

## Applying Data Mining to Analyze Travel Pattern in Searching Travel Destination Choices

<sup>1</sup>Pairaya Juwattanasamran <sup>2</sup>, Sarawut Supattranuwong , <sup>3</sup>, Sukree Sinthupinyo

<sup>1</sup>Technopreneurship and Innovation Management Program, Chulalongkorn University, Thailand

<sup>2,3</sup>Software Engineering Program, Department of Computer Engineering, Chulalongkorn University, Thailand

---

### Abstract

*The purpose of this research is to find a traveler's interest extracted from search behavior when the traveler searches for tourism destination. This research focuses on the travelers who use mobile devices as a tool to search for travel destination choices such as accommodation, tourist attraction, things to do, restaurant and souvenir shop. We applied data mining method and association rules technique to analyze the relationship between travelers' profile and their transactions. Knowledge discovered from database can be used as a rule set which provides travel information for travelers via mobile. Our framework is designed as a knowledge incremental learning. The paper demonstrates that applying data mining with tourism sector can increase opportunity for the competitive operations of tourism firm to respond the travelers' demand effectively.*

**Keywords :** Data Mining, Association Rule, Travel Information Searching, Destination Choices.

---

Date Of Submission: 10 April 2013



Date Of Publication: 30, April. 2013

---

## I. INTRODUCTION

The tourism industries have widely adopted information technology (IT) to enhance their operation efficiency and improve service quality and customer experience. Tourism firms often use the Internet as a channel of communication with the targeted customers [1], [2] because Internet can be easily accessible, cheaper and friendly to user [1]. Internet usages on the personal computers (PCs) are shifted to smartphones which are capable and more versatile. Traveler behaviors tend to change when new technologies come. They will plan less and search for information at the point of activities. Previously, they would search information from PC then switch to search on a mobile device instead because it is more flexible and convenient. This is consistent with a study by Hyde (2000) [3] indicated that traveler avoids vacation planning because flexibility of action and experiencing the unknown are essential amongst the hedonic experience they are seeking.

In the tourism sector, it is advantage to understand customers' needs to respond quickly to them with adequate offers – no matter in the online or the offline business. Innovation and Information Technology, these factors support an enhancing organizational performance [4]. This information is used to plan marketing strategy such as new product development which is appropriate to consumers, pricing, and public relations. The purpose of this paper is to show the method that applies data mining and association rules technique to analyze the rules of relationship between travelers' profile and their transactions. The tourism firm will be able to define marketing strategies and provide more affordable products and services for tourists who have various lifestyles appropriately using data mining as a tool to accumulate information and learned behavior in real-time.

## II. LITERATURE REVIEW

### A. Data Mining

Data mining is defined as a business process for exploring large amounts of data to discover meaningful patterns and rules [5]. Data mining is an information filter process in a large database using machine learning, statistical methods, mathematics, and, database method [6] which can result in improvements in the understanding and use of the data. The data mining system discovers patterns and relationships hidden in data [7], and actually is a part of a larger process called “knowledge discovery” which describes the steps that must be taken to ensure meaningful results. Fig. 1 illustrates main components of our prototype for knowledge discovery in databases. The core of the system is the discovery method, which computes and evaluates patterns on the way to be transformed into knowledge. The input to the discovery method includes raw data from the database, information from the data dictionary, additional domain

knowledge, and a set of user defined biases that provide high-level focus. The discovered knowledge can be sent directly to the user or back into the system as new domain knowledge.

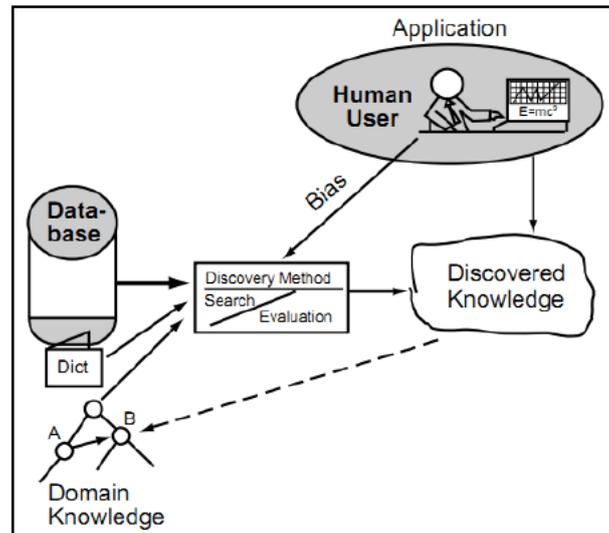


Figure 1. A Framework for Knowledge Discovery in Database [8]

This Knowledge Discovery in Databases (KDD) process consists of a sequence of the following steps [9].

- [1] Data cleaning – to remove noise and irrelevant data
- [2] Data integration – where multiple data sources are combined
- [3] Data selection – for retrieving from the database only the relevant data for the analysis
- [4] Data transformation – where data are transformed or consolidated into appropriate forms for mining
- [5] Data mining – the phase where the algorithms are applied in order to extract data patterns
- [6] Pattern evaluation – to find the interesting patterns which represents new knowledge
- [7] Knowledge presentation – when the visualization techniques are used to present the mined knowledge to the user

Data mining techniques such as association rules, clustering, decision trees etc. have been widely used for successfully segmenting and targeting customers across various industries. It provides an effective approach to discover and understand patterns in customer behavior thereby helping the decision maker to better group customers [10]. The on-line travel agency has high growth rate over the past ten years. It is estimated that more than 50% of all travel bookings happen on-line in the USA and Europe. The migration towards on-line travel agency continually grows and large online agencies like booking.com, hotel.com, expedia.com, etc. are expanding fast to supply to the emerging demand.

The related journals which applied data mining to improve service quality were showed in Emel [11]. The paper indicated that the better manager of a Destination Marketing Organization (DMO) understands traveler profiles and traveling patterns, the better they can market their destination. By using association rule mining, tourism organizations can identify different types of tourist profiling behavior. Wong, Chen, Chung, and Kao (2006) [12] adopted data mining techniques to analyze the travel patterns of Northern Taiwan Travelers and suggested that DMOs in Asian countries should promote their destinations in Taiwan.

In the tourism industry, knowing guests - where they are from, how much they spend, and when and on what they spend it- can help a company to formulate marketing strategies and maximize profits. Due to technological development, touristic companies have accumulated large amounts of customer data, which can be organized and integrated in databases that can be used to guide marketing decision [13]. Since identification of important variables and relationships located in these consumers -information systems could be a difficult task, some companies have attempted to raise the power of information by using data mining technologies. For example in hospitality area, the information systems have been used to assist the delivery of hospitality services. Some of key ways are [14]: improved capacity management and operations efficiency, central room inventory control, last room available information, yielding management capability, marketing, sales and operational reports, tracking frequency flyers and repeat hotel guests, internal management of operations from

transactions to human resources. Most of the items on the above list apply only to hotels and accommodation providers. In order to make high-quality marketing research and planning, data mining technology allows hotel companies to predict consumer behavior trends, which are potentially useful for marketing applications.

### **B. Consumer Behavior**

Consumer behavior refers to the problem recognition, searching, selection, purchase and consumption of merchandises and services for the satisfaction of their needs. There are different processes involved in the consumer behavior. In the marketing study, marketers need to understand traveler behavior to determine why customers consume or do not consume a product. The tourism firm will succeed if they can identify customer needs and respond their demands. [4]Kotler (1999) [15] explained the consumer behavior by S-R Theory shown in a model of consumer behavior. The model shows that globalization has changed tourist consumer behavior as it has the capacity to create impacts on 1) cultural criteria (culture, subculture, social class), 2) social criteria (reference groups, family, roles and status), 3) personal criteria (age and life cycle stage, occupation, economic circumstances, lifestyle, personality and self-concept), 4) psychological criteria (motivation, perception, learning, beliefs and attitudes). Smith (1977) [16] and King and Hyde (1989) [17] indicated that factors impacted by globalization dynamic, psychological factor of the tourists are considered to be the most important as it directly involves tourist consumer behavior. Smith (1977) [16] and King and Hyde (1989) [17] have explained classifications of persons who travel, which are proved to be very useful for tourism planning and marketing [15].Buhalis (2002) [14], said that the new generation travelers are more complex and highly demanding on quality of products. These travelers have known very well about attractions and tourism products. They had many experiences to spend time and money to travel. The new travelers like to compare details of products and choose the suitable items for themselves and they use the internet to search for information by themselves more than asking agency for services [14].Hyde (2000) [3] mentioned new generation travelers' behavior that they liked to plan a travel trip by themselves and wanted to face the enjoyable and exciting situation which is unpredictable rather than traveled following their plan, so searching travel information step in customer's decision making process is not necessary to be put into account before making a decision all the time.The new travelers are not just passive consumers anymore. They do not need advice from the agency and waiting for a long time anymore. They find themselves proactive and involve in decision making process of purchasing the travel products. The nature of the decision making to purchase those services must be flexible to be selected [18].

### **C. Innovation**

Innovation has become an important role in service sector [19]. The concept of Dorf and Byers (2008) [20] indicated that businesses can create a competitive advantage. Organizations should pay more attention to innovation or the ability for innovation which innovation can play an important role in the tourism industry as well [21]. Innovation in term of economics is meaningful to new products, new processes, new markets, acquisition for new sources and operation organization in new ways [22]. Innovation is about bringing new ideas transformed to new products, processes and services which create value. In the view of Amabile [23], creativity is the production of novel and useful ideas in any domain, In order to be considered creativity, a product or an idea must be different from what has been done before. (Few creativity theorists hold the strong position that a creative idea must be completely unique.) But the product or idea cannot be merely different for difference's sake; it must also be appropriate to the goal at hand, correct, valuable, or expressive of meaning. Innovation is the successful implementation of creative ideas within an organization. Rogers [24] defined an innovation is an idea, practice, or object that is perceived as new by an individual or other unit of adoption. It matters little whether the idea is "objectively" new as measured by the lapse of time since its first use or discovery. The perceived newness of the idea for the individual determines his or her reaction to it. If the idea seems new to the individual, it is an innovation. For the purposes of this paper, innovation is defined broadly as an organization's development and implementation of new products and services or new ways of doing things.

Literature review on innovation in tourism sector found that the tourism industry rather has not been very innovative while innovation is an important driven factor to increase the competitiveness [18]. In general, tourism research is likely to use qualitative research rather than quantitative research because it can reflect the phenomena observed and explained various aspects more than quantitative which has less flexibility. However, the studies of innovation in tourism also run quantitative research in parallel. The mainstream of the research is to find the intensity of innovation compared to other industries or compared with other countries or internationality [25].

Later on, innovation has been created into new products to the travel industry which internet is a key factor to drive the tourism economy. From the past used e-commerce to Radio Frequency Identification (RFID), Location Based Services (LBS), Web 2.0, mobile application which is a tool to link relationships between customers and tourism organizations [26], [27], [28]. Several studies showed that the application of information and communication technology (ICT) has changed the companies and tourism organizations. For the scope of this paper, we apply a data mining method to a system for providing travel information on mobile device which uses association rules technique to select the best item for users. The diffusion of this system innovation enables consumers to interact directly with the tourism service providers, leading to the reduction of the transaction costs. This is also a process which supports new service and new process via new product on mobile application forms.

### III. METHODOLOGY

We use a quantitative research which uses questionnaire as a research tool to collect data on Thai travelers aged between 18 to 40 years who live in Bangkok, Thailand and have ability to pay and make a decision independently by themselves. Sample size is 2,000 units. The questionnaire provides a content validity and reliability by collecting information from journals and theories in the field of travelers' behavior, market segmentation, and application of technology for travel information searching to design questionnaire. The questionnaire contained tourism products and services searching behavior case study in Chiang Mai Province. The variables are shown in table 1.

**Table 1.** Tourism Products and Services Searching Behavior Variables

Main Category	Tourism Products and Services Searching Behavior Variables
1. Accommodation	Accommodation Type: Hotel, Resort, B&B, Service Apartment etc. Accommodation Price (Baht/room/night): Below 500 B., 501-1,000 B., etc. Star Rating: 1 star, 2 stars, 3 stars, 4 stars, 5 stars Amenities: Restaurant in Accommodation, Pool, Free Wifi etc.
2. Tourist Attraction	Tourist Attraction Type: Palace, Museum, Temple, Natural Park etc. District Area: Chiang Mai City, Chiang Dao etc. Festival/Events: Song Kran Festival, Yi Peng Floating Lantern etc.
3. Thing to Do	Mountain Bike, Rafting, Massage/SPA, Elephant Riding etc.

We collected travelers' profile and behavior from questionnaire and executed those transactions by Rapiminer program with association rule technique. The result generated relationship rules of traveler behavior which called knowledge discovery in database (KDD).

The fig. 2 shows the framework of traveler behavior collecting and travel recommender innovative system

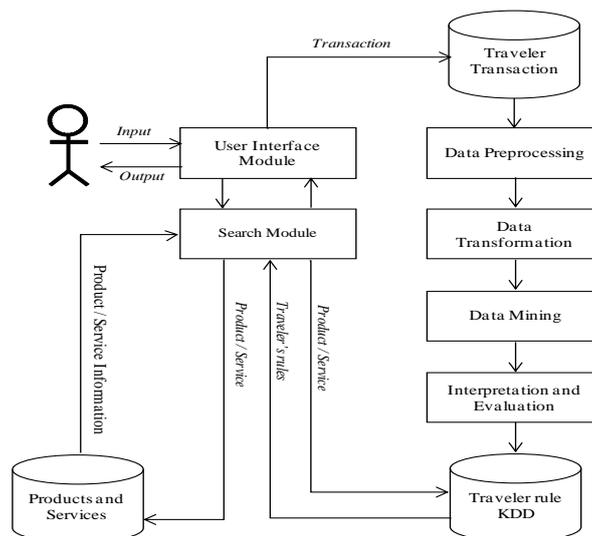


Figure 2. The Traveler Behavior Collecting and Travel Recommender Innovative System

We used the questionnaires as a tool to collect data. Then we preprocessed, transformed, and analyzed data by data mining method with association rule technique. The results were the traveler's searching pattern. The KDD is capable to be integrated with destination marketing organization (DMO) to create a mobile application which plugs the KDD into the server to calculate the traveler pattern and recommend the appropriate alternative destination choices to users.

IV. RESULT

We used RapiMiner program with training set of 2,000 transactions. We found 78 rules, three of which are shown in this paper as an example. Those 78 rules had minimum support 0.0230 maximum support 0.490 minimum confidence 0.211 and maximum confidence 1.000. The results are shown in table 2.

Table 2. Association Rule of Travelers' Behavior in searching travel information

No.	Premise	Conclusion	Support	Confident
1	Accommodation Price = 501-1,000 Baht/room/night, Accommodation Amenity = Restaurant in Hotel, Star Rating = 3 stars	Accommodation Type = Hotel	0.0335	0.9054
2	Attraction Type = Market	Festival/Events = Song Kran Festival	0.0230	0.4647
3	Facilities in Restaurant = Suitable for Taking Picture, Food Type = Thai Food, Restaurant Style = Family	Meal = Dinner	0.0275	1.0000

The result from 2,000 questionnaires found that when users searched for travel destination from 3 main categories (accommodation, tourist attraction and restaurant) in table 1, they had alternative to choose factors from variables which we determined and the sample obtained 3 association rules are shown below Rule 1 {Accommodation Price = 501-1,000 Baht/room/night, Accommodation Amenity = Restaurant in Hotel, Star Rating = 3 stars} => {Accommodation Type = Hotel} if a user chooses accommodation price between 501 – 1,000 Baht/room/night, restaurant in an accommodation, and accommodation 3 stars then that user will choose hotel at confident level 90.54%. Rule 2 {Attraction Type = Market} => {Festival/Events = Song Kran Festival} if a user choose market in attraction type then that user will choose the tourist attraction which concern with Song Kran Festival at confident level 46.47% Rule 3 {Facilities in Restaurant = Suitable for Taking Picture, Food Type = Thai Food, Restaurant Style = Family} => {Meal = Dinner} if a user chooses restaurant which suitable for taking picture and Thai food and restaurant in family style then that user will choose restaurant for dinner at confident level 100%. Those 78 rules were contained in database. We designed and developed the user interface of a web application and mobile application which connects to the server. Before travelers use these application they must register his/her information on web application, the system will collect their profile and analyze their travel pattern by association rule in database.

When traveler search for destination choices on mobile application he/she needs to log in to verify his/her authenticity (Fig 3.) User can search for destination choices by three categories, accommodation, attraction and restaurant. We illustrate the applying of rule 1 in Fig 4 – 7. If a user chooses accommodation price between 501 – 1,000 Baht/room/night, restaurant in an accommodation, and accommodation 3 stars the system will generate the general results of those criteria on the top of mobile page (Fig 6) and the system also generate the recommended destination choices by system shown below. The criteria will match to premise in table 2 and the recommend items will link to conclusion in table 2. Traveler will be given the alternative destination choices from travel pattern in database.



Figure 3. Log in page choices



Figure 4. Search by categories page



Figure 5. Accommodation



Figure 6. System recommended by rules in database

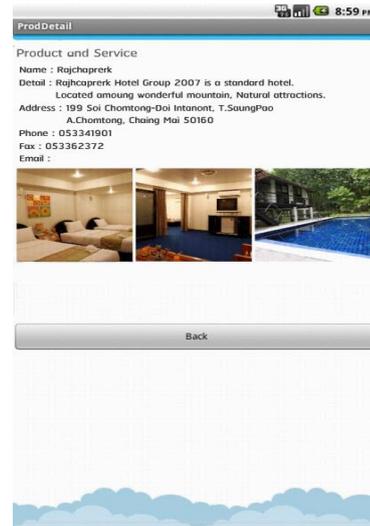


Figure 7. Detail of destination choice

### V. CONCLUSION

The rules in database is capable to utilize to be a database which is plugged into a server when the user registered to the system his/her profile will be collected in database and the system will select the suitable pattern for that user and when the user clicks to search any item, the system will deliver the alternative items which relates to the first selected choice (Fig 8). User will get many alternatives choices and suppliers like hotels, restaurants and tourist attractions also have more chance to promote their products and services.

#### Knowledge Incremental Learning Loop

1. Original Knowledge (from questionnaire)
2. User input new data
3. System:
  - 1) Data Collection
  - 2) Data Analysis
  - 3) Data Interpretation

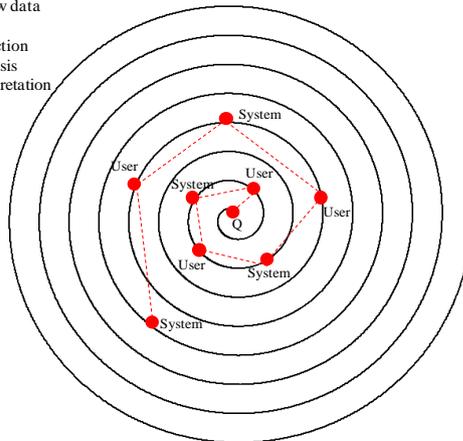


Figure 8. Knowledge Incremental Learning

The innovative system contributes new knowledge which gathers increasingly on database. In figure 8, Q is an original knowledge from questionnaires which are surveyed from 2,000 units. While new user is searching for travel information on mobile devices, the system will learn user behavior transaction which user clicks. The system will collect new data and analyze them then interpret to user. The system will learn increasingly when several users click more on mobile application. It will collect more data and repeatedly analyze the newly obtained data. This presents that the knowledge discovery in database will get bigger. Our framework implies to increase learning model. If the travelers' behavior changes, the pattern in database also changes. The same as tourism products and services changed, the knowledge in database will change to new rules as well. Therefore, the database developed in this study is capable to accumulate new knowledge and update all the time. The system will operate more accurate and work effectively along with the dynamics of the world.

## REFERENCES

- [1] Burke, R. R. (1997). "Do You See What I See? The Future of Virtual Shopping". *Journal of the Academy of Marketing Science*, 352-360.
- [2] Peterson, R. A., Balasubramian, S., & Bronenberg, B. J. (1997). "Exploring the Implications of the Internet for Consumer Marketing". *Journal of the Academy of Marketing Science*, 329-346.
- [3] Hyde, K. F. (2000) "Chapter10: A hedonic perspective on independent vacation planning, decision-making and behavior". Crouch, G. I.; Mazanec, J. A.; Oppermann, M.; Sakai, M. Y. *Consumer psychology of tourism, hospitality and leisure*. (pp 177-191).UK: CABI.
- [4] Anon, (1985). "AMA Board Approves New Marketing Definition," *Marketing Educator*, Spring, p. 1.
- [5] Michael, J.A., and Linoff, G. (2011). "Data Mining Techniques for Marketing, Sales, and Customer Relationship Management". Wiley Publishing.
- [6] Cabena, P., Hadjinian, Stadler, R., Verhees, and Zanasi, A. (1997). "Discovering Data Mining: From Concept to Implementation". Prentice Hall.
- [7] Edelstein H. "Data mining: exploiting the hidden trends in your data". *DB2 Online Magazine*. <http://www.db2mag.com/9701edel.htm>.
- [8] Frawley, W.J., Piatetsky-Shapiro, G., and Matheus, C.K. (1992). "Knowledge Discovery in Database: An Overview". *AI Magazine*: 57-70.
- [9] Wirth and Hipp, R. (2000). "Towards a Standard Process Model for Data Mining". In *Proceeding of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining*, Pages 29-39, Nanchester, UK
- [10] Hoontrakul, P. and Sahadev, S. (2005). "Application of Data Mining Techniques in On-Line Travel Industry: A case Study from Thailand". Website. [www.pongsak.hoontrakul.com](http://www.pongsak.hoontrakul.com)
- [11] Emel, G. G., Ta?kin, Ç., & Akat, Ö. (2007). "Profiling a Domestic Tourism Market by Means of Association Rule Mining". *Anatolia*, 18(2), 335-343.
- [12] Wong, J., Chen, H., Chung, P., & Kao, N. (2006). "Identifying Valuable Travelers and their Next Foreign Destination by the Application of Data Mining Techniques". *Asia Pacific Journal of Tourism Research*, 11(4), 355-373.
- [13] Danubianu, M. and Hapenciu, V. (2008). "Improving Customer Relationship Management In hotel industry by Data Mining Techniques", *Proceeding of "Competitiveness and Stability in the Knowledge-Based Economy"*, Vol: CD, 30-31 Mai, 2008, Craiova, Romania, ISSN/ISBN: 978-606-510-162-3, Page: 2444-2452
- [14] Buhalis, D. (2002). "eTourism: Information technologies for strategic tourism management". *Financial Times Prentice Hall*.
- [15] Kotler, P., Bowen, J. and Makens, J. (1999). "Marketing for Hospitality and Tourism", 2nd ed., Prentice Hall, USA
- [16] Smith, Valene., (1977). "Hosts and Guests: The Anthropology of Tourism", University of Pennsylvania Press, Philadelphia.
- [17] King, Brian and Hyde, G., (1989). "Tourism Marketing in Australia", Hospitality Press, Melbourne.
- [18] Poon, A. (1993). "Tourism, Technology and Competitive Strategies". New York: Cabi Publishing.
- [19] Miles, I. (2001). "Service Innovation: A Reconfiguration of Innovation Studies". University of Manchester.
- [20] Dorf, R. C. and Byers, T. H. (2008). "Technology Ventures: From Idea to Enterprise". New York: McGraw-Hill Companies Inc.
- [21] Hjalager, A. M. (2002). *Hospitality and Travel Marketing*. The United States: Delmar Thomson Learning.
- [22] Schumpeter, J. A. (1934). "The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest and the Business Cycle". Cambridge: Harvard University Press.
- [23] Amabile, T.M., Conti, R., Coon, H., Lazenby, J. and Herron, M. (1996). "Assessing the Work Environment for Creativity". *Academy of Management Journal*, 39: 1154-84.
- [24] Rogers, E. M. (1983). "Diffusion of Innovation". New York: Free Press.
- [25] Hall, C. M. (2009). "Innovation and Tourism Policy in Australia and New Zealand: Never the Twain Shall Meet? *Journal of Policy Research in Tourism*", *Leisure and Events*, 2-18.
- [26] Marcussen, C. H. (2009). "Trends in European Internet Distribution of Travel and Tourism Service". Called on 31 August 2011 from Research Centre of Bornholm Web Site: [http://www.cet.dk/UK/Staff/chm/P\\_CHM.htm](http://www.cet.dk/UK/Staff/chm/P_CHM.htm)
- [27] Weiermair, H. and Klein, S. (2005). "ICT-Enabled Innovation in Travel and Tourism. *Innovation and Product Development in Tourism*", 71-84.
- [28] Buhalis, D. and Law, R. (2008). "Progress in Information Technology and Tourism Management: 20 Years on and 10 Years after the Internet-The State of eTourism Research". *Tourism Management*, 609-623.