1), which is learned the interest trip based on the both Personalization Learner.

$$\hat{Q}(s_t, a_t) \leftarrow a \left[r + \gamma \max_{a_{t+1}} \hat{Q}(s_{t+1}, a_{t+1}) \right] \quad (1)$$

Whereas max Q is defined as:

if user clicks the provided trip information

$-1/n$ if user doesn't click the trip information, when n is number of trips per page

$1/p$ in case of the trip information that is not recommended, when p is total number of trips

Also a is the learning rate valued at 0.2

γ is the discount rate valued at 0.8

4.3. Personalization Ranking

The display area for Personalization Ranking was divided into two parts. Part one is main box. When a user explores a website to find any travel information, the engine will rank the trip by using reinforcement theory and given data from group properties, fundamental data that the all user registers such as ages and genders and historical data when visiting the websites.

Part two is Recommend Box. When a user explores a website to find any travel information, the engine will display trip information randomly at the first visit. After that it will display travel information which has been analysed, and learned from historical user transactions, and trip database. The travel information which is top five ranking will be offered on the web page.

The ranking score is evaluated from the equations:

$$Q_r = Q_t + Q_{xp} + Q_{mp} + Q_c + Q_d \quad (2)$$

Q_r = total score for each user transaction

Q_t = duration of each trip (days)

Q_{xp} = maximum trip price (bahts)

Q_{mp} = maximum trip price (bahts)

Q_c = trip category (art & craft, diving, eco tours)

Q_d = trip destination (country name)

Then Q_t , Q_{xp} , Q_{mp} , Q_c and Q_d are calculated by using input data from user transactions on exploring PTS web site and Q learning equation. After that the total score (Q_r) is the summation of Q_t , Q_{xp} , Q_{mp} , Q_c and Q_d . Next Q_r score from each transaction is ranked in descending order. The five maximum scores

are selected and recommended to the users. This process is named learning by user behavior.

For the second approach is learning by group property or clustering users by age and name. The ranking of trip provided to users is depended on user profile and user behavior or web surfing transaction. In this approach users are clustered into group by using age and gender. Then, the value of interesting trip in each group is calculated by using user behavior or transaction on PTS web site. The process of trip ranking in this approach is the same as the above paragraph.



Figure 3 travel information provided after learning.

The user interface for the web site consists of two parts: Part 1 is Main Box which offers the travel information appropriate for users clustering by age and gender. Part 2 is Recommend Box which offers the top rank travel information appropriate with usage behavior.

4.4. Process of Personalization Travel Support Engine

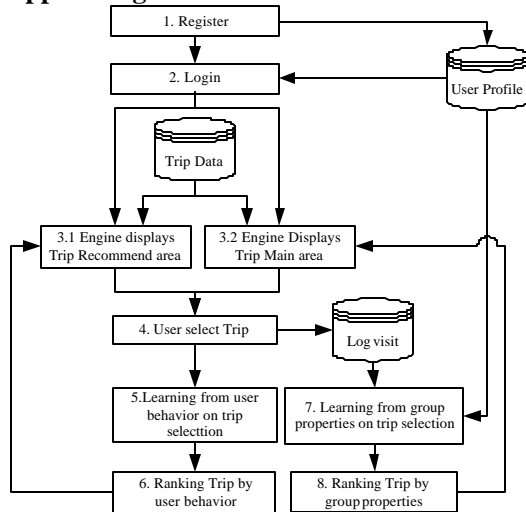


Figure 4 Process of Personalization Travel Support Engine.

Personalization Travel Support proposed eight stages engine in figure 4.

1. User registers and creates username and password.
2. User logs in to the website.

3. Engine displays information to user in two parts:

- 3.1 Engine offers a list of trip information at the Recommend Box. For initial state, the trip information is provided randomly.
- 3.2 Engine offers a list of trip information at the Main Box. This provided information given by PTS engine. This engine recommends trip by using data from group properties and from user transaction while surfing on PTS web site.

4. User can select the trip information in both the Main Box and the Recommend Box. Then PTS engine keeps a record of user behavior on the database.

5. Engine uses Personalization Learner to learn from user behavior and to find user interests on travel information. Personalization Learner calculation is based on a user behavior together with Equation 1. The ranking score for each transaction is calculated by using Equation 2. The results are showed on table 3.

6. Engine ranks the trip rewarded from process 5 from descending order of score. Then, the information of top five trips is provided at the Recommend Box.

7. Engine uses Personalization Learner to learn from user data clustering by Group Properties (age and gender) and find the travel information for user clustered. Learning is processing all day. At midnight, Personalization Learner is computed based on a group properties with Equation(1). The ranking score for each transaction is calculated by using Equation 2. The results are showed in table 4.

8. Engine ranks the trip reward from process 7 and lists the top five trip information to offer at the Main Box.

5. Experimental Results

This experiment describes the prototype of the personalization support engine which is implemented for recording and analysing the user interactions and behaviors. Then this engine presents and recommends interesting trips to user. User profile includes user name, age and gender. The trip list include Categories (art and culture, diving, shopping, ...and eco tour), Country (Thailand, Nepal, China), Duration (3, 4, 5 days), Minimal Price (400 bahts), and Maximal Price (10000 bahts).

The prototype of engine is implemented on this study which are approximately 50 trips. In each transaction, PTS automatically offers five trips in Recommend Box and 10 trips in Main box on one page. In this experiment there is 118 participants divided into 73 males and 35 females. The participants are undergraduate students who registered in one

subject Principle of Information Systems and/or subject Internet application for Commerce at a public University in Thailand. The participants are clustered by gender and age (group property) into 7 groups as showed in Table 1. Group 1, members are males who are younger than 15 years, there are three persons in this group. Group 2 members are males who are between 15 and 20 years old, there are 36 persons. Last Group 7, members are females who are between 25-30 years old, there are 3 persons.

AgeGender	Male	Female
< 15	3	
15 - 20	36	17
21 - 25	31	25
25-30	3	3
	73	45

Table 1. Number of user grouping by Age and Gender

ID	Trip Name	duration	Min price	Max price	Categories	Country
26	Mae Sot Package	3	4990	4990	Eco Tours	Thailand
27	Loei Package	3	3990	3990	Eco Tours	Thailand
29	Kanchanaburi Night Safari	3	3800	4300	Trekking	Thailand
43	Thais Travel Thai PKG	3	400	2300	Trekking	Thailand
34	Khao Yai Package Tour	3	4500	5500	Eco Tours	Thailand
32	Discovery Pattaya Package	3	2650	2950	Shopping	Thailand
37	Kanchanaburi Good Health	2	3950	4600	Health	Thailand

Table 2. List of Trip Information

Table 2 shows a list of trip information. Each trip record includes trip id, name, duration, minimum price, maximum price, categories, and country.

Rank	ID	Trip Name	Duration (Qr)	Min price (Qmp)	Max price (Qmp)	Category (Qc)	Country (Qd)	Total Score (Qr)
1	47	Thai Gulf Koh Tao-Koh Nang Yuan	0.410	0.100	0.522	0.001	0.410	1.421
2	51	Rafting Lum Num Kae Kaengsong waterfall	0.001	0.410	0.522	0.100	0.410	1.398
3	58	Mo Koh Surin	0.190	0.100	0.522	0.100	0.410	1.300
4	32	Discovery Pattaya Package	0.001	0.410	0.522	0.001	0.410	1.299
5	49	Mo Koh Surin	0.190	0.100	0.522	0.001	0.410	1.201
6	26	Mae Sot Package	0.001	0.100	0.522	0.001	0.410	1.001
7	27	Loei Package	0.001	0.100	0.522	0.001	0.410	1.001
8	29	Kanchanaburi Night Safari tour	0.001	0.100	0.522	0.001	0.410	1.001
9	37	Kanchanaburi Good Health	0.001	0.100	0.522	0.001	0.410	1.001
10	50	Rafting Khong Hin Boeng	0.001	0.001	0.522	0.100	0.410	0.990

Table 3. The ranking values of trip calculated by using user transactions as input data of Q-learning equation. (n=1)

Table 3 shows PTS analysis for one user. After learning from user transactions, the first rank ID 47 Thai Gulf-Koh Tao-Koh Nang Yuan-chumphon which its Duration 4 days is 0.410, Minimal Price 4500 bahts is 0.100, Maximal Price 4500 bahts is 0.522, Categories Beach Holiday is 0.001 and Country Thailand is 0.410. Total value is 1.421. This trip will be recommended to user firstly.

Rank	ID	Trip Name	Duration (Qr)	Min price (Qmp)	Max price (Qmp)	Category (Qc)	Country (Qd)	Total Score (Qr)
1	26	Mae Sot Package 3 days 2 nights	0.833	0.431	0.815	0.190	0.449	2.811
2	27	Loei Package 3 days 2 nights	0.833	0.522	0.815	0.190	0.449	2.481
3	29	Kanchanaburi Night Safari Tour	0.833	0.522	0.815	0.100	0.449	2.414
4	43	Thais Travel Thai PKG	0.833	0.001	0.815	0.100	0.449	2.084
5	57	Thai Gulf Koh Koh - Surfing	0.100	0.431	0.815	0.100	0.449	2.003
6	34	KHAO YAI PACKAGE TOUR	0.833	0.431	0.001	0.190	0.449	1.984
7	32	Discovery Pattaya Package	0.833	0.001	0.815	0.001	0.449	1.948
8	47	Thai Gulf Koh Tao-Koh Nang Yuan	0.100	0.431	0.815	0.001	0.449	1.934
9	58	Mo Koh Surin	0.001	0.431	0.815	0.100	0.449	1.893
10	54	Phu Luang - Phu Phan	0.100	0.522	0.815	0.100	0.449	1.832

Table 4. The Rank scores for Trip Learning by using Group Properties as input data of using Q-learning equation. (n=31).

Table 4 shows an example of the best ranked trip list by user profile clustering in one group who are males and between 21 – 25 years with 31 members. After learning, the first rank ID 26, Mae Sot Package, duration of trip-3 days is scored about 0.83, Minimal Price 4990 bahts is 0.651, Maximal Price 4990 bahts is 0.815, Eco Tours is 0.190 and Thailand 0.469. The total score is 2.811.

6. Conclusions

In this paper, recommended products are carefully displayed based on the user behavior and group properties of users. The engine starts learning from user profile, trip database and user historical transactions in accessing PTS web site. The learning process is using a Qlearning equation which is based on the reinforcement theory. An on-line personalized travel support engine is developed to assist in recommending trips for tourist industry. The system consists of three databases: (1) user profile, (2) trip database, and (3) user transactions. The main concept of the system is that users can surf on the PTS web site to find out interesting trips. The recommendation is based on user behaviors, and interests together with user profile. Then the top five trips are suggested for users after all candidate trips are ranked in terms of multiple criteria, these trips may be dynamically changed according to user behavior on PTS sites. With recommended trips based on significant data of user surfing and profile, it has the potential to increase the success rate of promotion, and user acceptance, and loyalty to the tourist industry.

Focusing on user's interest give the satisfied result since the information offered to the user is based on the statistics. Travel information can be offered to target users increasingly. The advantages of Reinforcement Learning Algorithm is easily understandable and implemented since there is no

need to find the best travel list but it can offer the most appropriate information at the point of time.

From this prototype, it can be expanded to the real online shops which enterprises can recommend interesting trips of users by personalized e-mail marketing for new trip or product promotions. Enterprises can increase numbers of sales and services growth through these personalized marketing.

7. References

- [1] C. Vassiliou, D. Stamoulis, and M. Drakoulis. "The process of personalizing web content: techniques, workflow and evaluation," *Proceedings of the SSGRR 2002w International Conference on Advances in Infrastructure for e-Business, e-Education, e-Science, and e-Medicine on the Internet*, 2002.
- [2] H. Lieberman, "Letizia: An agent that assist web browsing," *Proceedings of International Joint Conference on Artificial Intelligence*, pp.475-480, 1995.
- [3] R.S. Sutton, and A.G. Barto, *Reinforcement Learning: An Introduction*, MIT Press, Cambridge, 1998.
- [4] S. T. Yuan, "A personalized and integrative comparison-shopping engine and its Applications," *Decision Support Systems*, pp.139-156, January, 2003,
- [5] S.W. Changchien, Lee Chin-Feng, and Hsu Yu-Jung. "On-line personalized sales promotion in electronic commerce,". *Expert Systems with Applications*, pp.35-52, 2004
- [6] T. Joachims, D. Freitag, and T.M. Mitchell. "WebWatcher: A tour guide for the World Wide Web," *Proceedings of International Joint Conference on Artificial Intelligence*, pp. 770-775, 1997.
- [7] T. M. Mitchell. *Machine Learning*, The McGraw-Hill Companies, Inc., New York, 1997.
- [8] V. Galant, and M. Paprzycki. "Information Personalization in an Internet Based Travel Support System," *Proceedings of Business Information Systems 2002*, pp.191-202, 2002,
- [9] Y. W. Seo, and B. T. Zhang. "Personalized Web-Document Filtering Using Reinforcement Learning," *Applied Artificial Intelligence*, pp.665-685, 2001.
- [10] Z. Benyu, L. Wenxin, and X. Zhuoqun, "Personalized Tour Planning System Based on User Interest Analysis," *Proceedings of Business Information Systems 2002*, 2002.