Matthias Fuchs
Francesco Ricci
Lorenzo Cantoni (Eds.)

Information and Communication Technologies in Tourism 2012

ENTER

SpringerWienNewYork

SpringerWienNewYork

Matthias Fuchs
Francesco Ricci
Lorenzo Cantoni (Eds.)

Information and Communication Technologies in Tourism 2012

Proceedings of the International Conference in Helsingborg, Sweden, January 25-27, 2012

SpringerWienNewYork

Prof. Matthias Fuchs
Faculty of Human Sciences
Mid Sweden University, Östersund, Sweden

Prof. Francesco Ricci
Faculty of Computer Science
Free University of Bozen-Bolzano, Bozen-Bolzano, Italy

Prof. Lorenzo Cantoni
Faculty of Communication Sciences
Università della Svizzera Italiana, Lugano, Switzerland

SpringerWienNewYork is part of Springer Science + Business Media
springer.at

Typesetting: Camera ready by the author
Printing: Strauss GmbH, 69509 Mörlenbach, Germany

Printed on acid-free and chlorine-free bleached paper

With 87 figures and 110 tables

SPIN: 86039483

CIP Data applied for

ISBN 978-3-7091-1141-3 SpringerWienNewYork

Preface

Inaugurated in 1993 in Innsbruck, the 1st ENTER conference was established by social and computer scientists in strong collaboration with the travel and tourism industry interested in these emerging and promising new technologies. Since then, the ENTER conference steadily matured into probably the premier global conference on information and communication technologies within the travel and tourism domain. The unchanged aim and unique purpose of the ENTER conference series is to bring together delegates from regional and national tourism organizations, practitioners from the travel and tourism industry, as well as researchers with various academic backgrounds. This cross-disciplinary and highly synergistic focus is clearly reflected by the three major pillars of the annual conference programme: the destination track, the industry track and the research track, respectively.

The 19th International Conference on Information and Communication Technologies in Travel and Tourism, ENTER@2012, is taking place in Helsingborg, Sweden. The conference theme "e-Tourism: Present and Future Services and Applications" underscores the role of information technologies in co-creating superior and lasting tourism experiences; clearly, an intended thematic coincidence in one of the home countries of the "Scandinavian School of Service Management". The focus of the ENTER's research track is to present, critically discuss and disseminate high quality research findings on new research areas, theory development, new methods and innovative applications in the field of information and communication technologies within the travel and tourism domain. By following this aim, during the last months a total of 81 papers were submitted to ENTER by scholars from prominent research institutes around the world with a disciplinary background in the fields of social or computer sciences, respectively. Interestingly, a steady trend in popularity of research contributions can be observed in the fields of mobile technologies and services as well as social media and user generated content. However, surely triggered by this years' conference theme, the second frequently submitted contributions cover the study areas of electronic tourism marketing and customer research in the e-tourism domain.

In order to gain highest research quality standards, full paper submissions were rigorously double-blind reviewed by at least three qualified reviewers. Prior to final acceptance, the authors of the 51 accepted full papers and research notes were requested to revise their submissions according to reviewers' recommendations. Finally, 44 full papers were included in the present volume. They advance the state of the art in social media and Web 2.0, electronic tourism marketing, web site and search engine optimization, technology adoption and diffusion, online communities, tourism management and decision support, eLearning, mobile technologies and applications in tourism, recommender systems, e-intermediaries and networks in tourism, customer research in e-tourism and user generated content.

For the first time this year a restricted number of highly qualified researchers in the study field of e-Tourism have been invited to present their latest results at ENTER@2012. Thus, we are happy to list here four teams of authors who kindly followed this invitation to submit a research paper:

1) Nicoletta Di Blas, and Paolo Paolini (eTourism: Multimedia Brochures and Communication Chains; p. 72)

2) Federica Cena, and the PIEMONTE Team (Interacting with a Social Web of Smart Objects for Enhancing Tourist Experiences; p. 179)
3) Dietmar Jannach, Fatih Gedikli, Zeynep Karakaya, and Oliver Juwig (Recommending Hotels based on Multi-Dimensional Customer Ratings; p. 320)
4) Yuji Yoshimura, Fabien Girardin, Juan Pablo Carrascal, Carlo Ratti, and Josep Blat (New Tools for Studying Visitor Behaviors in Museum: A Case Study at the Louvre; p. 391)

In addition to the full papers, 7 research notes reflecting emerging ideas and technology prototypes were accepted for electronic publication in *e-Review of Tourism Research* (e-RTR). Thus, our sincere thanks are given to all authors of the full papers, invited papers and research notes for their excellent contributions and presentations at ENTER@2012 in Helsingborg, Sweden. Moreover, the high research standard of these proceedings could only be achieved through the outstanding efforts and constructive feedback of the members of the Research Programme Review Committee. In addition, we express our thanks to the additional reviewers that were recruited to help the Programme Review Committee to select the contributions published in this volume. Their names are listed at page XIII. Finally, we acknowledge the excellent support of Tatiana Chekalina in compiling the proceedings.

Matthias Fuchs, Francesco Ricci, and Lorenzo Cantoni

Helsingborg, January 2012

Contents

5. Technology Adoption and Diffusion

6. Online Communities

7. Tourism Management, Decision Support and E-Learning

8. Mobile Technology and Services I

9. Mobile Technology and Services II

10. Recommender Systems in Tourism

11. E-Intermediaries and Networks in Tourism

12. Customer Research I

13. Customer Research II

14. User Generated Content I

15. User Generated Content II

Index of Authors

Research programme Review Committee

Nineteenth International Conference on Information and Communication Technologies in Travel and Tourism, Helsingborg, Sweden, January 25-27, 2012.

Norman AU, Hong Kong Polytechnic University, China
Rodolfo BAGGIO, Bocconi University, Italy
Francois BEDARD, L'Université du Québec à Montréal (UQAM), Canada
Srikanth BELDONA, University of Delaware, USA
Thomas BIEGER, University of St. Gallen, Switzerland
Davide BOLCHINI, Indiana University – Purdue University, USA
Dimitrios BUHALIS, Bournemouth University, UK
Evangelos CHRISTOU, University of the Aegean, Greece
Roberto DANIELE, Oxford Brookes University, UK
Astrid DICKINGER, MODUL University Vienna, Austria
Roman EGGER, Salzburg University of Applied Sciences, Austria
Christer ELDH, Lund University, Sweden
Daniel FESENMAIER, Temple University, USA
Andrew FREW, Queen Margaret University, UK
Robert GOECKE, Munich University of Applied Sciences, Germany
Robert GOVERS, University of Leuven, Belgium
Ulrike GRETZEL, University of Wollongong, Australia
Martin HITZ, University of Klagenfurt, Austria
Wolfram HOEPKEN, University of Applied Sciences Ravensburg-Weingarten, Germany
Alessandro INVERSINI, Universitá della Svizzera Italiana, Switzerland
Timothy JUNG, Manchester Metropolitan University, UK
Hidenori KAWAMURA, Hokkaido University, Japan
Stefan KLEIN, University of Münster, Germany
Yohei KURATA, Tokio Metropolitan University, Japan
Rob LAW, The Hong Kong Polytechnical University, Hong Kong
Andy LEE, The Hong Kong Polytechnical University, Hong Kong
Miranda LEE, University of New Haven, USA
Maria LEXHAGEN, Mid-Sweden University, Sweden
Gang LI, Deakin University, Australia
Berendien LUBBE, University of Pretoria, South Africa
Bernd LUDWIG, Erlangen University, Germany
Carl MARCUSSEN, Center for Regional and Tourism Research, Denmark
Christian MAURER, IMC Fachhochschule Krems, Austria
Michael MCGARTH, Victoria University, Australia
Luisa MICH, University of Trento, Italy
Juline MILLS,Purdue Univerisity, USA
Valeria MINGHETTI, CISET-Ca' Foscari University, Italy
Nina MISTILIS, University of New South Wales, Australia
Lim Wai MUN, University of Plymouth, UK
Jamie MURPHY, University of Western Australia, Australia
Niels Christian NIELSEN, University of Southern Denmark, Denmark
Elena NOT, FBK, Italy
Peter O'CONNOR, ESSEC Business School, France
Paolo PAOLINI, Politecnico di Milano, Italy
Bing PAN, College of Charleston, USA
Harald PECHLANER, Catholic University Eichstätt-Ingolstadt, Germany
Birgit PROELL, Linz University, Austria
Sofia REINO, Queen Margaret University College, UK

Additional Reviewers for ENTER 2012

Perceptions and Strategies of Hospitality and Tourism Practitioners on Social Media: An Exploratory Study

Julian K. Ayeh,
Daniel Leung,
Norman Au, and
Rob Law

School of Hotel and Tourism Management
The Hong Kong Polytechnic University, Hong Kong
{julian.ayeh, leung.daniel}@connect.polyu.hk,
{hmnorman, hmroblaw}@polyu.edu.hk

Abstract

The potential benefits of social media for the promotion of hospitality and tourism businesses have been well acknowledged in the existing literature. Most studies have, however, examined social media use and perceptions from consumers' perspective. There is limited research addressing how service providers perceive and employ social media to influence the consumer's decision making process. Using in-depth interviews with hospitality and tourism practitioners in Hong Kong, this study explores practitioners' perceptions regarding social media and examines the strategies being applied in engaging social media to influence consumers' decision making. The findings reveal that practitioners generally understand what social media represent and recognise their potential for business promotion as well as inherent challenges. Furthermore, they employ various social media applications and strategies in attempting to influence consumers' decision making process. The findings hold implications for both practitioners and researchers.

Keywords: Social media; Web 2.0; practitioners; consumers' decision making process.

1 Introduction

Over the past decade, the increasing popularity of social media has been changing the way businesses employ the Internet for marketing and reputation management. Traditionally, companies have the overwhelming control of online information because consumers were simply passive content users. With the emergence of social media, consumers can now actively utilise different platforms like consumer review sites, blogs and social networking sites, among others, to create online content and share experiences about products and services. This "participatory culture" characterises the social media phenomenon which can now considerably determine a firm's reputation, sales and even survival (Kietzmann, Hermkens, McCarthy, & Silvestre, 2011). Given that these platforms present practitioners with numerous opportunities to interact with a wide range of stakeholders including potential and existing customers, employees and the general public, the ongoing changes hold important implications for business organisations, especially for those in the hospitality and tourism industry, "which has for many years used the Web as both a marketing medium and a selling mechanism and must now adapt to its evolving characteristics" (O'Connor, 2010, p. 762).

In the hospitality and tourism literature, scholars have widely discussed the potential benefits of social media and have advocated for practitioners to leverage it for viral

marketing, reputation management, customer relationship management, new product development, among others (Schmollgruber, 2007; Zehrer, Crotts, & Magnini, 2011). In view of the clear advantages that social media offer, a number of hospitality and travel businesses have been embracing social media to enhance consumers' travel planning process in general and information search experience in particular (Fuchs, Scholochov, & Höpken, 2009; Sánchez-Franco & Rondan-Cataluña, 2010). Yet, researchers in the general management literature suggest that many practitioners are either reluctant or unable to fashion appropriate strategies and allocate resources accordingly to successfully engage with social media. Hence, some business organisations often ignore or mismanage the opportunities and threats offered by these strategic tools (Berthon, Pitt, McCarthy, & Kates, 2007; Kietzmann et al., 2011).

With the growing popularity of the social media phenomenon, it is important to understand how practitioners view social media and how their organisations are utilising this tool to impact the consumer decision making process. Nevertheless, even as the potential of social media and Web 2.0 technologies in general for promoting hospitality and tourism businesses have been widely recognised in existing literature (e.g., Fuchs, Scholochov, & Höpken, 2009; Sánchez-Franco & Rondan-Cataluña, 2010), there is limited prior research investigating how practitioners in this industry are actually using social media in relations to influencing the consumers' decision making process. By means of in-depth interviews with industry practitioners from the catering, hotel, and travel agency sectors in Hong Kong, this study explores practitioners' views on social media and examines the strategies being employed in utilising this tool to influence consumers' decision making. The rest of the paper is organised as follows. The next section reviews previous literature pertaining to social media and the travel planning process. This is followed by an explanation of the research methodology adopted in the current study. The subsequent section presents the findings and discussions. At last, conclusions and implications of the study are discussed before offering suggestions for future research.

2 Literature Review

2.1 Social Media in Hospitality and Tourism

The term social media is generally understood as Internet-based applications that contain consumer-generated content and facilitate a higher level of social interaction among travellers (Blackshaw, 2006). From a customers' viewpoint, the value of using social media lies in its richness on personal experiences and trustworthiness as electronic word-of-mouth (Litvin, Goldsmith, & Pan, 2007). The impact of user-generated content by social media on a traveller's decision making is well recognised in various market research reports. Compete Inc. (2006) found that one in three travel purchasers who access information on social media agreed that the information helped with their purchase decision. In another study conducted by comScore (2007), 84 per cent of travel review users reported that travel reviews had a significant influence on their purchasing decisions. A number of prior studies have also examined the effect of social media on travellers' information search behaviour. Gretzel and Yoo (2008) suggested that travel reviews on social media sites help add fun to the planning process and increase confidence for travellers' decision by reducing risk. Xiang and

Gretzel (2010) clearly showed the increasing importance of social media on travellers who are looking for travel products online. Based on their search engine analysis on Google, the findings demonstrated the ubiquity of social media websites in online travel information search.

Undoubtedly, social media currently represent a dynamic online travel information source that can greatly influence travel consumers. Cox, Burgess, Sellitto, and Buultjens (2009) even suggest that the influence of user-generated content and social media on travel purchases is similar in scope to the effect that traditional forms of travel sites may have. Despite the growing popularity of social media in the tourism context, the existing literature on social media from supplier's perspective is limited. Leung, Lee, and Law (2011) attempted to examine the diffusion of social media in the Hong Kong hotel industry. Based on the content analysis of 109 hotel websites and in-depth interviews with four hotel managers, the researchers identified a trend of increasing adoption of social media as well as the factors contributing to the uneven adoption between chain and independent hotels. Sigala (2011) employed a similar approach to examine the exploitation and diffusion of social media by Greek tourism professionals. However, the researchers found that the social media adoption rate was surprisingly low owing to their perceived risk. Milano, Baggio, and Piattelli (2011) also revealed the contribution of integrating Facebook into the total number of visits on a tourism website. These prior studies, albeit useful, failed to offer insights into how hospitality and tourism practitioners exploit social media in assisting customers to plan their travel and thereby generate more business for their respective companies.

2.2 Travel Information and Travel Planning Process

Understanding the travel decision making process of travellers has been one of the most researched areas in tourism and consumer behaviour studies. Though the models to explain the process vary among different sources, extant literatures tend to follow the classical five-stage consumer decision-making model (Engel, Blackwell, & Miniard, 1990). Adapting to the travel and tourism context, the process commences when a traveller recognises the need to travel. An aroused traveller may then search for and process travel information to arrive at a set of alternatives, followed by the actual travel behaviour and concluding with the post-trip evaluation (see Fig. 1).

With the ease of use, interactivity, and flexibility that web-based interfaces offer, customers have been increasingly using the Internet to search for travel information. However, Fodness and Murray (1997) found that travellers usually depend on personal experience or opinions from friends and relatives when planning their trips. Hence, in addition to searching online travel information from websites by tourism suppliers, intermediaries and other parties, the trend of searching consumer-generated content in social media among customers has been on the ascendancy in recent years (Cox et al., 2009). In response to this trend and to remain competitive in the marketplace, most hospitality and tourism organisations have embraced social media as a marketing and communication strategy (Fuchs et al., 2009; Leung et al., 2011).

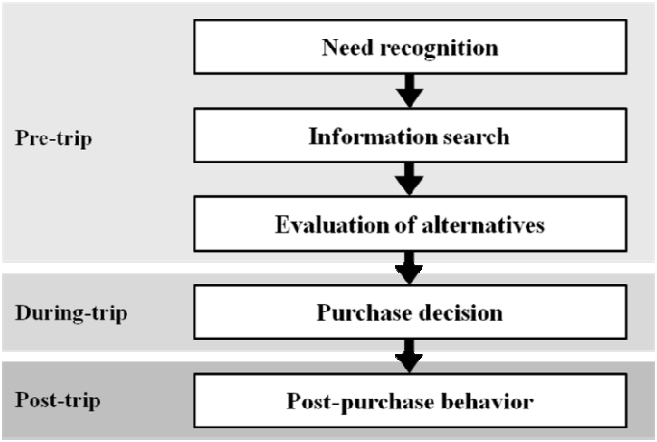


Fig. 1. Travel planning process
(Source: adapted from Engel, Blackwell, & Miniard, 1990, p. 482)

Sigala (2004; 2010) argued that practitioners must ensure that web applications provide adequate functional support to create value for their customers, and meet customers' need at every stage of their decision making process. In other words, the design and implementation of web applications as well as strategies should consider the values that customers seek. By the same token, an examination of the extent and ways practitioners engage social media to meet customers' need is essential but still under-investigated in the scholarly literature.

3 Research Design

3.1 Data Collection

To recap, this study aims to examine practitioners' perceptions on social media and to analyse the social media strategies being employed. In order to achieve these goals, this study adopted a descriptive research design and employed a qualitative approach on data collection. Specifically, in-depth interviews were conducted with five practitioners in the hospitality and tourism industry in Hong Kong. Semi-structured interviews were used because the approach ensured uniformity in asking questions. The approach was thus expected to help in arriving at a reasonable conclusion of the issues being investigated. A purposive sampling approach was employed in the selection of interviewees. Though the interviewees belonged to different hospitality and tourism-related businesses and held different positions, all of them were identified as the key persons in the design and implementation of social media marketing in their respective companies. These include the IT manager of a hotel (*Manager A*), the Director of Communication for another hotel (*Manager B*), the Senior Director of Marketing for a travel and tour company (*Manager C*), the General Manager for a wine bar and restaurant (*Manager D*) and an Account Executive in a travel and tour company (*Manager E*). The interviews were guided by interview schedules with open-ended questions. Each interview lasted for about 40-45 minutes, and the conversation was audio recorded and transcribed into word context.

3.2 Data Analysis

A content analysis technique similar to the one described by Miles and Huberman (1994) was employed in the analysis of the data collected. The audio-recorded data were firstly listened by the researchers to familiarise themselves with the data before transcribing. Starting with open coding, each researcher independently developed codes from the data using the inductive approach. Following the approach suggested by Strauss (1987), relationships between open codes were sought (axial coding) and some codes were later examined over and above others (selective coding). Afterwards, two researchers jointly discussed the coded data again. Three types of codes were subsequently developed – descriptive, interpretive and pattern (Miles & Huberman, 1994). Descriptive codes were firstly developed to label each section of the transcript. After developing descriptive codes, interpretive codes helped make some conclusions of the analysis. Pattern codes were finally developed to help identify themes and relationships in the data. The process of data analysis involved interpretations from the perspective of the interviewees on the theme of the study.

3.3 Data Trustworthiness and Limitations

Qualitative researchers are expected to demonstrate trustworthiness through internal validity, external validity, reliability and objectivity (Lincoln & Guba, 1985; Maxwell, 2005). To ensure the data obtained are credible and unbiased, a series of open-ended questions were asked to allow the interviewees to share their opinions and experiences on the issue under investigation without any interruption. In addition, the analysis and interpretation were done both individually and jointly by two researchers to ensure objectivity.

A major limitation of the study is the use of a single data collection approach – in-depth interviews. A triangulation approach could have improved trustworthiness (see Arksey & Knight, 1999). Nonetheless, this study did not aim to generalise the findings to the rest of the population, but to understand the phenomenon of interest from the respondents' perspective. Because meaning is subjective and differs from person-to-person, this study is still able to offer meaningful insights into the practitioners' perceptions and attitudes towards social media.

4 Research Findings and Discussion

To ensure clarity and enhance understanding of the subject under study, the findings are organised into three sub-sections: (1) General perceptions about social media; (2) Social media applications and strategies being employed; and (3) Discussion of the strategies in relations to the consumer's decision making process.

4.1 General Perceptions about Social Media

A primary question of the interview was to elicit practitioners' views on social media. The responses given by the practitioners can be considered from three perspectives – what it is, what it can be used for, and the challenge it presents. Generally, the practitioners considered social media as an information sharing platform where customers help themselves. Consumers often share their experiences with travel products and services on social media platforms. These tend to serve as a useful and

credible source of information for other travellers. *Manager D* from a restaurant described social media as "the modern-day word-of-mouth". Thus, practitioners are aware of the role of social media as a platform for sharing information among consumers. As social media represent a strategic tool for marketing and business generation, respondents also looked at social media from the perspective of what it can do for their businesses. Three key capabilities of social media were identified. First, social media was perceived as a "new channel" for promoting their products and services. During the interviews, phrases such as "a promotional medium", "a 'hip' method to arouse customers' attention", and "a new channel of marketing for our company to use" were used by practitioners to describe social media. Besides serving as a promotional channel, the practitioners perceived social media as a tool for connecting and building relationships with existing customers. They regarded this form of communication with customers via social media is casual and informal. Under such conditions, customers are more willing to share their experiences so that companies may understand what customers really like and dislike. In addition, the provision of exclusive offerings to customers in their online communities helps retain and build up long-term relationships with customers. Finally, the practitioners recognised the ability of social media in public relationship management, particularly for reputation management. Social media can be used for damage control because companies may take this opportunity to acknowledge, reply and even offer compensation to customers as a service recovery strategy.

While the benefits of social media are widely recognised by practitioners, they also perceived social media as a critical challenge to their businesses. Various reasons were given for this perspective. Since content posted on this platform is no longer controlled by companies, practitioners described social media as uncontrollable and difficult to monitor the quality of its content. Also, the fear of receiving negative reviews or comments from customers seems to determine the attitudes of some managers towards social media. For instance, one respondent from a travel and tour company was of the view that the top management of his organisation would be reluctant to allow the company to fully leverage social media because they are afraid of receiving negative comments from customers. Other challenges such as social media as a new trend whose use and potentials are difficult to fully comprehend as well as the possibility of fake comments were cited.

4.2 Social Media Applications and Strategies Being Employed

Sigala (2007) suggested that social networking site, blogs, Really Simple Syndication, podcasting, massively multiplayer online role playing game, tagging, mashups, and wikis, were the most commonly used social media applications in the hospitality and tourism industry. According to the discussions with industry practitioners, the types of social media applications being used by hospitality and tourism businesses studied include social networking sites (Facebook), micro-blogging sites (Twitter), media sharing sites (Youtube, Flickr, Picasa), third-party online travel communities and guides (TripAdvisor, Ctrip), customised online travel communities/forums, Google maps, tagsonomies and blogs.

Similar to the findings of Leung et al.'s (2011) study, Facebook was found to be one of the most popular social media applications used by practitioners as it was the most

frequently mentioned subject. Though the same platform is used for marketing or other business purposes, the strategies and approaches in using Facebook vary among different companies. One of the common approaches of social media use is to open an official account for the company and assign one staff member to be responsible for managing the account. As usual, the content shared on their official Facebook pages include but is not limited to photographs, videos, descriptions of their properties, industry news, information on products and services and promotional packages. In addition to sharing information, they encourage guests to review and leave comments about their services on their pages. Respondents also considered Facebook as useful for new product development. As suggested by *Manager C*, the "question" function on Facebook can be used to conduct polls in an attempt to understand the preferences of existing customers. Based on the descriptions by the practitioners, Facebook reflects a cultural shift from product "designing *for*" customers to "designing *with*" customers.

Another approach to social media use is for the company to create and manage a page on an existing platform like Facebook or on a unique domain sponsored by the company, without using its real identity. Using this approach, some organisations have created their own online communities that are relevant to the services they offer. For instance, *Manager E*'s company has created a special online travel community where any visitors interested in travel can post questions and share travel pictures or videos on the pages. The company has assigned one staff member to offer special help for travellers and give prompt responses to their questions. Instead of disclosing the real identities, practitioners believed and noted that the companies could still use these media to raise interest and awareness of their products, services and promotions and to organise polls in order to understand the needs and preferences of potential customers when developing new travel packages.

With regard to third-party online communities/guides, practitioners differed in the approaches they used. Some asked a number of staff and friends to pose and answer questions on these forums in an attempt to raise awareness and interest in the company's products and services. For instance, before its opening, *Manager D*'s company asked about 10 staff members and friends to pose questions on various forums catering for the interests of their target market regarding where to find certain type of wine. After a number of responses, they asked the same people to join the discussion and draw people's attention to the opening of a new wine bar and restaurant offering this type of wine in various varieties. Thus, social media represent a practical tool for niche marketing and interactive communication.

Dealing with negative comments has long been a concern for the companies which intend to use social media. According to the discussion on this issue, some practitioners chose to ignore it, others responded but either deny or apologise. One interesting approach worth explaining was suggested by *Manager D*. The management of his company does not respond to the negative reviews, neither does it ignore them. Instead, the management records the comments and correct the services accordingly if the complaints are found to be valid. Moreover, in order to minimise the negative impact of these negative comments, they post some positive reviews on the same site to push the negative ones downwards. As noted by *Manager D*, "if the

customer complained about the food, we add two or three more positive reviews about the wine". This approach may, however, raise credibility concerns.

4.3 Discussion of Social Media Strategies in Relations to the Consumer's Decision Making Process

In the light of the above discussions, it is apparent that the hospitality and tourism businesses involved in this study are embracing social media. The study further revealed that the practitioners' employ social media applications and strategies to address each stage of the consumer's decision making process.

According to Engel et al. (1990)'s model, consumers will firstly acknowledge their needs before making a purchase decision. These needs can be triggered by either internal or external stimuli (Kotler, Bowen, & Makens, 2003). In response to this, some practitioners employ social media to determine these needs by conducting polls and making use of questions asked by potential customers on some social media platforms. Some practitioners use the "question" function on Facebook to investigate customers' preferences and thereby utilising the responses to develop new products or customise existing products to meet their needs. Besides, the practitioners engage social media to "trigger" these needs through the creation of awareness and interests in their products and services. Sharing multimedia information is one of the approaches the businesses commonly adopt. Given that tourism products are intangible in nature and hence cannot be experienced beforehand, sharing pictures, videos or other multimedia on social networking and media sharing sites not only "tangiblises" the experience for the customers, but also arouses their interest in the product or service. Other approaches used which trigger consumers' interests include using social media to publicise promotional packages and to offer discount coupons.

In an attempt to satisfy their purchasing needs, consumers will search for and process information from various sources to come up with a set of alternatives. In line with this, practitioners promote their products and services on various social media platforms in order to maintain high exposure to potential and existing customers. Also, they continuously update the company's information and upcoming events on social media platforms in order to provide consumers with ample information and ideally raise their interest. In the information search process, potential customers may want to ask questions and communicate with service providers directly. Hence some of the practitioners assign a staff member from their companies to be responsible for answering enquiries from the general public through social media platforms. Thus the practitioners recognise customers' need for information and utilise social media accordingly to enhance their search experience.

Regarding the third stage of the decision making process, previous studies suggest that it is difficult for practitioners to manipulate customers' evaluation as customers have no single evaluation process in all buying situations (Kotler et al., 2003). Nonetheless, the practitioners interviewed do utilise social media in an attempt to influence potential consumers' evaluation of alternatives. Since customers tend to prefer personal information by previous customers (word-of-mouth) to commercial information by service providers (Crotts, 1999), comments and evaluations on the review sites and social networking sites are considered as invaluable information for supporting purchase decision making. Generally, customers prefer products or

services which are well perceived by previous consumers. In other words, positive comments may increase the chance of their products or services being selected by potential consumers. Hence, most of the businesses encourage their customers to leave comments about their products and services. Conversely, negative comments may adversely affect the company's image. As noted by *Manager D*, companies post positive reviews on the review sites to maintain a positive image on the electronic space. This approach, however, raises ethical concerns.

While an ultimate goal for employing social media is to convert the site visitors to customers, strategies on enhancing customers' purchase transactions are relatively few. One explanation is that some practitioners generally regard social media as a marketing and information dissemination platform, and not a distribution channel. In line with this, they provide hyperlinks from their social media pages to their e-commerce sites in order to increase the conversion rate. They also attempt to induce purchase intentions and decisions with promotional packages and complimentary coupons. Furthermore, some of the businesses studied have partnered with third-party online travel communities such as Ctrip to facilitate customer purchase transactions.

As all major purchases result in cognitive dissonance, a platform for customers to express their opinions is necessary and appropriate for a company's long-term development. Besides engaging social media as an arena for consumers to express opinions about products and services, practitioners should take actions to reduce consumer post-purchase dissatisfaction. Some practitioners respond to the consumers' comments on social media to address their dissatisfaction. Moreover, they take note of the customers' comments on product or service failure and rectify them accordingly. The aforementioned discussions demonstrate that the practitioners have recognised customers' needs at various stages of the purchase decision making process and attempt to employ appropriate strategies to meet these needs.

5 Conclusion and Implications

With the growing influence of user-generated content in assisting consumers' decision making, social media present unprecedented opportunities and challenges to hospitality and tourism businesses. On the basis of discussions with practitioners, this study offers valuable insights into how practitioners in the hospitality and tourism industry perceive and engage social media.

Goeldner and Ritchie (2009) have suggested that harnessing a newly developed media in business requires a solid understanding of its function and nature. In the current study, the practitioners interviewed understand what social media represents and recognise its capabilities as a channel for the promotion of their products and services, for connecting and building relationships with existing customers, and for reputation management. Generally, the organisations studied are able to exploit various social media applications and strategies to understand and proactively induce customers' needs. Contrary to the findings in Sigala's (2011) study, the organizations studied here are able to maintain a balanced value between firms and customers. They do not only enhance the benefits and values of their firms, but also, they expand this to create value for their customers. Regarding the potential challenges social media presents to their businesses, the practitioners do not underestimate them but they recognise these

challenges and ought to take the opportunity to address them. The findings thus contribute to helping address the research void regarding how service providers perceive social media.

This study further demonstrates how practitioners apply social media in order to influence the consumers' decision making process. Understanding consumers' needs and the buying process is the foundation of successful marketing (Kotler et al., 2003). While there is a growing body of research on social media in hospitality and tourism journals, most of the existing studies focused on the role and use of social media in travel information search (Xiang & Gretzel, 2010). Thus, this study adds to extant literature by looking at this phenomenon from service providers' perspective. The purchasing process starts long before and continues long after the actual purchase. Hence, Crotts (1999) and Kotler et al. (2003) have highlighted the need for marketers to focus on the entire process rather than just one stage. Though Cox et al. (2009) found that very few travellers relied on social media in the during-trip and post-trip stages, the practitioners interviewed employ strategies that are relevant to the various stages of the consumers' decision making process in pursuit of meeting consumers' needs via the applications of this strategic tool.

As noted above, the nature of the research design adopted for this study limits the generalisability of the findings. One direction for further research is to extend the sample size and the variety of managers to be interviewed to gain more insightful implications to the study's theme. On the basis of the current findings, future studies can aim at building a theory to describe practitioners' approach to social media. Future studies can employ a more post-positivist approach to compare the social media practices of hospitality and tourism practitioners with best practices suggested in the general marketing literature. Another suggestion for future work is to examine the possible differences in social media adoption and perceptions among managers from different sectors.

References

Arksey, H. & Knight, P. (1999). *Interviewing for Social Scientists*. London: Sage Publications.

Berthon, P., Pitt, L., McCarthy, I. & Kates, S. (2007). When customers get clever: Managerial Approaches to dealing with creative consumers. *Business Horizons* 50(1): 39-47.

Blackshaw, P. (2006). *The consumer-generated surveillance culture.* Retrieved August 17, 2011, from http://www.clickz.com/showPage.html?page=3576076

Compete Inc. (2006). *Embracing Consumer Buzz Creates Measurement Challenges for Marketers*. Retrieved August 17, 2011, from http://www.cymfony.com/files/pdf/Compete_Spark_12_06_Embracing_Consumer_Buzz_Creates_Measurement_Challenges.pdf

comScore. (2007). *Online consumer-generated reviews have significant impact on offline purchase behavior*. Retrieved August 17, 2011, from http://www.comscore.com/Press_Events/Press_Releases/2007/11/Online_Consumer_Reviews_Impact_Offline_Purchasing_Behavior

Cox, C., Burgess, S., Sellitto, C. & Buultjens, J. (2009). The Role of User-Generated Content in Tourists' Travel Planning Behavior. *Journal of Hospitality Marketing & Management* 18(8): 743-764.

Crotts, J. (1999). Consumer Decision Making and Pre-purchase Information Search. In A. Pizam and Y. Mansfeld (Eds.), *Consumer Behavior in Travel and Tourism* (pp. 149-168). New York: The Haworth Hospitality Press.

Engel, J. F., Blackwell, R. D. & Miniard, P. W. (1990). *Consumer behaviour* (6th Edition). Hinsdale, IL: Dryden Press.

Fodness, D. & Murray, B. (1997). Tourist information search. *Annals of Tourism Research* 24(3): 503-523.

Fuchs, M., Scholochov, C. & Höpken,W. (2009). E-Business adoption, use, and value creation: An Austrian Hotel Study. *Information Technology & Tourism* 11(4): 267-284.

Goeldner, C. R. & Ritchie, J. R. B. (2009). *Tourism: Principle, Practices, Philosophies* (11th Edition). Hoboken, New Jersey: John Wiley & Sons Inc.

Gretzel, U. & Yoo, K. H. (2008). Use and impact of online travel reviews. In P. O'Connor, W. Hopken, & U. Gretzel (Eds.), *Information and Communication Technologies in Tourism 2008* (pp. 35-46). New York: Springer-Wien.

Kietzmann, J. H., Hermkens, K., McCarthy, I. P. & Silvestre, B. S. (2011). Social Media? Get Serious! Understanding the functional building blocks of social media. *Business Horizons* 54(1): 241-251.

Kotler, P., Bowen, J. & Makens, J. (2003). *Marketing for hospitality and tourism* (3rd Edition). Upper Saddle River, NJ: Prentice Hall.

Leung, D., Lee, H. A. & Law, R. (2011). Adopting Web 2.0 technologies on chain and independent hotel websites: A case study of hotels in Hong Kong. In R. Law, M. Fuchs, F. Ricci (Eds.), *Information and Communication Technologies in Tourism 2011* (pp. 229-240). New York: Springer-Wien.

Lincoln, Y. S. & Guba, E. G. (1985). *Naturalistic Inquiry*. Beverly Hills, CA: Sage Publication.

Litvin, S. W., Goldsmith, R. E. & Pan, B. (2007). Electronic word-of-mouth in hospitality and tourism management. *Tourism Management* 29(3): 458-468.

Maxwell, J. (2005). *Qualitative Research Design: An Interactive Approach* (2nd Edition). Thousand Oaks: Sage Publications.

Milano, R., Baggio, R. & Piattelli, (2011). The effects of online social media on tourism websites. In R. Law, M. Fuchs, and F. Ricci (Eds.), *Information and Communication Technologies in Tourism 2011* (pp. 471-483). New York: Springer-Wien.

Miles, M. B. & Huberman, A. M. (1994). *Qualitative data analysis* (2nd Edition). Thousand Oaks, CA: Sage Publications.

O'Connor, P. (2010). Managing a Hotel's Image on TripAdvisor. *Journal of Hospitality Marketing & Management* 19(7): 754-772.

Sánchez-Franco, M. J. & Rondan-Cataluña, R. J. (2010). Virtual travel communities and customer loyalty: Customer purchase involvement and web site design. *Electronic Commerce Research and Applications* 9(2): 171–182.

Schmollgruber, K. (2007). *Hotel 2.0 –Interview with Opus Hotel CEO Blogger Daniel Craig*. Retrieved August 24, 2011, from: http://passionpr.typepad.com /tourism/2007/04/hotel_blog_opus.html

Sigala, M. (2004) Designing experiential websites in tourism and hospitality: A customer-centric value approach. In A. J. Frew (Ed.), *Information and Communication of Technologies in Tourism 2004* (pp. 162–171). New York: Springer-Wien.

Sigala, M. (2007). *WEB 2.0 in the tourism industry: A new tourism generation and newe-business models*. Retrieved August 29, 2011, from http://195.130.87.21:8080/ dspace/handle/123456789/386

Sigala, M. (2011). eCRM 2.0 applications and trends: The use and perceptions of Greek tourism firms of social networks and intelligence, *Computers in Human Behavior* 27(2): 655-661.

Strauss, A. (1987). *Qualitative research for social scientists*. Cambridge: Cambridge University Press.

Xiang, Z. & Gretzel, U. (2010). Role of social media in online travel information search. *Tourism Management* 31(2): 179-188.

Zehrer, A., Crotts, J. & Magnini, V. (2011). The perceived usefulness of blog postings: An extension of the expectancy-disconfirmation paradigm. *Tourism Management* 32(1): 106-113.

Acknowledgements

The authors would like to thank the three anonymous reviewers for their constructive comments on an earlier version of this paper. The authors further express their gratitude to the practitioners who made themselves available for this study. The work presented in this paper was supported by the internal grants funded by the Hong Kong Polytechnic University (Project Number: A-PK54).

Social Media Use and Impact during the Holiday Travel Planning Process

John Fotis[a],
Dimitrios Buhalis[a], and
Nicos Rossides[b]

[a]School of Tourism,
Bournemouth University, U.K.
{jfotis, dbuhalis}@bournemouth.ac.uk

[b]MASMI Research Group, Cyprus
nicos.rossides@masmi.com

Abstract

Through an empirical study among holiday travellers, residing in the Former Soviet Union Republics, this paper presents a comprehensive view of role and impact of social media on the whole holiday travel planning process: Before, during and after the trip, providing insights on usage levels, scope of use, level of influence and trust. Findings suggest that social media are predominantly used after holidays for experience sharing. It is also shown that there is a strong correlation between perceived level of influence from social media and changes made in holiday plans prior to final decisions. Moreover, it is revealed that user-generated content is perceived as more trustworthy when compared to official tourism websites, travel agents and mass media advertising.

Keywords: Social media; travel planning process; holidays

1 Introduction

As early as in 2006 it was argued that social media started replacing traditional sources of information (Jepsen, 2006). Later, it has been supported that social media might influence the buying behaviour process; at least as described in Inputs-Processing-Response models (Constantinides & Fountain, 2008). Moreover, De Valck, Van Bruggen, & Wierenga (2009) suggested that virtual communities are becoming important networks of consumer knowledge that influence consumer behaviour. Today, social media enjoy a phenomenal rise in their popularity among internet users: Facebook claims that its active users reached more than 750 million worldwide, spending more than 700 billion minutes per month (Facebook, 2011); Twitter hosts 175 million users who on an average week post one billion tweets (Twitter, 2011); YouTube users view daily over 3 billion videos (YouTube, 2011); and it is estimated that there are over 170 million blogs worldwide (BlogPulse, 2011).

In tourism, consumers' behaviour has always been influenced by developments in Information Communication Technologies (Buhalis, 1998; Poon, 1993), and especially it is Web 2.0 that dramatically changed how consumers plan and consume travel related products (Buhalis & Law, 2008). Gretzel, Kang, & Lee (2008) predicted that social media impacts in travel will be tremendous. In 2008 it was found that 82% of US online consumers have checked online reviews, blogs and other online feedback for their travel related purchasing decisions (eMarketer, 2008). Today, TripAdvisor, perhaps the leader among travel related consumer review websites (a

subset type of social media), serves more than 50 million users per month who seek advice about their travel plans and hosts more than 50 million travel reviews and opinions (TripAdvisor, 2011). Moreover, it is argued that social media "are taking an important role in travellers' information search and decision-making behaviours" (Yoo, Gretzel, & Zach 2011, p. 526).

To further describe the impact of social media in holiday travel, this study attempts to measure social media usage levels and scope during the whole travel planning process (pre, during and post-trip). Furthermore it attempts to measure perceived level of social media influence on destination and accommodation choice, as well as perceived level of trust to social media content in relation to other sources of travel related information.

2 Literature Review

2.1 Social Media

Kaplan and Haenlein (2010, p.61) define social media as "a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of User Generated Content". Although there are a number of approaches in terms of social media taxonomies (Constantinides, 2009; Fischer & Reuber, 2011; Kim, Jeong, & Lee, 2010; Mangold & Faulds, 2009) Kaplan and Haenlein (2010) proposed a taxonomy scheme that classifies social media types according to their (a) level of social presence / media richness, and (b) level of self presentation / self disclosure, identifying six types of social media: social networking websites (i.e. Facebook, Linkedin), blogs, content communities (i.e.YouTube, Flickr, Scrib, Slideshare, Delicious), collaborative projects (i.e. Wikipedia, Wikitravel), virtual social worlds (i.e. Second Life), and virtual game worlds (i.e. World of Warcraft). Still however, such a taxonomy neglects other types of social media such as microblogs (i.e. Twitter), consumer review & rating websites (i.e. TripAdvisor, Epinions) and internet fora (i.e. ThornTree, Fodor's Travel Talk).

2.2 Social Media in Travel

Holiday travel related purchases are considered complex due to the composite and experiential nature of the holiday travel product, involve high risks and as a result require extensive information search (Sirakaya & Woodside, 2005). Within such information search processes, consumers rely on other travellers' experiences as a mean to increase the exchange utility and decrease uncertainty (Kotler, Bowen, & Maken, 2010; Litvin, Goldsmith, & Pan, 2008; Yoo, Lee, & Gretzel, 2007). Just after the creation of the first virtual communities (Rheingold, 1993) it became apparent that their online content was perceived similar to recommendations provided by friends, family and "like-minded souls" (Fernback & Thompson, 1995; Wang, Yu, & Fesenmaier, 2002). During the Web 2.0 era social media applications "exploded" in popularity, usage levels providing a plethora of characteristics enabling self expression and sharing of content. Ad-ology supports that 23% of US Internet users were "somewhat" or "significantly influenced" by social media for their travel / holiday related decisions (eMarketer, 2010). Social media are therefore becoming increasingly important in travel planning, primarily for their function as vital

information sources providing access to other travellers' experiences (Chung & Buhalis, 2008; Yoo et al., 2011). At the same time, apart from their function as information sources, social media enable storytelling, a usual post-travel activity, on a '24/7' basis to large audiences, and also provide a sense of belonging into virtual travel communities (Gretzel, Fesenmaier, & O'Leary, 2006).

A number of studies focus on the impact and role of social media in travel related decisions: Gretzel, Yoo, & Purifoy (2007) found that online reviews posted in a travel related consumer review and rating website increase travellers' confidence during decision making, reduce risk, assist them in selecting accommodation and therefore facilitate decision making. Moreover, it was found that travellers read accommodation reviews throughout the various stages of the travel planning process: Before travel as a source of ideas, as a mean to narrow down choices, and post accommodation choice in order to confirm the choice made; during the trip; after the trip to compare and share experiences; but also as an ongoing process even if there is no trip ahead. Mack, Blose, & Pan (2008) studied the influence and credibility of travel blogs and found that traditional WOM is more trustworthy than blog posts, suggesting that WOM generated from sources with which travellers have strong social ties is more trustworthy than WOM from strangers. However, they found that those who post in blogs perceive the authoritativeness of blogs as similar to that of traditional WOM, leaving space for future improvement of blogs' credibility as the number of those who post to blogs increases over time (Technorati, 2010). Yoo, Lee, Gretzel, & Fesenmaier (2009) found that user generated content is perceived as more credible when posted to official tourism bureau sites rather than in review sites, travel blogs social networking sites and content communities. Vermeulen & Seegers (2009) studied consumer reviews' impacts on choice of accommodation and found that the consideration of a hotel is enhanced by exposure to both negative and positive consumer reviews. In their attempt to reveal the role of social media throughout the travel planning process, Cox, Burgess, Sellitto, & Buultjens (2009) found that social media are mostly used before the trip, while during and after the trip their use was very limited. More specifically, social media were primarily used after the holiday destination choice rather than during the evaluation of destination choices. Moreover, social media were perceived as less trustworthy than traditional sources of information (i.e. official tourism websites and travel agents). It may be argued however, that such a result may be attributed to the fact that the study's sample was derived from a mailing list of an official tourism website and therefore there may be a positive predisposition towards official sources of information. White (2010) suggests that travel related photos in Facebook generate interest to viewers and can very easily become part of the viewer's travel plans. Contrary to findings of Mack et al. (2008), and Yoo et al. (2009), Del Chiappa (2011) supports that trustworthiness of tourism-related blogs is second only to consumers' reviews and ratings found in online travel agents' websites.

In parallel with the argument made by Parra-López, Bulchand-Gidumal, Gutiérrez-Taño, & Díaz-Armas (2011), the above discussion also suggests that the majority of existing studies attempt to describe the role of social media either focusing on specific types of social media and specific communities, or at a specific stage of the travel planning process, evidencing that there is no adequate academic research on the role

and impact of social media as a whole throughout the holiday travel planning process. This study presents an attempt to address this gap and pursues the following research questions and resulting hypotheses:

RQ1: To what extent and for what reasons are social media used during the holiday travel planning process?

H1: Social media are predominantly used before the trip for information search purposes.

RQ2: Do social media influence holiday plans?

H2: The higher the perceived level of social media influence on holiday destination choice or accommodation choice, the more likely is that changes would be made to holiday plans.

RQ3: Are social media more trustworthy than traditional sources of holiday related information?

H3: Holiday travel related information provided in social media is more trustworthy than mass media advertising, travel agents and official tourism websites.

3 Methodology

3.1 Data Collection

An online questionnaire survey was conducted in October 2010 in a random subset of a commercial online research panel whose members are residing in 12 Former Soviet Union Republics (F.S.U.) including Russia. An e-mail invitation was sent to 900 panellists containing a link to the online questionnaire. Without any reminder 368 panel members completed the questionnaire achieving a response rate of 41%. Respondents had the option to complete part of the questionnaire, and as a result 22 questionnaires were incomplete, thus 346 questionnaires were used for further analysis. Through a screening question it was assured that all respondents have taken at least one holiday trip during the last 12 months. As a result the sample of the study consists of 346 internet users residing in 12 F.S.U. Republics who have taken at least one holiday trip within 12 months prior to the study. The questionnaire included an introductory statement explaining types of websites considered as social media within six broad categories: Blogs, photo & video sharing websites, microblogs, wikis, social networking websites and travel review websites. This taxonomy was created by the authors after a review of existing literature (Constantinides, 2009; Fischer & Reuber, 2011; Kim, Jeong, & Lee, 2010; Mangold & Faulds, 2009; Xiang & Gretzel 2010). For each of the above types a number of examples were given in order to increase the reliability of the instrument. Examples included websites not only among those dominating the western world but also websites operating exclusively within F.S.U. Republics such as Vkontakte.ru and Odnoklassniki.ru in terms of social networking websites.

3.2 Measures

The questionnaire was exclusively designed for the purposes of this study. Respondents were asked to recall their last holiday trip within the last 12 months, the

destination country of this trip and the nights spent there. Three sets of multiple response questions were used to measure social media use before, during and after the holiday trip. A total of 12 statements providing different reasons for social media use were used, of which five were borrowed from Cox et al. (2009). Perceived level of social media influence on destination and accommodation choice was measured a 7-point scale ranging from "Not Influential At All" (1) to "Very Influential" (7) transforming a five point scale used by Lo, Cheung, & Law (2002) to a seven point scale. Perceived level of trust to holiday travel related information sources was measured by a seven point Likert scale ranging from "Strongly Disagree" (1) to "Strongly Agree" (7) as per Cox et al (2009). Of the seven statements used to describe an equal number of information sources, four were borrowed from Cox et al. (2009) to enable comparison of findings. The questionnaire also included, among other, measures to provide socio-demographic characteristics and travel behaviour.

4 Results

4.1 Profile of Sample

The sample had the following characteristics: As per the gender, 65.6% were females and 34.4% males. The over representation of female respondents has been also observed in other studies (Cox et al., 2009; Gretzel et al., 2007), and may to an extent be related to gender differences in holiday-related roles (Decrop, 2006; Mottiar & Quinn, 2004). As per the age 18.2% were less than 25 years, 38.4% from 25 to 39, 30.9% from 40 to 54, and 12.5% 55 or older. As per the education, 30.3% have completed secondary school, and 69.7% were university graduates. Moreover, 64.2% were residing in Russia, and 35.8% in the other FSU Republics, a proportion that approximates the distribution of actual internet users between the two regions (InternetWorldStats, 2010).

4.2 Use of Social Media

The vast majority of respondents (97%) have visited at least one social media website during the last 12 months. Almost half (49%) of respondents visited social media websites several times a day, 36% almost every day, 9% only sometimes per week, and 3% very rarely. Level of use was found to be differentiated by age, $\chi^2(15, N = 346) = 34.99$, $p = .01$, and level of education, $\chi^2(5, N = 346) = 11.92$, $p = .04$. The vast majority (94%) of respondents "below 25" visit social media websites "several times a day" or "almost every day", whereas among those "55 and over" the same usage level decreases to 72%. In terms of level of education, 88% of university graduates visit social media websites "several times a day" or "almost every day", whereas among those who completed secondary school the same usage level decreases to 79%.

4.3 Social Media Usage Levels and Reasons for Use During the Holiday Planning Process

Figure 1 reveals that among the three stages of the travel planning process (before, during, and after the trip) it is after the trip that social media are predominantly used for the purpose of sharing experiences and photos with friends and/or other travellers. More specifically before their holiday trip 45% of respondents searched for ideas on

where to go for holidays, and 42% seek ideas and information on excursions and other leisure activities. The analysis did not reveal any significant differences among age groups, level of education, and region of residence. Gender differences were observed only in usage levels for the purpose of confirming that a good destination choice was made, $\chi^2(1, N = 346) = 4.64$, $p = .03$, since 35% of females used social media for that purpose as opposed to 24% of men.

During the holiday trip, respondents primarily used social media for a non-holiday related reason: to stay connected with their friends (50%), and to a lesser extent to find holiday related information (30%). Providing comments and reviews about the holiday experience, while at the destination, engaged only 17% of respondents, however significant differences among age groups were observed, $\chi^2(3, N = 346) = 10.62$, $p = .01$: One every four holidaymakers "below 25" posted reviews during trip, a percentage that drops to 9% for those aged 40 and above.

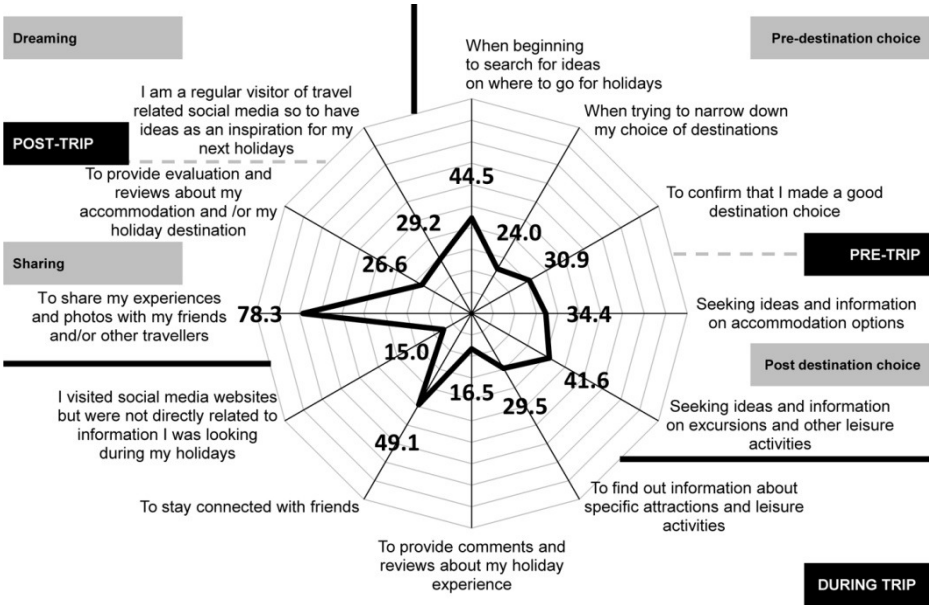


Fig. 1. Percentage (%) of respondents that used social media during their travel planning process and reasons of use

It is after the holiday trip that respondents primarily used social media, as 78% used them to share experiences and photos with friends and / or other travellers. Providing reviews and evaluations engaged 27% of respondents, however such activity level differentiates among age groups, $\chi^2(3, N = 346) = 12.28$, $p = .01$. More than a third (36%) of respondents between 25 and 39 engaged in such activity, as opposed to only 12% of those aged over 55.

From the above discussion it is evident that H1 (Social media are predominantly used before the trip for information search purposes) is rejected. This finding is in contrast

with Cox et al. (2009) who found that social media are predominantly used during the before the trip stage for information search purposes.

4.4 Social Media Influence on Holiday Plans

Perceived level of social media influence on holiday planning was assessed in terms of destination choice, accommodation choice and changes made to holiday plans before the final decision. Social media perceived level of influence on destination choice was measured at 4.84 (SD = 1.63) and on accommodation choice on 4.61 (SD = 1.81) both on a 1-7 scale. Independent samples t-tests and one-way analysis of variance did not reveal any significant differences in gender, education level, age, frequency of travel, or region of residence.

To assess change in holiday plans, respondents who visited social media websites while they were still planning their holiday trip (n=273) were asked: "Before your final decisions about your last holiday, did you make any changes to your original holiday plans because of other travellers' opinions, reviews, photos, videos, or other information that you found in social media websites?" Table 1 presents the findings.

Table 1. Social media level of influence and changes in original holiday plans

	% respondents (n=273)	Influence on destination choice Mean scores* (SD)	Influence on accommodation choice (SD)
I am not sure / cannot remember if I made any changes	2.2%	2.75 (1.50)	2.75 (2.36)
I did not make any changes	33.7%	4.10 (1.80)	3.93 (1.91)
I did make few changes to my original holiday plans	49.5%	5.25 (1.37)	5.02 (1.58)
I did make significant changes to my original holiday plans	14.7%	5.35 (1.21)	5.00 (1.73)
Pearson Correlation test:		$r(268) = .34$, $p < .001$	$r(268) = .27$, $p < .001$

(*) Measured on a 7-point Likert scale, where: 1 = Not influential at all, 7 = Very influential

The Pearson correlation test was employed to test H2 (The higher the perceived level of social media influence on holiday destination choice or accommodation choice, the more likely is that changes would be made to holiday plans). In both cases (Table 1) $p<.001$ evidencing that (a) as influence from social media on destination choice increases the more likely is that there are changes in the holiday plans, and (b) as influence from social media on accommodation choice increases the more likely is that there are changes in the holiday plans. As a result hypothesis H2 is accepted.

4.5 Trust Towards Social Media and Traditional Sources of Information

Respondents were asked to indicate their agreement or disagreement with seven statements in the form of "I trust information about holidays provided by..." followed by the information source. As it can be seen in Table 2, friends and relatives are the most trustworthy source of information, followed by information provided by other travellers in various websites, whereas advertisements in mass media are the least trustworthy among the information sources examined.

Table 2: Perceived level of trust in holiday-related information sources

I trust information about holidays provided by… (N=346, 7-point Likert scale:1=*Strongly Disagree*, 7 = *Strongly Agree*)	Mean (SD)
Friends and relatives	6.36 (1.14)
Information provided by other travellers in various websites	5.15 (1.41)
Social media	4.61 (1.37)
Official tourism websites (state / government owned)	4.36 (1.57)
Shows or documentaries in TV, in radio, or articles in newspapers and magazines	3.97 (1.45)
Travel agents	3.83 (1.36)
Advertisements in TV, radio, newspapers and magazines	3.56 (1.35)

The above findings are in disagreement with of Cox et al. (2009) who using the same scale found that official tourism websites and travel agents are more trustworthy in comparison to "comments made by travellers on third party sites e.g. TripAdvisor". One potential reason to explain the increased level of trust in official tourism websites observed in Cox et al. (2009) may be that their sample was taken from the database of e-mail subscribers of a destination's official tourism website. As a result, it may be argued that the sample had a positive predisposition or even a preference towards official sources of information.

To test H3 (Holiday travel related information provided in social media is more trustworthy than mass media advertising, travel agents and official tourism websites) one-sample t-tests evidence that the value of 5.15 (observed as a mean level of trust of information provided by other travellers in various websites) is significantly different to the means observed for official tourism websites $t(345) = -9.442$, $p = .000$; travel agents, $t(345) = -18.019$, $p = .000$; and mass media advertising $t(345) = -21.890$, $p = .000$. As a result, hypothesis H3 is accepted.

5 Conclusions, Limitations and Implications

This paper contributes to the current literature on the use and impact of social media on holiday travel planning. Contrary to the majority of existing studies that are medium or community specific, or focus in a specific stage of the travel planning process, this study attempts to provide a comprehensive picture on the impact of social media as a whole, throughout the travel planning process. In relation to the study of Cox et al. (2009) this study attempts to describe social media use and impact during the last holiday trip, therefore it attempts to depict actual and not hypothetical travel behaviour (Beaulieu & Schreyer, 1985, cited Um & Crompton, 1990). Moreover, by focusing at travellers residing at F.S.U. Republics, a market with distant national culture characteristics to those of the Australian culture, it attempts to add to the discussion raised by Gretzel et al. (2008) that there are differences in terms of social media use among national markets.

This study shows that social media are used during all stages of the holiday planning process (before, during and after holidays) however, to a different extent and for a different purpose. In contrast to the findings of Cox et al. (2009) who found that social media are predominantly used during the information search stage of the travel planning process, this study evidences that travellers residing in the F.S.U. Republics

use social media predominantly during the post-trip stage for sharing experiences and photos with friends and/or other travellers. It may be argued that such a finding is associated primarily with Russian's high level of engagement with social networking websites (comScore, 2010), although residents of the other F.S.U. Republics do not seem to differentiate. The second most popular use of social media was observed during holidays: as means to enable travellers to stay connected with friends. Both the first and the second most popular uses of social media observed in this study may be attributed to the very low individualist / very high collectivist nature of the Russian culture (Hofstede, Hofstede, & Minkov, 2010).

In terms of perceived level of influence, social media were rated as "somehow influential" on both destination and accommodation choice. However, more than six out of ten (65%) respondents stated that they indeed made some sort of changes to their original plans because of exposure to user-generated content in social media websites, with 50% describing them as "few changes" and 15% as "significant changes". Moreover, the present study reveals a strong correlation between social media level of influence on destination and accommodation choice, and the changes made in holiday plans before final decisions were taken. More specifically, as the perceived level of influence from social media on destination choice increases, the more likely is that there were changes to holiday plans in terms of destination selection. Similarly, as perceived influence from social media on accommodation choice increases the more likely is that there were changes in the holiday plans in terms of accommodation selection.

This study also examined the perceived level of trust among seven holiday related information sources: Official tourism websites, publicity and advertorials in mass media (i.e. TV or radio shows and documentaries, newspapers and magazines' articles), advertisements in mass media, travel agents, social media, friends and relatives, and information from other travellers in various websites. Among those, friends and relatives were rated as the most trustworthy source, followed by information from other travellers in various websites. In contrast to the findings of Cox et al. (2009) who found that state tourism websites and travel agents outscored, in terms of trustworthiness, comments by other travellers found in third party sites (i.e. TripAdvisor), blogs and social network sites, this study found that information from other travellers in various websites is trusted more than official tourism websites, and travel agents.

Finally, it should be stated that a comparison of the findings of this study to that of Cox et al. (2009), at least for the constructs that are identically operationalized in both studies, contributes to the empirical confirmation of the claim made by Gretzel et al. (2008) that there are differences in terms of social media use among national markets. More than that, such a comparison suggests that among national markets, apart from differences in social media adoption and use, there are also indications for differences in the perceived level of trust towards social media. Such differences may not be present in closely related national cultures (as this study evidences by examining Russians and residents of the other F.S.U. Republics), however, there are indications that differences are more evident when comparing distant national cultures such as Australians (as per the case in Cox et al. 2009) and Russians. Based on Hofstede et al. (2010) Australians and Russians differ significantly in all four cultural dimensions: in

power distance (Russia: 93, Australia: 36); in individualism (Russia: 39, Australia 90); in uncertainty avoidance (Russia: 95, Australia: 51); and in masculinity (Russia: 36, Australia: 61).

5.1 Research Limitations

Three limitations apply for this study: (a) The sample is not random due to the self-response nature of the specific online panel survey; (b) there was no treatment for non-responses; and (c) findings apply only to the specific geographical context (F.S.U. Republics) and therefore cannot be generalized especially to other national markets with distant cultural characteristics.

5.2 Managerial Implications

This study provides insights and guidelines to industry practitioners to enable them fine tune their social media campaigns in the F.S.U. Republics. First, this study provides evidences that social media are being used during all stages of the holiday travel planning process, however to a different extent and for a different scope. Second, in combination with the work of Cox et al. (2009) this study provides preliminary indications that national source tourism markets behave differently not only in terms of social media adoption and usage levels, but also in terms of scope of use, as well as in perceived levels of trust among travel related information sources. Such a finding clearly suggests that national source markets should be studied individually prior to the design and implementation of social media campaigns. Third, the "during the holidays" stage remains a challenging domain, as social media seem to be used primarily for maintaining contact with friends, while providing comments and reviews on the spot is a least preferred activity. Fourth, in terms of the context of this research, there are only minor differences between Russia and the other F.S.U. Republics source markets in terms of the impact of social media on the holiday planning process, therefore, there is no need for major differentiation in social media campaigns aiming at travellers in those two regions.

5.3 Suggestions for Further Research

The findings of this study suggest that social media travel related research should place equal emphasis in all stages of the travel planning process. Moreover, taking into account the findings of Cox et al. (2009) among Australian internet users and the findings of the present study, it can be claimed that the impact of social media on holiday related travel planning differs among tourism source markets, with cultural differences contributing to such a difference. As a result, an adequate number of cross-cultural studies are needed to substantiate such a claim.

References

BlogPulse. (2011). *Blogpulse Stats.* Retrieved from www.blogpulse.com.

Buhalis, D. (1998). Strategic use of information technologies in the tourism industry. *Tourism Management* 19(5): 409-421.

Buhalis, D. & 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* 29(4): 609-623.

Chung, Y. & Buhalis, D. (2008). Information Needs in Online Social Networks. *Information Technology & Tourism* 10(4): 267-281.

comScore. (2010). *Russia has the most engaged social networking audience worldwide*. Retrieved from http://www.comscore.com/Press_Events/Press_Releases/2010/10/ Russia_ Has_Most_Engaged_Social_Networking_Audience_Worldwide

Constantinides, E. (2009). *Social Media / Web 2.0 as Marketing Parameter: An Introduction.* Paper presented at the International Marketing Trends Conference. Retrieved from http://marketing-trends-congress.com/2009_cp/Materiali/Paper/Fr/Constantinides.pdf

Constantinides, E. & Fountain, S. (2008). Web 2.0: Conceptual foundations and marketing issues. *Journal of Direct, Data and Digital Marketing Practice* 9(3): 231-244.

Cox, C., Burgess, S., Sellitto, C. & Buultjens, J. (2009). The Role of User-Generated Content in Tourists' Travel Planning Behavior. *Journal of Hospitality Marketing & Management* 18(8): 743-764.

De Valck, K., Van Bruggen, G. & Wierenga, B. (2009). Virtual communities: a marketing perspective. *Decision Support Systems* 47(3): 185-203.

Decrop, A. (2006). *Vacation decision making.* Wallingford, UK: CABI Publishing.

Del Chiappa, G. (2011). Trustworthiness of Travel 2.0 applications and their influence on tourist behaviour: an empirical investigation in Italy. In Law, R., Fuchs, M. & Ricci, F. (Eds.), *Information and Communication Technologies in Tourism* (pp. 331-342). New York, NY: Springer.

eMarketer. (2008). *Online reviews sway shoppers.* Retrieved from http://www.emarketer.com/ Article.aspx?R=100640

eMarketer. (2010). *Online Leisure Travel: Six Post-Recession Trends.* Retrieved from http:// www.emarketer.com/Report.aspx?code=emarketer_2000638

Facebook. (2011). *Statistics.* Retrieved 8 July 2011, from http://www.facebook.com/ press/info.php?statistics

Fernback, J. & Thompson, B. (1995). *Virtual communities: Abort, retry failure?* Retrieved 14 January 2010, from http://www.well.com/user/hlr/texts/ VCcivil.html

Fischer, E. & Reuber, A. R. (2011). Social interaction via new social media: (How) can interactions on Twitter affect effectual thinking and behavior? *Journal of Business Venturing* 26(1): 1-18.

Gretzel, U., Fesenmaier, D. & O'Leary, J. (2006). The transformation of consumer behaviour. *Tourism Business Frontiers: Consumers, Products and Industry* 9-18.

Gretzel, U., Kang, M. H. & Lee, W. J. (2008). Differences in consumer-generated media adoption and use: a cross-national perspective. *Journal of Hospitality & Leisure Marketing* 17(1-2): 99-120.

Gretzel, U., Yoo, K. & Purifoy, M. (2007). *Online travel review study: Role and impact of online travel reviews.* Laboratory for Intelligent Systems in Tourism, Texas A & M University. Retrieved from http://www.tripadvisor.com/pdfs/OnlineTravelReview Report.pdf

Hofstede, G. H., Hofstede, G. J. & Minkov, M. (2010). *Cultures and organizations: Software of the mind: International cooperation and its importance for survival* (Rev. 3rd ed.). New York, NY: McGraw-Hill.

InternetWorldStats. (2010). *Internet World Stats: Usage and Population Statistics.* Retrieved from http://www.internetworldstats.com/ stats.htm

Jepsen, A. (2006). Information Search in Virtual Communities: Is it Replacing Use of Off-Line Communication? *Journal of Marketing Communications* 12(4): 247-261.

Kaplan, A. & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons*, *53*(1), 59-68.

Kim, W., Jeong, O. R. & Lee, S. W. (2010). On social Web sites. *Information Systems* 35(2): 215-236.

Kotler, P., Bowen, J. & Makens, J. C. (2010). *Marketing for hospitality and tourism* (5th ed.). Boston, MA: Pearson Education.

Litvin, S. W., Goldsmith, R. E. & Pan, B. (2008). Electronic word-of-mouth in hospitality and tourism management. *Tourism Management* 29(3): 458-468.

Lo, A., Cheung, C. & Law, R. (2002). Information search behaviour of Hong Kong's inbound travelers-a comparison of business and leisure travellers. *Journal of Travel & Tourism Marketing* 13(3): 61-81.

Mack, R., Blose, J. & Pan, B. (2008). Believe it or not: Credibility of blogs in tourism. *Journal of Vacation Marketing* 14(2): 133-144.

Mangold, W. & Faulds, D. (2009). Social media: The new hybrid element of the promotion mix. *Business Horizons* 52(4): 357-365.

Mottiar, Z. & Quinn, D. (2004). Couple dynamics in household tourism decision making: Women as the gatekeepers? *Journal of Vacation Marketing* 10(2): 149-160.

Parra-López, E., Bulchand-Gidumal, J., Gutiérrez-Taño, D. & Díaz-Armas, R. (2011). Intentions to use social media in organizing and taking vacation trips. *Computers in Human Behavior* 27(2): 640-654

Poon, A. (1993). *Tourism, technology and competitive strategies.* Wallingford, UK: CAB International.

Rheingold, H. (1993). *The virtual community: Homesteading on the electronic frontier.* New York: Addison-Wesley.

Sirakaya, E. & Woodside, A. G. (2005). Building and testing theories of decision making by travellers. *Tourism Management* 26(6): 815-832.

Technorati. (2010). *State of the Blogsphere 2010.* Retrieved from http://technorati.com /blogging/article/state-of-the-blogosphere-2010-introduction/

TripAdvisor. (2011). *Fact Sheet.* Retrieved 8 September 2011, from http://www.tripadvisor. com /pages/about_us.html

Twitter. (2011). *Twitter blog: #numbers.* Retrieved from http://blog.twitter.com/2011/03/ numbers.html

Um, S. & Crompton, J. L. (1990). Attitude determinants in tourism destination choice. *Annals of Tourism Research* 17(3): 432-448.

Vermeulen, I. & Seegers, D. (2009). Tried and tested: The impact of online hotel reviews on consumer consideration. *Tourism Management* 30(1): 123-127.

Wang, Y., Yu, Q. & Fesenmaier, D. (2002). Defining the virtual tourist community: implications for tourism marketing. *Tourism Management* 23(4): 407-417.

White, L. (2010). Facebook, Friends and Photos: A Snapshot into Social Networking for Generating Travel Ideas. In N. Sharda (Ed.), *Tourism Informatics: Visual Travel Recommender Systems, Social Communities, and User Interface Design* (pp. 115-129). Hershey, PA: IGI Global.

Yoo, K.-H., Gretzel, U. & Zach, F. (2011). Travel Opinion Leaders and Seekers. In R. Law, M. Fuchs & F. Ricci (Eds.), *Information and Communication Technologies in Tourism 2011* (pp. 525-535). New York: NY, Springer.

Yoo, K.-H., Lee, Y., Gretzel, U. & Fesenmaier, D. R. (2009). Trust in travel-related consumer generated media. In Höpken, W., Gretzel, U. & Law, R. (Eds), *Information and communication technologies in tourism 2009* (pp. 49-59). New York: NY, Springer.

Yoo, K. H., Lee, K. S., & Gretzel, U. (2007). The role of Source Characteristics in eWOM: What Makes Online Travel Reviewers Credible and Likeable? In Sigala, M., Mich, L., Murphy, J. & Frew, A. (Eds.), *Proceedings of the 14th International ENTER Conference in Ljubljana, Slovenia, January 24-26, 2007* (pp. 23-34). UK: Axon Imprint.

YouTube. (2011). *Statistics.* Retrieved 6 September 2011, from http://www.youtube.com /t/press_statistics

Xiang, Z. & Gretzel, U. (2010). Role of social media in travel information search. *Tourism Management* 31(2): 179-188.

An Assessment of Contagion on Social Networking Sites

Iis P. Tussyadiah

School of Tourism & Hospitality Management
Temple University, USA
iist@temple.edu

Abstract

Social network influence has been recognized as an important determinant for consumer behaviour. Through a web-based survey using restaurant consumption as a research context, this study explores social influence resulting from two distinct social reference processes: communication and comparison. The relationship between social interactions and social influence is moderated by opinion leadership and attitude towards status consumption, which are conceptualized to represent consumers' competitiveness. Consumers' status consumption contributes negatively to recommendation-based consumption, but positively to competition-based consumption. In other words, the more competitive the consumers are, the less they tend to follow others' recommendation, the more they consume products and services to establish or maintain their status in the social network. Several managerial implications are provided.

Keywords: territoriality; mobility; location-based media; mobile technology.

1 Introduction

The development of social network platforms on the internet has brought a tremendous impact to the facilitation of global social interconnections. The so-called social media offers unparalleled constant connectivity for their users, allowing them to share, collaborate and establish an online community. Consumers are using these media to share information, exchange opinions and recommendations, and display certain consumption behaviour. In that, they disseminate positive and negative word-of-mouth on various products and services (De Bruyn & Lilien, 2008), influencing the behaviour of many others. Discussions and sentiments about products and services are found extensively on consumer opinion platforms, blogs and microblogs, online communities, and social networking sites (Hennig-Thurau, Gwinner, Walsh, & Gremler, 2004; Jansen, Zhang, Sobel, & Chowdury, 2009; Zhang, Lee, & Zhao, 2010). Consequently, online social network has the ability to significantly impact reputation, sales, and even survival of product and service providers (Kietzmann, Hermkens, McCarthy & Silvestre, 2011).

Indeed, social network influence has been recognized as an important factor in shaping consumer behaviour. Word-of-mouth (WOM) communication, which is an informal person-to-person communication among non-commercial communicators and receivers about products and services, is believed to be a powerful tool for advertising and promotion (Hennig-Thurau & Walsh, 2004; Harrison-Walker, 2001; Westbrook, 1987). Consequently, marketing concepts based upon referral within social network, such as electronic word-of-mouth (eWOM) marketing, relationship marketing, and viral marketing (De Bruyn & Lilien, 2008; Ferguson, 2008; Helm, 2010), have been considered a relevant marketing strategy in the social media era, particularly for service and experience providers such as tourism and hospitality. The approach to these marketing concepts suggests that marketers can leverage the power

of interpersonal networks to promote their products and services by transforming the communication networks into influence networks (i.e., using social network for referral marketing).

These marketing concepts are based on social network theories emphasizing information seeking, sharing and adoption behaviour among members of online communities. Most studies on social influence through online social networks presume consumers' need of information and referral to make informed decisions or to validate certain consumption behaviour. Hence, research focus has been on whether or not information seekers are willing to follow others' recommendations based on the characteristics of information, information providers, and social ties between information seekers and providers (Brown & Reingen, 1987; Steffes & Burgee, 2009). Lacking, though, is research on social influence resulting from communication among peers in online social networks beyond referral, but through the processes of social comparison (i.e., often labelled as peer influence or social contagion), which is an important characteristic of social life. As social influence through comparison implies different marketing strategies for product and service providers, it is important to conceptualize this social media phenomenon further.

Social comparison, which is the act of comparing one's features to those of others and vice versa (Buunk & Mussweiler, 2001), has been a research interest in various social context, including health and wellbeing as well as productivity and career performance, since the conceptualization of social comparison theory by Festinger (1954). It is suggested that when making decisions, people might be comparing their choices with others and are sensitive to the social comparison cues (Bearden & Rose 1990), including social reactions to their choices. However, studies on social comparison in terms of consumption behaviour through online social networks are still scant. Among the notable few studies on social comparison online are the works of Loewenstein (1991) as well as Wu and Lee (2008). Identifying the reference shifts from physical to virtual social contexts, Wu and Lee (2008) uncovered that consumers have stronger purchase intention when social comparison is present.

Due to the prevalent use of social media, particularly social networking sites (SNS), for day-to-day communication, it is important to explore how interpersonal communication on SNS manifests in social influence. Hence, the goal of this study is to investigate social influence resulting from peer-to-peer interactions on SNS. Specifically, the study aims at: (1) defining social influence from referral and comparison, (2) identifying the relationship between the extent of communication on SNS and the adoption of social influence behaviour, and (3) investigating the role of competitiveness and social status in shaping the relationship between the extent of communication and social influence. The scope of this study is on peer-to-peer communication (i.e., in-group reference point) in order to focus on the above objectives.

2 Literature Review

2.1 Social Comparison and Contagion

Social comparison theory is the idea that there is a drive within individuals to look to outside images in order to evaluate their own opinions and abilities (Festinger, 1954). According to the theory of reasoned action (TRA) (Fishbein & Ajzen, 1975) and its extension, theory of planned behaviour (TPB) (Ajzen, 1985), subjective norms, the influence of people on one's social environment, is believed to be an important construct that influences his/her intention to adopt a certain behaviour. Leenders (2002) argues that people are appropriately taking into account the opinions and behaviours displayed by others, combined with the considerations of other constraints and opportunities, to establish their own opinion and behaviour. He labelled this process 'social contagion' (Leenders, 2002). Also, based on their study on telecommuting behaviours, Wilton, Páez and Scott (2011) identify that when presented with a choice people refer to the experiences of others to make informed decisions.

When comparing themselves with others, people tend to select a person or group to serve as a point of comparison or as a reference group (Khan & Khan 2005, Schiffman & Kanuk, 2000). Leenders (2002) identifies two distinct processes that lead to social contagion: (1) *communication*, when people use others with whom they are directly tied as their frame of reference, and (2) *comparison*, when people use others they feel similar to as their frame of reference (Leenders, 2002). *Communication* implies direct contacts between people and their influencers. On the other hand, according to Burt (1987), *comparison* is triggered when people are in competition with one another. In this sense, when people compete with one another, they use each other as reference frame through indirect communication (e.g., displayed behaviour).

Therefore, it can be hypothesized in this study that the extent of engagement in SNS, which results in direct and indirect communication, will lead to two different types of social contagion: referral-based and competition-based influence. A person who is well connected with others on SNS will refer to their friends for (a) recommendations to follow and (b) displayed behaviours to compare her/himself to. The more a person is engaged in SNS, the more she/he is exposed to recommendations from others and likely to follow them. In the case of social competition, a high level of engagement in SNS means more access to displayed behaviours of others and leads to the consumption behaviour that will likely put her/him ahead of others on SNS. The following hypotheses are suggested:

Hypothesis 1a: There is a direct positive relationship between the level of engagement on social networking sites and the level of referral-based social influence.

Hypothesis 1b: There is a direct positive relationship between the level of engagement on social networking sites and the level of competition-based social influence.

2.2 Competition and Status Consumption

A few scholars have addressed the question as to when a social actor is most strongly influenced by their peers; most have pointed out competition as the important mechanism of social influence (Bothner, 2003; Burt, 1987). Indeed, Burt (1982; 1987) argues that social actors compete and, thus, monitor and affect each other's choices. The effect of social comparison exists due to the fact that people are concerned or care about reactions of others/reference groups (Bearden & Rose, 1990). The theory of self-concept (Grubb & Grathwohl, 1967), which is stated to be the result of consumers' interactions with their peers, indicates that consumers value consumption that results in recognition and reinforces reactions from the social network so as to strengthen the conception about themselves. Consumer behaviour literature refers to these attitudes as conformism vs. snobbism (Amaldoss & Jain, 2005; Corneo & Jeanne, 1997). Conformism in consumer behaviour typically refers to consumption behaviour that follows conventional standards, leaning towards social acceptance. On the other hand, snobbism refers to consumption behaviour as a way of establishing superiority within social networks.

Based on their study among members of SNS, Iyengar, Han and Gupta (2009) identified that consumers with different attitude toward of social status react differently on friends' online purchase. They high-status consumers are well connected, but react negatively to friends' purchase. On the other hand, low-status consumers are not influenced by it. Mid-status consumers react positively towards friends' influence and exhibit the behaviour of "keeping up with Joneses" (Iyengar, Han & Gupta, 2009). High-status consumers are often characterized with conspicuous consumption (Amaldoss & Jain, 2005; Corneo & Jeanne, 1997), where display of consumption behaviour is seen as a means of attaining and maintaining social status.

It can be summarized from the literature that the attitude toward status consumption will moderate the influence of the displayed behaviour of others. To confirm the findings from Iyengar, Han, and Gupta (2009), it is hypothesized that attitude toward status consumption negatively moderates the relationship between engagement on SNS and the level of referral-based influence. Additionally, since competition signifies the attainment and maintenance of social status, it is hypothesized that attitude toward status consumption positively moderates the relationship between the level of engagement on SNS and competition-based influence. The following hypotheses are suggested:

Hypothesis 2a: Consumers' attitude toward status consumption negatively moderates the relationship between the level engagement on social networking sites and the level of referral-based social influence.

Hypothesis 2b: Consumers' attitude toward status consumption positively moderates the relationship between the level engagement on social networking sites and the level of competition-based social influence.

3 Methodology

3.1 Research Design

The theoretical framework of this study is illustrated in Fig. 1. The goal of this study is to identify the relationships between the level of engagement on SNS and the level of referral-based and competition-based social influence and to identify the moderating effects of status consumption. This framework was tested in the context of online interactions among members of SNS and its influence on restaurant selection and experience.

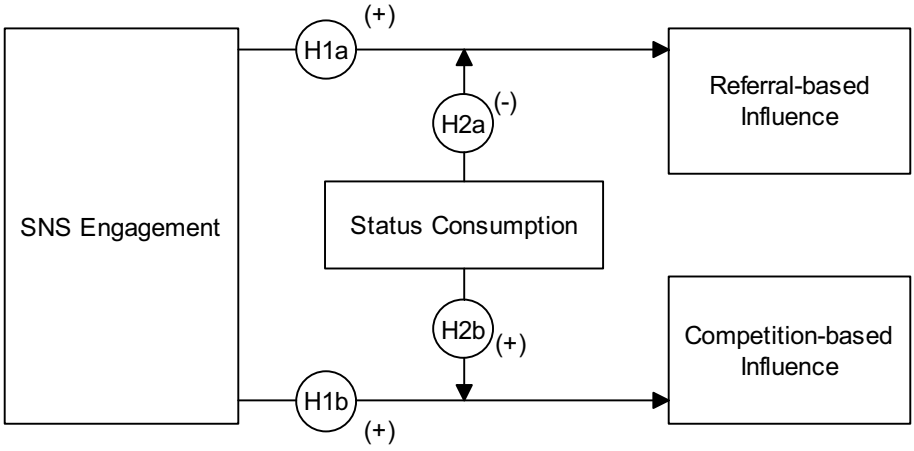


Fig 1. Foundation of the Study

Established constructs from literature were used to represent items measuring SNS engagement (Tang, 2011) based on the dimensions of communication and social connection on SNS participation level (Parent, Plangger & Bal, 2011). To measure status consumption, the dimensions of social prestige and conformity to consumption trend (Zhou, Teng & Poon, 2008) were adapted.

These two dimensions load into one factor (labelled Status consumption) with Cronbach's Alpha of .946. Items measuring referral-based and competition-based influence were developed following an extensive literature review, focus group discussions, and expert evaluation. Items were presented in 7-point Likert-type scale with Agree–Disagree anchor statements.

Table 1. Measurement Items

Constructs – Items	Internal Consistency (Cronbach's Alpha)
Engagement on SNS (Mean: 4.23; S.D.: 1.59)	.942
I share my experiences regularly on SNS	
My friends and I comment on each other regularly on SNS	
My friends and I converse regularly on SNS	
My friends and I are well connected on SNS	
Status Consumption (Mean: 3.26; S.D.: 1.33)	.946
I select restaurants…	
…that signify my trendy image.	
…that symbolize my social status.	
…that are associated with the symbol of prestige.	
…that represent the latest lifestyle.	
…that make me have good impression on others.	
…that make me feel good in my social group.	
Referral-based Influence (Mean: 4.84; S.D.: 1.29)	.886
I would try a new restaurant if my friends posted on SNS that they have been there.	
I would try a new restaurant if my friends positively raved about it on SNS.	
I would frequent a particular restaurant if my friends posted on SNS that they do so.	
Competition-based Influence (Mean: 2.77; S.D.: 1.50)	.914
I would try new restaurants to be able to post new and exciting updates ahead of my friends on SNS.	
I would dine out more often if my friends posted on SNS that they do so.	
I would try new restaurants to keep up with my friends on SNS.	

3.2 Data Collection and Analysis

A web-based survey was administered for data collection. The survey questions consist of four sections. In Section One, respondents were asked about their typical use of SNS, which include the types of SNS applications, devices used and venues for using SNS. In Section Two, respondents were asked about their network engagement using SNS and the state of opinion leadership. In Section Three, respondents were asked to state their behaviour toward restaurant selection that count for conspicuous consumption and state their behaviour toward network influence on restaurant selection. Lastly, several demographic control variables were collected.

An invitation to participate in the online survey was sent in August 2011 to 5,000 Americans randomly selected from a tourism email list. All recipients requested travel information about the US Midwestern states over the past 3 years. An incentive to win a $100 or one of two $50 or one of two $25 dining certificates from *restaurant.com* was provided. Following three reminders, 232 responses were collected (4.6% response rate). A total of 158 complete responses were included in the analysis. To measure the hypothesized relationships and the moderating effect of status consumption, regression analysis was performed.

4 Result and Discussion

The majority of respondents (74.4%) were female. In terms of age, respondents were slightly older, with 73.9% of them older than 45 years (33.3% 45-54 years old, 24.8% 55-64 years old, and 15.7% over 65 years). The rest of respondents were 35-44 years old (15.7%) and 25-34 years old (10.5%). Most respondents were highly educated, with 49% attended or graduated college and 32% had a graduate/advanced degree. Around 97% of respondents resided in the US during data collection.

As illustrated in Fig. 2, respondents reported an overwhelming majority of *Facebook* use (86.2%). Following are the use of *LinkedIn* (19.5%), *YouTube* (13.8%), *Twitter* (12.6%), and personal blogs (8.8%). Respondents also reported the use of other SNS such as *Tumblr*, *Foursquare* (on mobile devices), and email. In terms of devices used for communicating on SNS, 98.1% respondents used personal computers (including laptops). Some respondents used mobile phones (36.1%) and a small number of respondents (10.8%) used tablet PCs, such as iPAD, to access SNS.

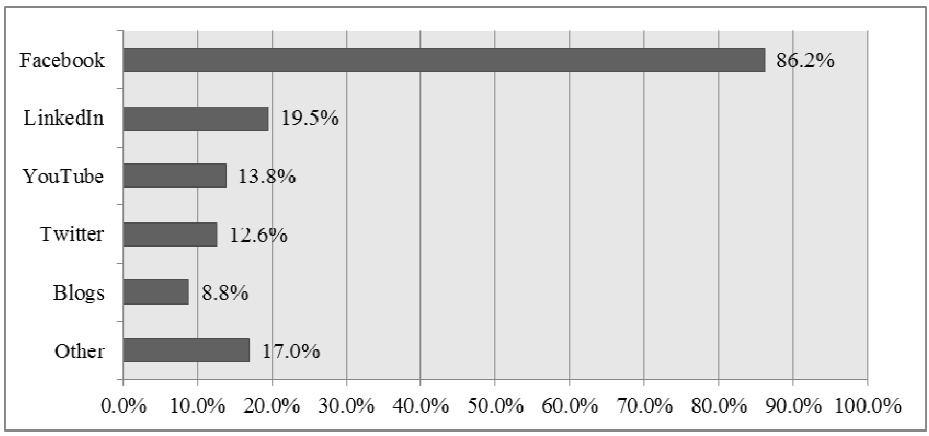


Fig 2. The Types of SNS Used by Respondents

The majority of respondents stated that they dine out frequently. About 30% respondents dine out 2-3 times a week, 28.5% once a week, and 28% 2-3 times a month. Only about 13% respondents stated they dine out once a month or less. Most respondents are highly engaged in SNS (Mean: 4.23; S.D.: 1.59), but exhibit a low degree of status consumption (Mean: 3.26; S.D.: 1.33).

4.1 Referral-based Influence

Based on the composite mean of the referral-based construct (Mean: 4.23; S.D.: 1.59), it was identified that, influenced by their friends' recommendation, respondents tend to be willing to consume the same thing as their friends (e.g., visiting the same restaurants). Further, based on the results of regression analysis (see Table 2), the level of engagement on SNS positively affect the level of referral-based influence (*Hypothesis 1a* supported). The base model (Model 1) indicates a statistically significant relationship with 28.1% variance in the referral-based influence explained by model. Model 2 represents the direct effect of engagement as well as the moderating effect of status consumption on the relationship between engagement and referral-based influence. The results show that the negative moderating effect of status consumption is not statistically significant (*Hypothesis 2a* not supported). It can be suggested that consumers who are highly connected with their peers on SNS are most likely to exhibit consumption behaviour based on recommendations from their friends.

Table 2. Regression Models for Referral-based Influence (N = 158)

	Model 1	**Model 2**
R^2	.281	.284
F (Sig.)	60.312 (.000)	30.362 (.000)
Independent Variables (Beta (Sig.))		
Engagement	**.530** (.000)	**.529** (.000)
Engagement* Status Consumption		-.052 (.442)

4.2 Competition-based Influence

Based on the composite mean, it was identified that respondents tend not to receive competition-based influence (Mean: 2.77; S.D.: 1.50). Only a few respondents stated they would consume more (i.e., visit more restaurants or dine out more often) in order to compete with their friends. Table 3 represents the results of regression analysis for competition-based influence. The level of engagement on SNS positively affects the level of competition-based influence (*Hypothesis 1b* supported). The base model (Model 1) indicates a statistically significant relationship with the variance explained of 21.2%.

Table 3. Regression Models for Competition-based Influence (N = 158)

	Model 1	**Model 2**
R^2	.212	.557
F (Sig.)	41.380 (.000)	24.405 (.000)
Independent Variables (Beta (Sig.))		
Engagement	**.480** (.000)	**.466** (.000)
Engagement* Status Consumption		**.173** (.013)

Model 2 represents the moderating effect of status consumption on the relationship between engagement and competition-based influence. This model has a better fit with the variance explained of 55.7%. The results show a positive moderating effect of status consumption (*Hypothesis 2b* supported). It can be concluded that interpersonal communication and engagement on SNS prompts social influence due to the process of social comparison. This effect is magnified among consumers, who put a high value on status consumption.

5 Conclusion and Implication

This study explores social influence by way of social competition and comparison enabled by the facilitation of global interpersonal communications and connectivity on social media. Specifically, this study investigates how consumers' engagement in SNS and the level of connectedness with their peers lead to two types of social influence: referral-based influence (i.e., consumption behaviour based on friends' recommendations on past behaviour) and competition-based influence (i.e., consumption behaviour to stay ahead of friends). The results from the analysis supported these two hypothesized relationships, confirming that social media provide not only potentials for electronic word-of-mouth in terms of referrals and recommendation, but also clues for social comparison (e.g., displayed consumption behaviour) that spurs competition.

This study further confirms that displayed consumption behaviours online would generate two effects: similar consumption by those who follow recommendations and consumption by those competing with the information providers. Consequently, marketers can capitalize on this opportunity by creating a competitive arena where consumers have more opportunity to display their consumption behaviour online through platform development for instant updates for ease of sharing. In doing so, marketers can encourage consumers to be the leaders in their online social networks.

Furthermore, this study also investigates the moderating effect of consumers' attitude toward social status on the relationships between engagement on SNS and social influence. The moderating effect on competition-based influence is supported, confirming that the consumers' competitiveness positively magnify the tendency for competition-based influence. In that, when consumers tend to consume products and services to signify their social status online, they use their peers' displayed consumption behaviour as reference frame for them to compete. However, the negative moderating effect of status consumption on referral-based influence is not supported. The cause for this issue might be the low number of respondents or the nature of restaurant consumption, not differentiating between luxury and conventional brands, as a base for competition.

The scope of this study context poses a limitation to the interpretation of the research findings. This study focuses on restaurant visitation as consumption stimuli due to the simplicity of the situation to test the hypothesized model. As the study focuses on social influence resulting from communication and comparison processes only, other factors that might influence visitation behaviour, such as those associated with consumption situation, restaurant quality, etc., are excluded from the analysis.

Future study should include these factors to control for the unidentified forces of social influence in this study. Furthermore, future research should be directed to other types of consumption situation, particularly destination choice, to retest the applicability of the measurement items and confirm the hypothesized model.

References

Ajzen, I. (1985). The theory of planned behavior. *Organizational Behavior and Human Decision Processes* 50: 179-211.

Amaldoss, W. & Jain, S. (2005). Reference groups and product line decisions: An experimental investigation of limited editions and product proliferation. *Management Science* 56 (4): 621-644.

Bearden, W.O. & Rose, R.L. (1990). Attention to social comparison information: An individual difference factor affecting consumer conformity. *Journal of Consumer Research* 16: 461-471.

Bothner, M.S. (2003). Competition and social influence: The diffusion of the sixth generation processor in the global computer industry. *American Journal of Sociology* 108(6): 1175-1210.

Brown, J. J. & Reingen, P. H. (1987). Social ties and wordof-mouth referral behavior. *Journal of Consumer Research* 14: 350-62.

Burt, R.S. (1982). *Toward a Structural Theory of Action: Network Models of Social Structure, Perception, and Action*. New York: Academic Press.

Burt, R.S. (1987). Social contagion and innovation: Cohesion versus structural equivalence. *American Journal of Sociology* 92: 1287-335.

Buunk, B. P. & Mussweiler, T. (2001). New directions in social comparison research. *European Journal of Social Psychology* 31: 467-475.

Corneo, G. & Jeanne. O. (1994). *Conspicuous Consumption and the Existence of Upward Sloping Demand Curves*. Discussion Paper Serie A 461. University of Bonn, Germany

De Bruyn, A. & Lilien, G. (2008). A multi-stage model of word-of-mouth influence. *International Journal of Research in Marketing* 25: 151-163.

Ferguson, R. (2008). Word of mouth and viral marketing: taking the temperature of the hottest trends in marketing. *Journal of Consumer Marketing* 25(3): 179-182.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations* 7: 117-140.

Fishbein, M. & Ajzen, I. (1975). *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*. Reading, MA: Addison-Wesley.

Grubb, E. L. & Grathwohl, H. L. (1967). Consumer self-concept, symbolism and market behaviour: A theoretical approach. *Journal of Marketing* 31(4): 22-27.

Harrison-Walker, L. J., (2001). The measurement of word-of-mouth communication and an investigation of service quality and customer commitment as potential antecedents. *Journal of Service Research* 4(1): 60-75.

Helm, S. (2010). Viral marketing – Establishing customer relationships by „word-of-mouse. *Electronic Markets* 10(3): 158-161.

Hennig-Thurau, T., Gwinner, K.P., Walsh, G. & Gremler. D.D. (2004), Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the internet? *Journal of Interactive Marketing* 18 (1): 38-52.

Hennig-Thurau, T. & Walsh, G. (2004). Electronic word-of-mouth: Consequences of and motives for reading customer articulations on the internet. *International Journal of Electronic Commerce* 8 (2): 51-74.

Iyengar, R., Han, S. & Gupta, S. (2009). Do Friends Influence Purchases in a Social Network? *Harvard Business School Working Paper* 09-123.

Jansen, B. J., Zhang, M., Sobel, K. & Chowdhury, A. (2009). Twitter power: Tweets as electronic word of mouth. *Journal of the American Society for Information Science and Technology* 60(11): 2169-2188.

Khan, G. & Khan, N. (2005). Susceptibility to informational social influence on purchase decisions of designer label apparel: The mediating role of gender. *The Business Review* 4(1): 32-36.

Kietzmann, J. H., Hermkens, K., McCarthy I. P. & Silvestre, B. S. (2011). Social media? Get serious! Understanding the functional building blocks of social media. *Business Horizons* 54:241-251.

Leenders, R. T. A. J. (2002). Modeling social influence through network autocorrelation: Constructing the weight matrix. *Social Networks* 24: 21-47.

Loewenstein, G. F. (1988). Frames of mind in inter-temporal choice. *Management Science* 34(2): 200-214.

Parent, M., Plangger, K. & Bal, A. (2011). The new WTP: Willingness to participate. *Business Horizons* 54(3): 219-229.

Steffes, E. M. & Burgee, L. E. (2009). Social ties and online word of mouth. *Internet Research* 19(1): 42-59.

Tang, J. C. (2011). *Consumer Engagement in Mobile Social Networking Activities*. Unpublished Masters Project. Temple University, USA.

Westbrook, R. A. (1987). Product/consumption-based affective responses and postpurchase processes. *Journal of Marketing Research* 24: 258-270.

Wilton, R.D., Páez, A. & Scott, D. M. (2011). Why do you care what other people think? A qualitative investigation of social contact and telecommuting. *Transportation Research Part A: Policy and Practice* 55(4): 269-282.

Wu, L.-L. & Lee, L. (2008). Online social comparison: Implications derived from Web 2.0. Proceedings of the 12th Pacific Asia Conference on Information Systems. Suzhou: July 3-7, 2008.

Zhou, L, Teng, L. & Poon, P.S. (2008). Susceptibility to global consumer culture: A three-dimensional scale. *Psychology & Marketing* 25(4): 336–351.

ConTur: An Intelligent Content Management System for the Tourism Sector

Germán Herrero Cárcel[a], Aida Campo[b], Maria Gil[b], Ana Garcia[c], David Martin, Ion Zugasti[d], Sonia Bilbao[e], Amalia Perez[f], Hristo Koshutanski[g], Antonio Maña[g] , and Iñaki Perez De Albeniz[h]

[a] Atos Research & Innovation, Atos Spain, Spain
{german.herrero}@atos.net

[b] Andago Labs,Andago, Spain
{aida.campo, maria.gil}@andago.com

[c] Dome Consulting, Spain
{agarcia}@dome-consulting.com

[d] Cooperative Research Center in Tourism, CICtourGUNE, Spain
{david.martin, ion.zugasti}@tourgune.org

[e] Tecnalia Research & Innovation, Spain
{sonia.bilbao}@tecnalia.com

[f] Universidad de Burgos, Spain
{apromero}@ubu.es

[g] Computer Science Department, Universidad de Malaga, Spain
{hristo,amg}@lcc.uma.es

[h] Barcelo Viajes, Spain
{i.perez}@barceloviajes.com

Abstract

Nowadays, tourists have access to a lot of different web sites in order to find information about destinations. This has a direct impact on tourism destination management organizations and tourism providers. These stakeholders have a lot of problems in order to find reliable and up-to-date information about their destinations. This paper presents an innovative solution that helps to promote tourism destinations through an intelligent, reliable and comprehensive content management platform that integrates different online heterogeneous and distributed information sources. This platform is described and validated in two different scenarios: "Portal Destino 3.0", a tourism portal managed by a public organization and "Mi Guia Multidestino", a personalized online web travel guide.

Keywords: intelligent content management system, semantic annotation, intelligent travel guides, semantic content integration.

1 Introduction

Although often underestimated as a marginal sector, the tourism sector plays a strategic role with a considerable market share for the European continent. Europe is still the leading tourism market worldwide with more than half of the global market

share (488 million arrivals) in 2008[1]. The importance of Tourism as a strategic market sector is also underpinned by its share of the European GDP (5% of the total European GDP is produced directly and 10% of it indirectly by Tourism) as well as the number of jobs depending on it (7-8M jobs directly and 20M jobs indirectly). However, the tourism sector has its strengths and weaknesses. Some of the most relevant weaknesses in the European and Spanish tourism are:

- Emerging destinations growing rivalry (problems of competitiveness)
- Weak sector integration: the value-chain of the tourism sector is weakly integrated due the emergence of Internet and the fragmentation of tourism service providers.
- Nowadays, one of the trends that has had most effect on the sector in the recent years is the ever-growing number of website pages providing tourism information and the use of social-media that is starting to pervade every single internet services and defines the online user behaviour and consumer decisions

To solve the above-mentioned issues is more necessary than ever to develop a system that helps to promote tourism destinations through and efficient management of content partners and integrate the actors of the tourism value-chain related to the destination. ConTur faces these challenges and its main objective is to provide an intelligent, reliable and comprehensive management of different online heterogeneous resources (personal websites, social media, documents and other resources). ConTur platform has a global reach within the tourism sector. The benefits provided and the problems it intends to address concerns all actors which are part of the value chain (public tourism authorities, service providers, online agencies, tourists).

2 Background

Currently, the World Wide Web (WWW) is based mainly on documents written in HTML, a mark-up language used to create hypertext on the Internet. The Semantic Web aims to overcome the limitations of HTML by introducing explicit descriptions of meaning, the internal structure and the overall structure of content and services available on the WWW (Berners-Lee, 2001). For this, the Semantic Web technologies mainly used are RDF, SPARQL, and OWL, as well as XML, mechanisms to help turn the Web into a global infrastructure where it is possible to share and reuse data and documents between different types of users (W3C).

In the particular case of the tourism sector, where different kinds of information, provided by different sources are generated, ontologies can be applied to improve interoperability between different sources provided by the actors that take part in the tourism sector. Ontologies act as another tool in the content management systems or knowledge management. They transform the information, and increase it with structures that give content and meaning to information.

Moreover, at present, exchanging information between companies is based on the different standards. These standards have been used as agreements between different actors in a given sector to ensure that components made by different organizations can

[1] See: European Travel Commission – Fact Sheet, September 2009:
http://www.etc-corporate.org/resources/uploads/ETC_Factsheet_20090831.pdf

be mounted on a set and work in coordination. After several years of standardization, leading companies are beginning to adopt Service-Oriented Architecture (SOA) to integrate disparate data, providing the availability of data using Web services. Web services provide standardized and open protocols to provide a means unified access to information from a diverse set of platforms (different operating systems, programming languages, and applications). These Web services can be reused to provide new services and applications within the same organization or between different, generating a high flexibility for the business.

Finally, a key aspect of the current web is the emergence of the concept of *mashup*. Mashups technologies can be used through the philosophy of Software as a Service (SaaS) (Piñero, 2007). A mashup is a web application that uses resources from more than one source to create a complete service. The data used by the *mashup* is normally obtained from third parties, accessing public interfaces or APIs they provide, through interoperable items, which are integrated through syndication tools.

Thus, it can be concluded that the application of technologies based on the Semantic Web can improve access to tourist information integrated from multiple sources of information scattered (Murua, 2005).

3 Objectives

The main objective of ConTur is to develop a Content Management System for tourism destinations to enable tourism stakeholders (e.g. SMEs) to be present on the Internet directly to travellers or receptive businesses. ConTur platform integrates content from different sources: tourism companies at destination, unstructured information on the Internet (e.g. social media, blogs), DMOs and local institutions.

With ConTur, any stakeholder can filter, catalogue, aggregate, merge and integrate tourism related contents coming from different, heterogeneous and distributed web sites (e.g. private web sites, social networks, institutional catalogues, etc.) in an intelligent, reliable and robust way.

The scope and purpose of the platform is multiple and varies depending on the stakeholder of the tourism value chain. In the case of the institutional agencies of tourism, its aim is to encourage the access to their content (e.g. brochures, videos, cultural agendas, etc.) and the promotion of the destination itself through different indirect channels (e.g. online agencies, social network of travellers, etc.). Similarly, they may rely on the platform to enrich their portals with the inclusion of non-institutional content available on the Web. Destination service providers will increase direct sales channels due to the contents being captured in the Web, and will facilitate the dissemination and promotion of its services through indirect channels. In the case of online agencies, they will be able to offer value added services, based on the creation of personalized contents. The offer will become more attractive thanks to the inclusion of content from the destination. Finally, tourists, consumers and producers of the tourism sector, will be able to choose their destination and prepare a more exhaustive travel plan thanks to the information that is acquired from different sources.

4 ConTur Architecture

ConTur has to acquire and unify content from different sources that is usually structured using different data formats. This way, the proposed architecture has to face the following technological challenges:

- Data structures that facilitates the description of tourism contents through a tourism ontology.
- Web interfaces to capture and annotate tourism content (automatically and semi-automatically).
- Mapping mechanisms that allow the transformation of content's metadata already described by their suppliers to the ConTur ontology.
- Ontology extension mechanisms to facilitate the identification of new tourism concepts that were not initially included in the ontology.
- Reasoning engines to classify and infer new data based on the acquired contents.
- Content aggregation engines that fuse content selectively according to the requirements of the recipient.
- Web interfaces in order to facilitate the visualization of content by different users.
- Security and confidence models that allow the definition of access policies and the allocation of degrees of reliability to the content.

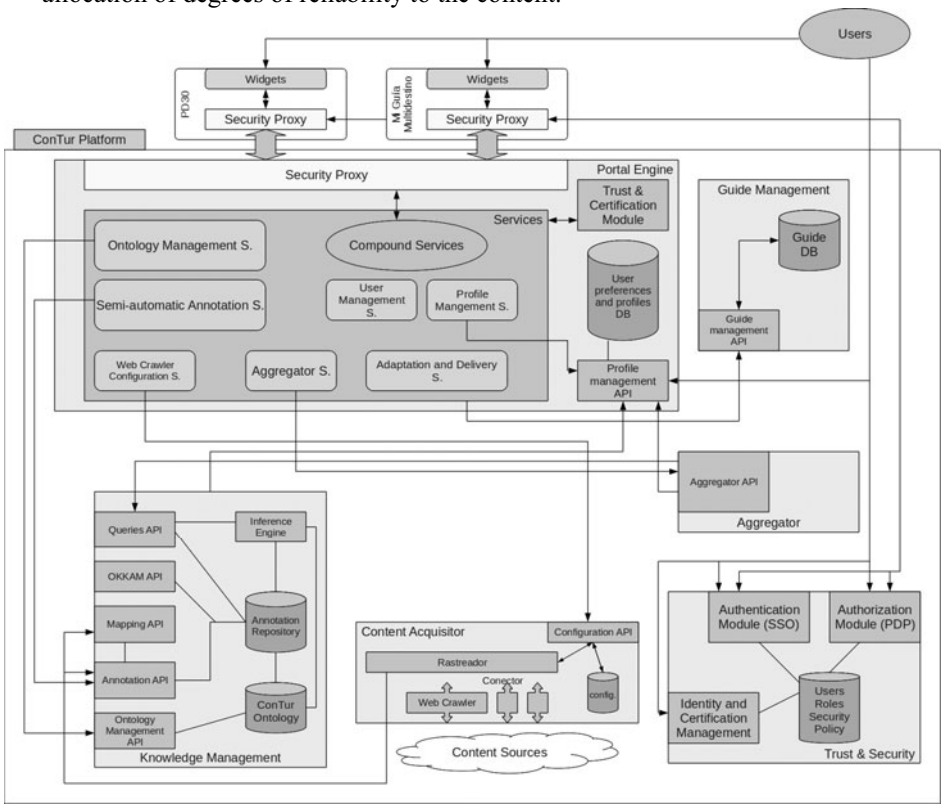


Fig. 1. ConTur Architecture

As shown in Fig.1, ConTur is divided into six different modules. Each of these modules is described in the following sections.

4.1 Portal Engine

The portal engine is the entry point to the platform. Here is where services basic and compound services are found, so that users from different use cases have the opportunity to access the functionality offered by the platform.

ConTur platform contains a Content Management System (Generic Access Portal) from which widgets that can be embedded in other web pages can be downloaded. These widgets are web user interface components that allow the communication with the platform's services. Each widget communicates with a compound service, which can communicate with other simple services (or *mashups* of simple services). Communication between widgets and services is carried out through SOAP (Simple Object Access Protocol) web services and through security proxies.

In addition, in this Content Management System, a management interface for administrators is provided, and includes the following functions: registration of new users, configuration of the crawler (including the web pages relevant to the tourism domain and other parameters), extension of the ontology, etc. Beta version of Portal Contur (the Generic Access Portal of ConTur's project) can be accessed at this URL: http://portalcontur-demo.andago.com/

4.2 Knowledge Management

This module manages all the information that is gathered from tourism web sites. This way, the information that is annotated by the users of the platform is sent to the Knowledge Manager. Here, this data is saved and merged into the triple store, avoiding data inconsistencies and redundancy. This information can be queried in order to be used by the rest of the modules of the platform.

Data is structured in RDF format according to the defined semantic model. Two different stores have been configured. The first one contains the structure of the ontology (TBox) and the second one contains all the annotations in the form of triples (ABox). The module has been implemented using the Jena semantic framework .

The Knowledge Manager is divided into several modules that expose high level APIs to the rest of the platform.

- Ontology Management API: it is used to extend the proposed ontology. Like that, new classes and properties can be created and modified. This module has an extra store where all the created/modified classes and properties are stored in order to be validated by the administrator.
- Annotation API: this module receives the created annotations and inserts them into the repository of annotations. It uses the Okkam entity name system (http://www.okkam.org) in order to assign or reuse unique identifiers to the triple instances.
- Mapping API: this is used to create mappings between web contents that are already structured using ontologies and the ontology used by the platform.

- Query API: it is used to retrieve data from the Knowledge Manager using the SPARQL query language.
- Inference engine: this module is used to make inferences based on the asserted triples and the defined rules. The inference process generates new triples that are also stored in the repository of annotations.

The created ontology has been designed according to the state of the art in the field of tourism semantic modelling (Prantner, 2007; Werthner, 2009). This ontology is defined using the OWL ontology language and contains 153 classes and 116 properties. It uses other ontologies like the w3c time (http://www.w3.org/2006/time) and geo (http://www.w3.org/2003/01/geo/wgs84_pos) ontologies in order to model the time and the location entities.

4.3 Content Acquisitor

This component is responsible of collecting touristic contents stemming from different and heterogeneous sources scattered around the web, such us private web pages, social nets, institutional catalogues, etc. and annotate them automatically based on the ontology of ConTur. Content *acquisitor* also includes semi-automatic annotation system to accomplish editorial revisions over annotations performed automatically, which combines the active learning and self-training in order to reduce the efficiency of the annotation process.

The Content *acquisitor* is composed by the following modules; some interact with APIs that reside inside ConTur platform:

- Crawler configuration API: This component allow to ConTur administrator to set parameters to configure the crawler, which are stored in DB configuration. It can be set the seed URLs from which the crawler begins or the refreshing interval to re-crawl the pages among other properties.
- Crawler: This process crawl the web to recollect sources of unstructured format (mainly HTML) with tourism value. It starts from the seed URL's and crawl connected pages so long as they continue to be considered relevant to tourist domain topic. The relevance of a page is determined after downloading its content. To calculate the relevance to a predefined set of topics web documents are analyzed by annotation API. Relevant pages are sent to content indexing and their contained URLs are added to the crawl frontier, pages that fall below are discarded. ConTur Crawler has been implemented by building a focused (vertical) crawling mechanism on top of Nutch open source crawler, which uses MapReduce distributed model to make the crawler scalable.
- Text processing: Text is processed to detect semantically related concepts already present in ConTur domain ontology. It uses annotation API to do the Natural text processing.
- Annotation repository: Extracted touristic concepts are annotated against ConTur domain ontology and annotations are stored on ConTur repository by calling annotation API.
- Semi-automatic annotation API: This component provides results from automatic annotation process. A user can inspect these results and refine or correct them if required.

The semi-automatic annotation API is based on Freeling API (Freeling, 2010), an open source suite of language analyzers. The annotation API receives the corpus from the crawler and executes a natural language process through a text tokenization and entity detection. When the entities are identified, the API produce a query against the repository to check if the entity already exists, in order to have the URI of the semantic entity from the ConTur ontology. The new annotation is added to the repository in the form of a triple as follow:

{<URLsource>, <ConTur:isAbout>, <ConTur:SemanticEntityfound>}.

If the entity identified in the corpus is not in the repository, the Annotation API uses the Okkam entity name system in order to assign a unique identifier to the triple instance before add the annotation to the repository. Based on Freeling capabilities, the current version of the Annotation API supports languages as Spanish, English, Italian or Portuguese. Ontologies are a good modelling approach in order to provide a common vocabulary and harmonize information coming from heterogeneous sources. Furthermore, reasoning techniques can be applied in order to infer new knowledge from explicit information that can be valuable for the rest of the platform.

4.4 Aggregator

As the portfolio of products and services of a destination is so numerous, it becomes increasingly important to offer tools that help travellers find those that match their demand and expectations.

The aggregator is a search engine that returns the set of tourism products and services that match the user preferences as expressed in a query. It has been designed as a configuration tool of products and services. By configuration we understand the definition given by Mittal & Frayman (Mittal et al, 1989).

> *"Given: (A) a fixed, pre-defined set of components, where a component is described by a set of properties, ports for connecting it to other components, constraints at each port that describe the components that can be connected at that port, and other structural constraints (B) some description of the desired configuration; and (C) possibly some criteria for making optimal selections.*
>
> *Build: One or more configurations that satisfy all the requirements, where a configuration is a set of components and a description of the connections between the components in the set, or, detect inconsistencies in the requirements."*

In the case of the tourism domain the components can be hotels, holiday packages, flights, restaurants, car rentals, etc. Each component is described by a set of properties such as the name, rating, popularity, price, brand, amenities and so on.

Components are associated or "connected" to a destination and to other resources. These resources can be multimedia resources (audio, photo or video) or websites (blog, social network or official website).

A semantic model has been defined in order to translate into a machine readable description the query performed by the user. The tourist preferences are expressed as constraints on the components properties. These constraints are used to filter the amount of results. In addition, the semantic model allows defining ranking criteria to sort the results returned by the aggregator. Finally, the model allows specifying the

data to be retrieved for each type of component that is part of the query's results, e.g. name and address for hotels, description for activities and price for car rentals. (Bilbao, 2010).

4.5 Guide Management

This module provides the CRUD functionality (Create, Read, Update, Delete) to manage travel guides. Each travel guide is composed by a list of destinations and a period of time in which to visit each destination. By means of the aggregator, the user can search for tourism products and services associated to each destination. If the product returned is of interest for the user, he can add it to his travel guide. Users can decide the level of access security for each of their guides: private, protected or public.

4.6 Trust & Security

The ConTur platform is designed to offer touristic contents to end users via a federation of Web portals. Given that, it was developed a security architecture providing a *circle of trust* between Web portals and the ConTur platform, where both Web portals and ConTur platform have to be certified by a ConTur certification authority to be part of the circle of trust.

The security solution builds upon single sign-on (SSO) for federations and provides single authorization for federated content provisioning. It was adopted SAML (SAML, 2005) specification for building SSO-based federated identity management capable of expressing various authentication contexts such as X.509 (X.509, 2005) certificates – the case of FNMT[1] and DNIe[2] certificates, or credentials of type userID/password.

A *single authorization* protocol allows for multiple unified user accesses to Web portals within the ConTur federation. Users get authorized for service access once (within an SSO session) and for each Web portal of the federation obtain interface access without being required to authorize for the already authorized services.

A ConTur Security Provider (CSP) actor implements "Trust & Security" functionality of Fig.1 as logically independent entity from the platform. The separation facilitates management of security aspects from those of the business logic of the platform. The CSP serves as an anchor of trust for the federation. Web portals trust the CSP for proper user authentication and authorization when users interact with Web portals' interfaces, while the ConTur platform trusts the CSP for certifying ConTur Web portals and providing authorization statements about end-users when Web portals access platform's services on behalf of end users.

Fig. 2 shows a high-level view of security interactions between the main actors. A security proxy was developed to facilitate secure and trusted communications between Web portals and CSP, and between Web portals and the ConTur platform. The proxy encapsulates all authentication and authorization aspects providing high-level security abstraction.

[1] See: http://www.fnmt.es
[2] See: http://www.dnielectronico.es

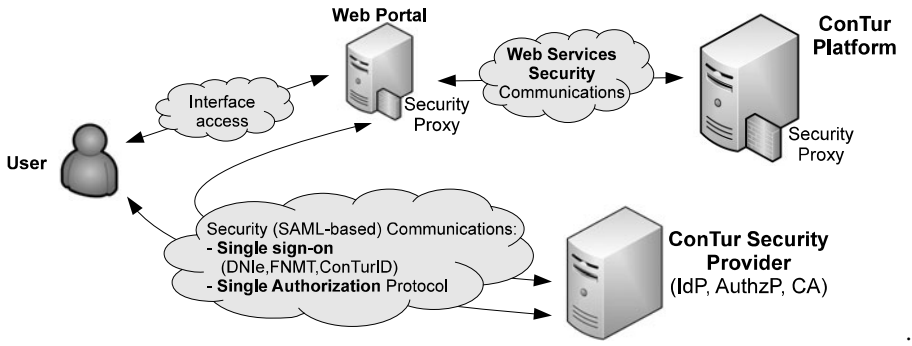


Fig. 2. ConTur Security Communications

5 Scenarios

5.1 Portal Destino 3.0

The main goal of the "Portal Destino 3.0" is to allow the management, access and display of contents of the destination. This site presents the user, through the use of dialog boxes, the contents automatically processed and added by the content manager, whose sources include institutional and non institutional content. This tourism portal is based on the integration of existing information distributed by other portals, institutional or otherwise, on tourism and related topics. For its construction, data sources provided by ConTur platform are used. In this case, ConTur provides a mechanism that allows the integration of automated information on tourism and related topics. Also, ConTur allows an intelligent and adapted integration, to both the site where it integrates, in this case the "Portal Destino 3.0", and to the user's vision.

This is an advanced tourism Website that allows the enrichment of its contents with those produced by non-institutional travellers, hotel chains and any other agent that may have information about the destination. Given the multiplicity of possible sources of content, security mechanisms and reliability, enabled by the content manager, are of particular relevance. Among the features included are:

- Find useful information, reliable, rich and of quality, on the destinations of interest based on key concepts, instead of searching based on keywords.
- Have a centralized inquiry point of information about tourist destinations.
- Request and filter content.
- Add content and update content from other points
- Know the experience of other travellers through social networks

The portal will thus have two potential target users. On the one hand, the public administration that will use the portal for the aggregation and content management and on the other hand, the tourists, that can access the portal via the Internet to view the contents, allowing the user the preparation of their visits in a more practical way.

“Portal Destino 3.0” is the technical basis for the generation of functionalities available to a fully functional Web portal, in which pages or articles that describe a tourist destination that is being offered and promoted through this infrastructure, are shown.

The objective is to be able to present the end-user, a configurable infrastructure with which he can modify the characteristics of access and display of contents offered by the manager. That is, an end user must have the ability to decide which components want, in order to obtain content, data or information, and the way the end user want them displayed. Similarly, this is a requirement for receptive managers that act as intermediaries between the obtaining of the contents of the various sources and the displaying in a particular format to the end user. These content integration capabilities are achieved through interoperable interfaces of items or services that are usually called gadgets or widgets. The experience of use of the “Portal Destino 3.0” is different from the one experienced by visitors of tourist sites today. The information that complements the destinations they are searching, come from annotated sources with semantic values. This allows an automatic system, reasoning if the results are appropriate to the user or not.

The use case "Portal Destino 3.0" in ConTur, has as its main theme, the destination Atapuerca[1] (Burgos, Spain). The use case focuses on aspects of tourism (hotels, restaurants, attractions and activities) around the archaeological site of Atapuerca, and its purpose is to provide contextualized information and of interest to the visitor, through the widgets installed in it. Widgets downloaded from Contur[2] can be installed in Portal Destino 3.0. Some of them are already developed and being tested, as beta prototypes.

5.2 Mi Guia Multidestino

One of the objectives of the travel agencies like Barcelo Viajes is to help customers on their travel plans, trying to supply them with much information as possible about the destinations and other offered services. Besides there is a lot of information of tourist destinations and services, sometimes, this information is old and not complete, so the information provided to the customer has a poor quality.

“Mi guia multidestino” (“My multi-destiny guide” in English) is a web application which address this requirement The main objective of the web application is to provide the functionality to organize the route of the leisure travels based on high-quality information collected from heterogeneous sources of information about destination and tourism services. From the technological basis, travel agencies request a centralized repository that offers up-to-date, integrated and high-quality information about the destinations and tourism services. Moreover, above this repository, the web application should provide functionalities and services to manage the data, users and configuration.

The solution proposed by ConTur is to offer a unique point of reliable and trusted information about the tourism resources. This tool allows search, filter and update of

[1] See: http://www.atapuerca.org/
[2] See: http://contur-demo.andago.com

tourism information in an easy way and results in a cost and time reduction of the travel agencies employees. Furthermore, as the information provided to the customer is of top-quality, the relationship between agencies, stakeholders, destinations and the tourist has a positive impact. Moreover, the communication interfaces of the platform are based on technological and tourism standards to handle the integration between travel agency tools and ConTur.

As indicated, the "Mi guía multidestino" web application provides the travel agencies a useful data information provider about user desires, experiences and behaviors. The results of the data warehouse of this information will allow marketing departments of the travel agencies and destination a better effectiveness of their campaigns. "Mi guía multidestino" is part of the Barcelo Viajes Intranet provided to their business customer partners. Nowadays is at prototype and validation phase, and is predicted to be part of the main Barcelo Viajes web-site[1] in the next months.

6 Impact and Future Research

In the different market reports about the current situation of the tourism sector, one of trends identified is the growing demand of tourism contents by the main source markets. The quality, trust and variety of these contents are the base for a subsequent selection of tourism products by the tourists or final customers. Moreover, these reports also includes the shared idea that the emergence of Internet in the tourism chain-value as the technology used by the tourist to select and buy their holidays and tourism products. Internet has boosted the audience (at an international level) of the tourism products and makes it easier for tourism providers to commerce with their products. Furthermore, Internet positions itself as one of the key component in the tourism product distribution through the whole chain-value. In the last years, several private and public entities are working in the development of new technological standards for the tourism market. This research is focused on the ecommerce through the use of some specifications based on XML messages. Some example of this research is the Travel Technology Initiative (TTI)[2] and Open Travel Alliance (OTA)[3]. In both cases, TTI and OTA, it is identified a lack of content share and data interoperability about tourism between the different actors of the market. ConTur tries to cover part of this lack and bring to the actors and easy way to share contents through its platform and open-source software and services. Because of this and taking into account some figures about the market analysis reports and its trend, it is considered that there is a niche for ConTur, to bring the contents of various service providers to consumers, where the technological solution explained in this paper is the most suitable solution to solve the market request.

7 Conclusions

In this paper the ConTur platform is presented, an intelligent content management system for the tourism sector. ConTur proposes an alternative technological solution for the current challenge of providing integrated added-value access to the content

[1] See: www.barceloviajes.com
[2] See: http://www.tti.org
[3] See: http://www.opentravel.org

information sources which allow a more trusted and easy interaction between tourist and service providers. ConTur acts as the unifying channel of contents coming from different sources (web pages, social media, public catalogues…) about tourism products and destinations. The technological solution proposed in this paper allows to filter, catalogue, aggregate, merge and integrate contents from heterogeneous sources in a intelligent, reliable and trusted way. Furthermore, two use case scenarios, "Portal Destino 3.0" and "Mi guia multidestino", demonstrates how different stakeholders, like public providers and private travel agencies, can merge and provide up-to-date information to their customers.

References

Barta, R., Feilmayr, C., Pröll, B., Grün, C. & Werthner, H. (2009). Covering the Semantic Space of Tourism - An Approach based on Modularized Ontologies. In *CIAO' 09: Proceedings of the 1st Workshop on Context, Information and Ontologies*.

Berners-Lee, T. H. (2001). The Semantic Web. *Scientific American*, May 2001.

Bilbao, S. (2010). *E4.1 y E4.2 Modelo semántico de restricciones y requisitos de configuración*. Retrieved from http://contur.labs.andago.com

Freeling (2010), *FreeLing User Manual 2.2. Talp Research Center in Universitat Politècnica de Catalunya*. Retrieved from http://nlp.lsi.upc.edu/freeling/doc/userman/userman.pdf

Mittal, S. & Frayman, F. (1989). Towards a Generic Model of Configuration Tasks. In *Proceedings of the Ninth International Joint Conference on Artificial Intelligence* (pp. 1395–1401). Menlo Park, Calif.

Murua, I. L. (2005). La web semántica aplicada al empaquetamiento dinámico de servicios y productos turísticos. *Conferência IADIS Ibero-Americana WWW/Internet 2005.*

Piñero, R. (2007). *TechTear. Los mashups uno de los pilares de la web 2.0*. Retrieved from http://www.techtear.com/2007/03/26/los-mashups-uno-de-los-pilares-de-la-web-20

Prantner, K., Ding, Y., Luger, M., Yan, Z. & Herzog, C. (2007). Tourism Ontology and Semantic Management System: Stateof-the-arts Analysis. *IADIS International ConferenceWWW/Internet 2007 (WWW/Internet2007)*. Vila Real, Portugal.

SAML (2005). *Security Assertion Markup Language. OASIS Standard.* Retrieved from http://saml.xml.org/saml-specifications

W3C: World-Wide Web Consortium. Semantic Web Brief Guide. Retrieved from http://www.w3c.es/Divulgacion/GuiasBreves/WebSemantica

X.509 (2005). The directory: Public-key and attribute certificate frameworks. *ITU-T Recommendation X.509:2005 | ISO/IEC 9594-8:2005.*

Acknowledgements

This paper is supported by Avanza research Program from the Spanish Ministry of Industry, who provides funds to the project "CONTUR: *Plataforma para la gestión inteligente de contenidos en el ambiente de turismo*". We would also like to acknowledge the archeological excavation research team at Atapuerca and to thank the research Project nº CGL 2009-12703-C03-03 under the General Directorate.

Measuring Emotional States in Touristic Events: The Case of the *Emocionómetro*

Aurkene Alzua-Sorzabal,
Ion Gil Fuentetaja, and
Maria Peralta Ariza

CICtourGUNE, Cooperative Research Centre in Tourism, Spain
{aurkenealzua@tourgune.org; iongil@tourgune.org; mariaperalta}@tourgune.org

Abstract

The measurement of emotions has been stated as an integral part of the satisfaction identification process and, thus, a helpful tool for better future strategies on the design of touristic products and services. This research presents an adaptation of the PAD dimensional model of emotional measurement to unique events. Deriving from the literature, the research aims at better understanding the holistic system of the emotional measurement methods, as well as developing a practical application for smartphones. The validation of the tool has been performed by means of a pilot test within the framework of the last phase of the election of Donostia - San Sebastián as the winning bid to become the European Capital of Culture during the year 2016, by means of measurement of the emotional states of the delegations in Donostia - San Sebastián, as well as in Madrid. After the recollection of the emotional states, the time and the location of the measurements, the analysis has identified some differences on the patterns during the days of the pilot test.

Keywords: Hedonic experiences; emotions; measurement; PAD theory; emotional fluctuation; smartphone application.

1 Introduction

Tourism is based on hedonic consumption and, as such, experiences are an integral part of the phenomena (Vogt & Fesenmaier, 1998, p. 558; Govers & Go, 2009, p. 17). Experiences are multi-sensorial phenomena that interconnect senses, cognitive processes and affective processes (Govers & Go, 2009, p.17). These experiences, beyond the ephemeral events, will be the basis of future memories. According to literature, the overall satisfaction of the experience seems to influence future choices (Fredrickson, 2000). Therefore, the understanding of the experiences of individuals when consuming touristic products and services is essential for further planning processes. Consumers' satisfaction is greatly influenced by the emotional situation of the individual (Bigné & Andreu, 2004b). Consequently, being able to identify and measure the emotional state of consumers, tourism agents and operators would be capable of better adjusting their services to the individual needs and expectations of consumers.

The intensity of the emotions experienced by each user will be determinant to assess the overall emotional situation of each individual and group them into different segments (Bigné & Andreu, 2004a). According to the Pleasure-Arousal-Dominance (PAD) theory, emotional states can be measured by means of the mentioned three items. Adopting the dimensional model of measurement, the scale of emotions will identify the intensity and degree of each variable in concrete places and time

moments. In this generic framework, the present paper works with an adapted version of the general PAD model to tourism activities and services and develops a precise application, tested on a pilot test, to be implemented in the case of application based on the development of a novel device to measure the emotional state of the visitors and travellers.

2 Theoretical Background

Emotions are an essential part of every individual, but their definition has been widely discussed and approached from very diverse disciplines (Kleinginna & Kleinginna, 1981). Even the very roots of emotions have been approached from different perspectives. Emotions are psycho-physiological states of mind that represent ways of adaptation to some external environmental or to internal stimuli (Myers, 2004, p. 500). From the point of view of psychology, emotions alter attention, reinforce individuals' behaviour and activate associative networks, which are relevant to memory. From a more physiological point of view, in order to establish an optimal internal mean for the most effective behaviour, emotions organize responses to the different biological systems, including facial expressions, muscles, voice, activity of the autonomous nervous system, as well as to the endocrine system. Therefore, emotions are answers that, depending on the individual mood, vary on the intensity and relation to the stimuli. Even though all these responses derive from the stimuli, according to Kleinginna and Kleinginna (1981), there are different approaches to define emotions, grouped in diverse categories, depending on the main emphasis.

Among these main categories, the most important ones from the point of view of the present research are a) the affective definitions, which put the emphasis on hedonic values and on the arousal level of feelings; b) the cognitive definitions, which emphasize the perceptual and thinking aspects of the emotions, linked to cognitive processes, and c) external-stimuli definitions, which highlight the importance of the external stimuli on the manifestation of an emotion, conversely to the internal stimuli of motivations. Arising from all the different approaches, Kleinginna and Kleinginna (1981) elaborated an operative definition of emotions, which is currently widely used by researchers from various fields of specialization. They stated that "emotion is a complex set of interactions among subjective and objective factors, mediated by neural/hormonal systems, which can (a) give rise to affective experiences such as feelings of arousal, pleasure/displeasure; (b) generate cognitive processes such as emotionally relevant perceptual effects, appraisals, labelling processes; (c) activate widespread physiological adjustments to the arousing conditions; and (d) lead to behavior that is often, but not always, expressive, goal-directed, and adaptive" (Kleinginna & Kleinginna, 1981, p. 355).

At the light of the different approaches to the understanding and consequent analysis of emotions and emotional states, the present research will treat them from a more cognitive-subjective-experiential perspective. This specific approach will foster a better comprehension of what users feel when experiencing a unique event. Therefore, the application of this approach of measurement along a wide time span will identify fluctuations on the general emotional state of every individual and not in a single moment. The identification of these patterns provides an accurate picture of the emotional state of users regarding the activity, lifetime moment or specific event.

2.1 Theoretical Model of PAD

Due to the multi-component nature of emotions, as affirmed by Derbaix and Pham (1989), there are three main methods to measure them: behaviour observation, psycho-physiological measurements and subjective reports. The psycho-physiological measurements, as well as the observational methods of behaviour are non-verbal methods that require specialized tools, because they aim at measuring the activity of the nervous system. This fact can entail a handicap for the measurement itself, because devices and tools are expensive and heavy and can prompt discomfort among individuals.

Verbal methods, on which the subsequently explained PAD model is based, conversely, utilize questionnaires with emotional scales to measure the state of individuals. This type of methods present some advantages regarding the non-verbal ones, mainly the greater easiness to reach the target sample. Authors such as Bagozzi, Gopinath and Nyer (1999) suggest that verbal methods based on emotional scales are the most suitable ones to perform more accurate measurements, because of a) the adaptation capacity to the diverse reactions under analysis; b) the anonymity and minimization of reactive effects due to the physical absence of the observer; c) the non-requirement of special equipment; d) the easiness to reach wider samples; and e) the possibility of measuring the direction (positive or negative), as well as the content of the emotions. Anyway, as stated by Derbaix and Pham (1989), there are some inconveniences related to the verbal methods. Among them, the most widespread ones are difficulties for the translation of the questionnaires, difficulties to obtain additional information due to the subjective character of the sample and retrospection problems.

The measurement used on these verbal methods can be classified as discrete or as dimensional. The discrete approach to the measurement presents dichotomous pairs of feelings and forces individuals to choose one of the items of each pair in order to identify their emotional state. The dimensional measurement of emotions, on the other hand, allows users to have greater flexibility when defining their emotional state. Emotions are not static feelings and, as have been mentioned before, the emotional state varies on intensity. Thus, compartmentalized emotions will not distinguish between the different degrees and intensities on individuals.

A theoretical approach to this measurement process is represented on the Pleasure-Arousal-Dominance one, widely known by its acronym PAD. This model, as defined by Mehrabian and Russel (1974), Russell (1980) and Mehrabian (1995), assumes that emotional spaces consist of a reduced number of abstract dimensions, on which specific emotional states are based. The fundamental premise of this dimensional approach is that emotions vary on a continuum in some limited affective dimensions. That is, intensity and direction of emotions fluctuate along the range of some specific affective dimensions. These specific dimensions are defined by this approach as pleasure (P), arousal (A) and dominance (D). This theoretical approach to the measurement of emotions, deriving from the environmental psychology, describes the emotional situation of every individual according to the fluctuation and combination of the recently mentioned dimensions. Even though they respond to very abstract items, these dimensions can be described as:

- Pleasure regarding to an either positive or negative affective state of the user or, in other words, to a subjective feeling of the delight perceived by the individual.
- Arousal concerning the activation degree of the individual, that is, a state of feeling that varies along a unique dimension that ranges from asleep to a frenetic activity.
- Dominance relating to the degree of freedom or its absence that the individual feels has to act.

Table 1. Definition of pairs of the emotional dimensions of Mehrabian and Russell (1974)

Pleasure	**Arousal**	**Dominance**
Annoyed - Pleased	Sluggish-Frenzied	Submissive-Dominant
Unsatisfied-Satisfied	Unaroused-Aroused	Influenced-Influential
Unhappy-Happy	Sleepy- Wide Awake	Controlled-Controlling
Bored-Relaxed	Calm-Excited	Guided-Autonomous
Melancholic-Contented	Relaxed-Stimulated	Cared For-In Control
Despairing-Hopeful	Dull-Jittery	Awed-Important

Therefore, these three dimensions are ranked in scales that define and delimit their higher or lower intensity. The combination of the displayed values in each dimension will constitute, according to the PAD theory, the basis of the identification of the global emotional situation of the analysed individual.

For the present research this dimensional PAD model has been chosen as the optimal one to measure emotional states of individuals in the pilot test represented by a unique and unrepeatable event, such as the nomination of the candidature of Donostia - San Sebastián as the one awarded with the European Capital of Culture for the year 2016[1].

2.2 Application of the PAD Model in the Tourism Realm

This theoretical model has been lately applied to specific fields of research. It has been traditionally used for marketing purposes, due to the great importance of emotions of consumers to analyse their behaviour (Laros & Steenkamp, 2005; Bagozzi, Gopinath, & Nyer, 1999; Smith & Bolton, 2002; Hirschman & Stern, 1999). The understanding of the emotions and consequent behaviour of consumers is at the core of any successful marketing process. In this sense, the analysis of the emotional situation and fluctuations of consumers is important both for a deeper understanding of the consumers (Hirschman & Stern, 1999), as well as for de development of services, marketing strategies and failure prevention measures planning (Smith & Bolton, 2002). Therefore, the identification of the emotional state of consumers is a helpful tool to better manage the personal, but also social and cultural impacts of the design products and services. Thus, the measurement of emotions and emotional states is related to the identification of the satisfaction degree of consumers.

Tourism industries should be aware of the potential of measurement of emotions on all their services, because tourism is, in fact, a hedonic consumption (Go & Govers,

[1] The specifications of this event will be explained more in detail later.

2009, p. 17). The identification of the emotions throughout the whole consumption process of touristic services allows delimiting their fluctuation. The analysis of these general emotional states and fluctuations from a global and an individual perspective supports the better knowledge of the satisfaction of tourists related to specific events and services. In other words, the measurement of emotional states of tourists' favours, on the one hand, the identification of the success or failure of any service in relation to emotions of the consumers and, on the other, the delimitation of potential groups of consumers based on their emotional states and fluctuations. This twofold approach to the measurement of emotional states of consumers will allow touristic agents and operators to better meet the future needs and expectations of individuals by means of a segmentation of both tourists and services.

As the continuous measurement of the emotional states of consumers is practically non possible, the measurement during key significant different events or consumption of services can beforehand foster the further improvement of design and management processes. If tourism industries' stakeholders are aware of the emotional strengths and potential weaknesses of their services and products, they could be able to better perform in an increasingly hedonic competitive market.

3 Methodology

At the light of these theoretical approaches, the Cooperative Research Center in Tourism - CICtourGUNE has developed an application to measure the emotions of visitors. The main objectives of the present research are: (1) to adapt the dimensional model of PAD to the practical case of the final stage of the selection of the European Capital of Culture 2016; (2) to validate the application developed by CICtourGUNE as means of an advanced emotional measurement tool.

3.1 Empiric Pilot Test: the Case of DSS2016

The city of Donostia - San Sebastián (Basque Country, Spain) aimed at becoming the European Capital of Culture during the year 2016. The bid prepared by the responsible organisation was to be defended and evaluated by the commission of experts selected by the European Commission and the Ministry of Culture of the Spanish Government between the 27th and the 28th of June of 2011. This event represented an optimal pilot test to test the measurement system developed by CICtourGUNE.

In this context, the developed system of measurement was applied during the days 26, 27 and 28 of June of 2011, coinciding with the final stage of the election of the winning bid to become the European Capital of Culture on the 2016. As understood, this event is unique due to its singularity, but it also presents an unrepeatable character being appropriate to be selected as a practical case for the validation of the model and application developed by CICtourGUNE. The emotional measurement was done in intervals of two hours during the mentioned days within the time frame starting at 11:00 in the morning and finishing at 21:00 at night.

Participants on the pilot test of measuring the emotional situation with the developed tool were the staff components and representatives of the candidature of Donostia - San Sebastián 2016 (DSS2016). These participants represent a small sample, but,

since it is a pilot test, it has been assessed as optimal for the validation of the measuring system and the application. The sample consists of the staff of the candidature, public representatives, collaborators of the bid, as well as voluntary citizens. Even though the further extrapolation of result should be taken with caution, this pilot test has validated the system and application suggested.

The evaluation and selection of the winning bid was made in different places of Madrid on the 26th, 27th and 28th of June, as mentioned before. The 26th of June a rehearsal of the defence was performed at the Museo del Traje of Madrid, the place where the following day the defence itself took place. Finally, the 28th of June, the election of the winning bid took place in the Ministry of Culture of the Spanish Government. Moreover, part of the team of the candidature did not travel to Madrid, but stayed in Donostia - San Sebastián. For this reason, as will be explained on the section about the developed tool, the device allowed to store the location of each participant, in order to better understand obtained data.

Before starting the measurement of the emotions, participants were given a brief questionnaire to define their socio-demographical profile and two specific questions identifying their attitude towards the candidature, as well as the expectations to be the winning bid. Regarding the socio-demographical profile of the sample, six out of ten participants were females and more than half of the sample was between 26 and 40 years old. It is worth to mention that, whilst 94.0 per cent of the participants affirmed to have a totally positive attitude towards the candidature, only half of the total sample stated to be absolutely confident on being the winning one. In fact, four out of ten participants had some confidence and an additional six per cent a low degree of confidence on the success of the bid.

3.2 Adaptation of the PAD Model to the Measurement of DSS2016

As mentioned among the main objectives of this research, an adaptation of the general PAD model has been carried out for the specific measurement of the emotional fluctuations of the components of the candidature of DSS2016. To simplify the measurement by users, the diverse pairs of items used on other applications have been reduced to the three general emotions. Anyhow, for a better understanding of what each emotion describes, a table with the embraced different pairs of items was delivered to each participant.

Having these definitions in mind, participants were asked to define the intensity of each emotion on a scale from 0 to 10, the lowest values addressing the lowest presence of the emotion and the highest ones the greatest presence.

Taking into account that the analysed case presents three very specific events such as the rehearsal, the defence and the final decision of the committee, the research aimed at identifying fluctuations on the emotional states of participants regarding their time frame in relation to these milestones. Thus, the research identifies whether the measurement has been performed before or after the rehearsal, the defence or the final election for each specific day.

This distinction will allow relating potential fluctuations on the emotional states with the period of measurement. To better understand the emotional flow, the analysis has

grouped all the measurement into six specific milestones, namely before and after the rehearsal, before and after the defence, as well as before and after the election[1].

Similarly, as the research assumes that the spatial location of participants can affect their emotional state, apart from a general measurement, another additional one has been carried out taking into account if the participants where in-situ in Madrid, in Donostia - San Sebastián or somewhere else. This differentiation supports further analyses related to geographical location of the participants.

Finally, one of the difficulties that the research had to face was, as the literature highlights, the translation of the emotions and the pairs to the three languages used during the pilot test. Due to the co-official nature of both Basque and Spanish in Donostia - San Sebastián, these two languages were the core ones when developing the application. Additionally, because of the international character of the team of the candidature, English was also implemented. The difficulty of finding the exact translation for each emotion was one of the handicaps this research has faced.

3.3 Technical Specifications of the System: the *Emocionómetro*

The developed system, named *Emocionómetro* or Emotionometer in English, is divided into two distinct modules. On the one hand, (1) an Android client application developed for the collection of the emotional state and, (2) on the other, a data storage module.

Module 1: Android client Application

First of all, due to the growing importance of mobile devices and more specifically smartphones among individual users (Want, 2009), the system was developed to be used on these devices. The client application developed for the collection of emotional state is developed under the Android operating system. The choice of this operative system is based on the forecast of Android being the system of around half of the market of worldwide smartphones by the end of 2012[2]. This way, users can even utilize their own devices to perform the measurement after a simple installation of the application.

The application consists of an alarm manager developed for the regular collection of the emotional state. This alarm manager specifies the frequency of measurement of the emotional state, as well as the time frame of activation of the system. When the application detects the time set for the measurement by the alarm manager, launches the emotional collection interface. In this specific case, it was developed for tactile screen, in order to facilitate the identification of the degree of each emotion by the participants. This interface collects on the 0-10 scale the level of pleasure, arousal and dominance of the situation.

[1] These six different milestones are reflected on the figures as AR = ante rehearsal; PR = post-rehearsal; AD = ante defence; PD = post-defence; AE = ante election; and PE = post-election.
[2] Gartner Newsroom, April 2011. See: http://www.gartner.com/it/page.jsp?id=1622614. Last accessed 01/09/2011.

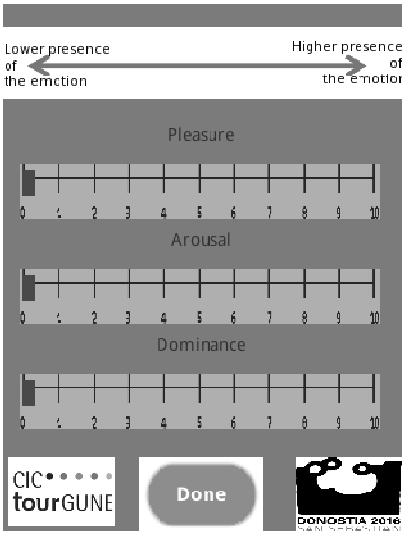


Fig. 1. Snapshot of the interface of the device

In addition to storing the emotional state, it also stores the position and exact time of the measurement. After the recollection of the emotional state, the system returns to its previous state. Once the whole process of capturing the emotional state has been completed, the system sends all the information recorded to data servers of CICtourGUNE in XML format.

Module 2: Data Storage

The data storage module consists of a Servlet and a database. The Servlet is responsible of the reception of the data in XML format and then parse them to introduce the information in the database. This information corresponds to the emotional measurements made by the individual throughout the measurement cycle.

The database used is PostgreSQL with a PostGIS module that will give data a GIS character. With this feature, along with the delimitation of the target areas of Madrid and Donostia - San Sebastián made by CICtourGUNE, the analysis has also supported a segmentation of the measurement of the emotional states by location.

4 Overview on the Results of the Analysis

After the completion of all the emotional measurements performed during the specified three days, all data were translated into statistically usable information for further treatment with IBM SPSS Statistics 19.0 software.

Related to the overall emotional means, the average values pointed above six in the ten points scale. Anyway, there are some slight differences among them, because the mean value for pleasure is 6.67 and for arousal 6.90, whilst the value of dominance is placed on 6.21. Additionally, looking at all the collected values, it can be deduced that most of the measurements related to pleasure range from six to eight.

This same pattern is identified when dealing with the arousal measurements, but the ones about dominance show lower values. In fact, taking them as a whole, most of the measurements regarding dominance range are places in a span between five and six of our scale.

Differences among the three emotional dimensions identified by this research have not been excessively high during the three days of the measurement. In general terms, the rehearsal day, as well as the defence one present more homogeneous emotional patterns than the day of the election of the winning bid. Similarly, the general levels of pleasure and arousal show a greater difference between highest and lowest values than the dominance ones. In fact, levels of intensity of dominance fluctuate in a range of more uniform values.

Analysing each of the emotions individually, it can be affirmed, as seen on figure 2, that pleasure and arousal show a generally increasing pattern throughout the measurement process, while dominance displays a different general pattern, with more marked peaks and lows of intensity.

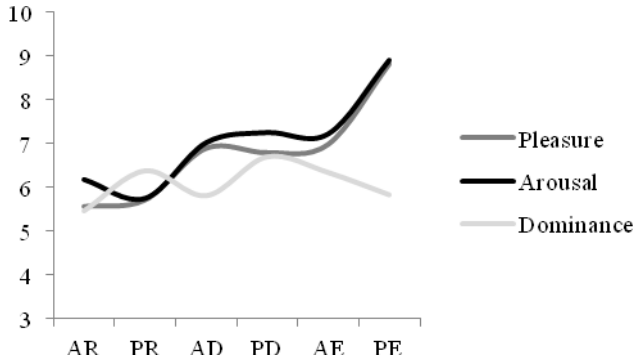


Fig. 2. Graphic of the general emotional flows

Thanks to the storing of the location of each emotional register, the analysis has been able to differentiate a more detailed emotional pattern depending on the specific location. From this locally individualized analysis, it can be concluded that there are some stressed differences between the general patterns of the participants in Donostia - San Sebastián and the ones in Madrid. Levels of pleasure and arousal in Donostia - San Sebastián present a quite homogeneous pattern, until the moment the candidature DSS2016 was awarded as the winning one. In this moment, the intensity of these two mentioned emotions increases significantly from between six and seven to above nine in both cases. Degrees of intensity of dominance in Donostia - San Sebastián increase gradually from the beginning, but still remain below seven in the final moments, showing the most homogeneous pattern of the three emotions. Regarding the emotional patterns of the participants in Madrid, it shows a greater variance on the emotional intensity. The pleasure indicator increases after the rehearsal of the 26th of June and maintains a constant gradual increase.

The dominance of participants in Madrid, conversely, displays great irregularities, defined by two sharp decreases on the levels measured before the defence in front of the committee and after the verdict granting them the European Capital of Culture for the year 2016.

Consequently, from the point of view of the analysis by location of the participants, according to the results displayed on figure 3, it can be concluded that emotional patterns among participants in Donostia - San Sebastián have been significantly homogeneous, with similar values during the three days, except for pleasure and arousal after the election of the committee. Components of the delegation in Madrid, on the contrary, display a greater variety on the intensity of the emotions and a pattern with abrupt peaks and lows, especially when dealing with the dominance. In fact, from being the emotion with the highest value at the beginning with a measurement over eight in our scale, it becomes the one with the lowest value after the election, since it scores below four.

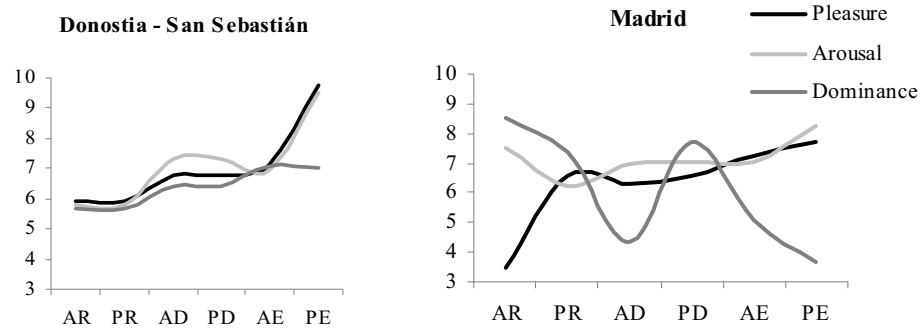


Fig. 3. Graphic of the emotional flows in Donostia - San Sebastián and Madrid

Summing up, the developed tool has measured the emotional states of participants of the candidature of Donostia – San Sebastián during the final phase of the selection process to become the European Capital of Culture on 2016. According to the obtained results, there are evidences pointing out that the period and the location of measurement foster different intensities on the utilized three emotional dimensions in this case. This research presupposes that the greater variability on the emotional intensity of the delegation in Madrid can be linked to the closer experience of all the actions, reaching a very significant expression of their emotional state after the election of the candidature DSS2016 as the winning one. At this point, pleasure and arousal reach their highest values among participants in Madrid, while the dominance displays its lowest. From the point of view of this research, this state could be labelled as a hardly controllable joyful excitement, answering to the achievement of a goal insistently pursued.

5 Conclusions and Further Implications

The practical application of the *Emocionómetro* with the team responsible of the DSS2016 candidature has proven to be of high interest. On the one hand, it has permitted to adapt the dimensional measurement of PAD to a specific event and, consequently, has opened the opportunity to apply this adaptation to future touristic events. On the other hand, the device developed by CICtourGUNE has validated the evaluation system.

Regarding the adaptation of the PAD dimensional model for the measurement of emotions, no special problems have been identified when implementing it. Additionally, having been a well-established measurement and satisfaction identification system in the field of marketing, results obtained from the application of the model to unique and unrepeatable events can help touristic activities managers in further design and development processes. Since the emotional state is an integral aspect of the overall satisfaction and, thus, might influence future choices, event, activities and product managers can improve their strategic planning processes by a better understanding of the real emotional state of the consumer.

Thanks to the adaptation of the PAD model and to previous researches on the measurement of emotions in touristic and leisure related environments, a more holistic overview of the emotional measurements has been obtained.

Concerning the application of the developed system, it should be stated that, once overcome the initial translation problems of the emotions, the pilot test has been successful. Although verbal measurements require a really high degree of participation and involvement of the participants, no special difficulties have been faced to reach a representative sample and obtained necessary data.

Furthermore, the developed tool for the emotional measurement has presented a great acceptance among the participants of the pilot test. The system stored the emotional state, as well as the temporal and geographical coordinates of users with no remarkable difficulty. The use of smartphones has been key due to its advanced features. Similarly, the simple, intuitive and user-friendly interface has been assessed as successful. No negative reports were gathered concerning the usability of the interface. Therefore, there are evidences to state that the potentiality of developing further similar devices to measure the emotional states of consumers is high.

The present research and application enhance our understanding of tourism as a hedonic experience. Along with it, the new comprehension will support tourism operators to better match the needs and expectations of visitors and travellers. Evidences show that there is room to develop new approaches, methods and tools to get innovative insights of the tourism phenomenon, so that new steps can be taken to improve it as a scientific discipline.

References

Bagozzi, R. P., Gopinath, M. & Nyer, P. U. (1999). The Role of Emotions in Marketing. *Journal of the Academy of Marketing Science* 27(2): 184-206.

Bigné, J. E. & Andreu, L. (2004a). Emociones del consumidor. Aplicación de la escala agrado-activación en la segmentación de visitantes de servicios de ocio y turismo. *Revista Española de Investigación de Marketing ESIC* (pp. 11-36).

Bigné, J. E. & Andreu, L. (2004b). Modelo cognitivo-afectivo de la satisfacción en servicios de ocio y turismo. *Cuadernos de Economía y Dirección de la Empresa* 21: 89-120.

Derbaix, C. & Pham, M. T. (1989). Pour un développement des mesures de l´affectif en marketing: systhèse des prérequis. *Recherche et Applications en Marketing* IV(4): 71-87.

Fredrickson, B. L. (2000). Extracting meaning from past affective experiences: The importance of peaks, ends, and specific emotions. *Psychology* 14(4): 577-606.

Gartner Newsroom (2011). *Gartner Says Android to Command Nearly Half of Worldwide Smartphone Operating System Market by Year-End 2012, April*. Retrieved September 1, 2011, from http://www.gartner.com/it/page.jsp?id=1622614.

Govers, R. & Go, V. (2009). *Place branding: Glocal, Virtual and Physical Identities, Constructed, Imagined and Experienced*. Basingstoke-New York: Palgrave MacMillan.

Hirschman, E. C. & Stern, B. B. (1999). The Roles of Emotion in Consumer Research. *Advances in Consumer Research* 26: 4-11.

Kleinginna, P. R. & Kleinginna, A. M. (1981). A categorized list of emotion definitions, with suggestions for a consensual definition. *Motivation and Emotion* 5(4): 345-379.

Laros, F. & Steenkamp, J. (2005). Emotions in consumer behavior: a hierarchical approach. *Journal of Business Research* 58(10): 1437-1445.

Meharabian, A. (1995). Framework for a comprehensive description and measurement of emotional states. *Genetic, Social & General Psychology Monographs* 121(3): 339-361.

Meharabian, A. and Russell, J. A. (1974). *An approach to environmental psychology*. Cambridge, Massachusetts: MIT Press.

Myers, D. G. (2010). *Psychology* (9th ed.). New York: Worth Publishers.

Russell, J. A. (1980). A Circumplex Model of Affect. *Journal of Personality and Social Psychology* 39(6): 1161-1178.

Smith, A. K. & Bolton, R. N. (2002). The Effect of Customers' Emotional Responses to Service Failures on Their Recovery Effort Evaluations and Satisfaction Judgments. *Journal of the Academy of Marketing Science* 30(1): 5-23.

Vogt, C. A. & Fesenmaier, D. R. (1998). Expanding the functional information search model. *Annals of Tourism Research* 25(3): 551-578.

Want, R. (2009). When Cell Phones Become Computers. *IEEE Pervasive Computing*, April-June: 2-5.

Acknowledgements

The research team of CICtourGUNE would like to acknowledge the kind collaboration of the candidature DSS2016 and of all the participants from the delegations in Donostia - San Sebastián and Madrid for the success of the pilot experience.

Analysing the Emotional Appeal of a Website through Verbal and Pictorial Stimuli

Anke Schneider, and
Barbara Stangl

Institute for Tourism and Leisure Studies
Vienna University of Economics and Business, Austria
{anke.schneider, barbara.stangl}@wu.ac.at

Abstract

When looking for information about a tourism product in the majority of cases websites are the first contact most tourists have with a tourism product. Moreover, the first impression of a destination or tourist facility is crucial for the evaluation of these. As pictures and text are a main part of a tourism website, a tourist's opinion will be influenced by them. So, through the visual and textual stimuli an experience and especially emotions are evoked. Therefore, it is of major importance to know what a picture and a text initiates and why. Hence, this research study analyzes pictures and text on tourism websites related to their emotions in order to catch the visual appeal and in succession the induced image. Results show that verbal stimuli are used skilfully while a third of the pictorial stimuli are perceived negative. Therefore improvements are necessary to create a positive overall perception.

Keywords: Emotion, Picture, Text, Image

1 Introduction

Tourism as a part of the service industry is characterized by intangibility and heterogeneity (Sheldon, 1997). Therefore, it is very important for tourists to look for detailed information, at first internally and in a second step externally (Engel, Blackwell & Miniard, 1995), here especially on the internet (Zhang, Pan, Smith & Li, 2009; Xiang & Pan, 2009; Pan & Fesenmaier, 2006). So mostly, websites are the first contact tourists have with the tourism product. Additionally, the first impression of a destination or tourist facility is crucial for the evaluation of these (Haahti & Komppula, 2006). This means if users find a website visually appealing, other factors (like usability, missing information) play a secondary role (Lindgaard, Fernandes, Dudek & Brown, 2006).

When thinking of the form of a website, the major components are the structure, the pictures and the text. To improve the visual appeal, these three segments have to harmonize with each other. A lot of research was done in usablitiy studies concerning the structure (Moss, Gunn, & Heller, 2006; Ruso & Mayer, 2006), but the minor elements of the website, namely pictures and text, which build the structure, are not really analysed. As pictures and text are a main part of a tourism website a tourist's opinion will be influenced by them.

Gunn (1988) called these views in the mind of the tourists formed through websites, brochures, etc. induced image. Thus, an experience and especially emotions are evoked. Therefore, it is of major importance to know what a picture and a text initiates. Although research studies in measuring emotions evoked through text

(Strapparava & Mihalcea, 2008) and pictures exist (Lang, Bradley, & Cuthbert, 2008; Schweiger & Wiklicky, 1986), most marketing campaigns only think about the ideal theme and not really about the evoked emotion. Hence, this research study analyzes pictures and text on tourism websites related to their emotions in order to catch the visual appeal and in succession the induced image. The following research is devoted to which emotions are evoked through pictures and text on selected websites and what could be the reason that these emotions arise.

2 Theoretical Background

2.1 Induced Visual and Textual Image

The experiences and impressions which a potential tourist has when visiting a website, reading a book or looking at a brochure, form the so called induced image (Gunn, 1988). Through the use of pictures on a website, a certain image is created called the visual image. This visual image is defined as "the interest that a picture will create if potential tourists observe it and it inspires for a first and repeated visits" (Ye & Tussyadiah, 2011). Based on this literature background, the use of text creates also a special image which is from now on called the textual image.

So far there are various Image-Analyses (Baloglu & McCleary, 1999; Pike, 2002; Mazanec, 2010), however, none of these are analysing the emotional component. Most of them, focus on the cognitive dimension of an image only (Bigné Alcaniz, Sánchez García, & Sanz Blas, 2009). Moreover, there are few studies which compare the projected image meaning the image transported by the tourism facility and the perceived image meaning the image of the visitor (Grosspietsch, 2006).

Due to the inspection of pictures and text from websites, in this paper the visual and textual image and in succession the induced image caused through selected websites regarding their emotional component is detected.

2.2 Evoked Emotions

It could be said that a tourism website has to fulfill two big challenges. On the one hand it is important to build trust in order to create a relationship between the offered product and the prospective tourist as the user of the website. On the other hand the website has to arise first experiences and inspirations.

Experiences and also trust are mostly created through emotions (Dunn & Schweitzer, 2005). This means that through emotional stimuli a person will be activated and will be strongly connected to the tourism product. Emotions are the key to attract potential tourists.

The imparted information will be easier memorized and stays longer in the memory (Foscht & Swoboda, 2007) if the appropriate emotional stimuli are used. Bagozzi, Gopinath and Nyer (1999) call these kinds of advertisement "feeling ads". Especially, they can be faster addressed as rational information (Lindgaard, Fernandes, Dudek, & Brown, 2006).

An exact definition of emotion is hard to generate as scientists from different disciplines have different accesses to the term (Scherer, 2005; Kleinginna &

Kleinginna, 1981). However, there are some characteristics which can be found in all definitions. Based on these, emotions are current physical conditions of humans which are from a specific quality, intensity and duration. Normally, they are related to an object and lead to physiological and behavioral modifications of the person (Meyer, Reisenzein, & Schützwohl, 1993).

So, the challenge in emotion research is at first to identify emotions and furthermore to analyze the specificity of an emotion in order to describe it and to relate a certain emotion to all the others (Ulich & Mayring, 1992). There are three different strategies to structure emotions: basic emotions, dimension analysis and classification systems. Hence, basic emotions are tried to be identified to bridge the gap between huge numbers of different emotion theories and to generate one common basis (Plutchik, 1962; Izard, 1977; Russell & Mehrabian, 1977).

By two or more basic emotions emerge secondary emotions. Basic emotions can be mapped in a two-dimensional space defined by two orthogonal bipolar dimensions. However, dimension analysis is strongly criticized due to their methodological design (Scherer, 2003; Roseman, 2001). All strategies share a common problem to the number of used emotions, which is traced back to the emotion definition problem.

Scherer combines in his research two of these approaches to create a new measuring tool. He uses basic emotions as categories to classify words. His list of emotions includes four emotion families (high control, low control, positive and negative) and 16 emotions (pride, elation, happiness, satisfaction, relief, hope, interest, surprise, anxiety, sadness, boredom, shame/guilt, disgust, contempt, hostility and anger) in total (Scherer, 2005).

3 Method

In a first step an adequate website has to be found from which pictures and text could be taken for further analysis. Then the selected pictures and text were measured concerning the emotions which evoke.

3.1 Website Selection

Starting point of this research was the brochure “Night of Museums 2010” including 59 Viennese museums. All these museums have their own official website which was the base for further analysis. In order to find the most interesting museums which should be part of deeper research, a scoring model was implemented. This model consisted of six main categories and 19 sub-categories (see Table 1) which were evaluated on a score from 0 (not available) to 3 (outstanding). One or two points were given for available but not above average. The categories were mostly chosen based on literature (Weideman & Mgidana, 2004; Choros & Muskala, 2009; Hong & Kim, 2004) and weighted to show their importance.

One category called the *outline* was weighted 25%. Within this category two parts were of main interest. First, the interactivity of the website which was analyzed based on user-generated content like guest books, blogs or RSS-Feeds, and second, the choice set of languages.

Another category with a weighting of 25% concentrated on the *usability* of the website. The logical command structure and the number of points in the navigation were crucial in this category as well as the design of the surface of the website.

Table 1. Scoring Model

Main category	Sub-category	Weight sub-category	Weight main-category
Outline	Languages	40%	25%
	Interactivity	60%	
Usability	Construction	40%	25%
	Hierarchy	30%	
	Surface engineering	30%	
Information	Price	20%	15%
	Opening hours	20%	
	Events	15%	
	News	15%	
	Map	15%	
	Contact information	15%	
Design	Fits identity	40%	15%
	Foreign ads	20%	
	Innovative Elements	20%	
Currentness	Current news	70%	15%
	Last update	30%	
Search tool	Looking for price	40%	5%
	Looking for opening hours	40%	
	Looking for driving directions	20%	

Also the *information* presented on the website was analyzed and weighted with 15%. The information of interest were prices, opening hours, events, news, map, contact information. Moreover, the availability of interactive tools as Google maps had an impact on the score.

The criteria *design* with 15% weight was analyzed in terms of fit with the topic and the availability of distracting foreign advertising.

Currentness of data was also of main interest and therefore weighted 15%. Not only the currentness of the website was analyzed, but also if current news are presented on the website.

When looking for specific information on the website a *search tool* would be helpful. Therefore also the availability of this tool and the correctness of the results were part of the scoring model. It was weighted with 5%.

Based on the results of the scoring model two museums were chosen. One is the Albertina which had the highest score. The other one is the Technical Museum Vienna which was among the 10 highest rated and was chosen because of its completely different topic.

3.2 Picture Inspection

For a further picture analysis, 29 pictures were taken from the Albertina website and 22 pictures were taken from the website of the Technical Museum Vienna, more precisely from each site of the main menu all pictures were taken. These pictures included known and unknown photographs and paintings as well as pictures from the museums itself. In order to find a relationship between a specific picture and an emotion, an online survey was conducted. For this purpose, the online tool Websort (http://www.websort.net/) was used for generating a closed card sorting study.

50 respondents had to allocate the pictures to specific words standing for emotions based on Scherer's list (see 2.2 Evoked Emotions). Because not all of the emotional words fit to museum websites, only an applicable selection was chosen: admirable [bewundernswert], interesting [interessant], fascinating [faszinierend], pleasant [angenehm], nostalgic [nostalgisch], surprising [überraschend], happy [fröhlich], cheerful [heiter], relaxing [entspannend], positive [positiv], negative [negativ]. It was possible to allocate to each emotion more than one picture. After all respondents finished the survey, the data was downloaded from the website and cumulated.

3.3 Text Inspection

For the text analyses the webpage evaluating tool Webeval was used. The tool was developed by Wöber[1] and can be found under http://www.tourism.wu-wien.ac.at/webeval.html. This tool enables to list all words in the text of the website with their frequency.

For the text study, the number of words was cut by non meaning full words like articles or pronouns first. Secondly, all words which would be hard to allocate to the emotions on Scherer's list were eliminated. The final 45 words from the Albertina website as well as 30 words from the website of the Technical Museum Vienna were used in the card sorting questionnaire on Websort. The same procedure was arranged as in the picture analysis. 50 respondents had to allocate the words to emotions which were the same emotions as mentioned above.

4 Analysis

The output data from the card sorting studies (frequencies) was used as input data for further cluster analysis. For the cluster analysis the neuronal network program TRN (Topology representing Network) was used.

[1] See: http://www.modul.ac.at/woeber

The advantage over traditional cluster analysis is that it provides more robust cluster (Ganglmair & Wooliscroft, 2001). TRN is a non parametric method that relies on vector quantification (Martinetz & Schulten, 1991).

After the cluster analysis was accomplished a correspondence analysis was added to get better results. Therefore, the diverse emotions were categorized in a further study into positive, negative and neutral. These resulting categorizations from text and pictures serve as input data for following correspondence analysis.

5 Results

5.1 Picture Study

TRN provides the wSSI to find the right cluster solution which is in this study 0.46, and advices a five cluster solution. Moreover, for this solution a robustness of 0.93 is given. By means of the provided prototypes table, a description of the clusters can be made (Table 2 shows a summary).

Cluster 1 named *grounded colours* includes 13.73% of the pictures. Besides a picture of the Albertina itself, the Dürer rabbit, the logo of the technical museum as well as a concrete mixer are included in this cluster. The colours brown or off-white are dominated in this cluster. It is mostly associated with positive and pleasant (Table 2).

The second cluster is concentrated on interaction and therefore also called *interaction*. Most pictures show people doing something or object which are standing for speed like a car. It is associated with interesting and with 13.73% it is together with cluster 1 the second biggest cluster (Table 2).

The pictures in cluster 3 (size 29.41%) are dominated by *high luminance*, this means that mostly light colours are used or the illumination is quiet high. Emotions which come up with this cluster are interesting and negative (Table 2).

Concerning cluster 4 mostly *dark colours* are presented. The cluster includes 27.45% of the pictures and is associated with negative and nostalgic. The second smallest cluster with 15.69% of the pictures is called *few colours* as most pictures are dominated by two to three colours. Colours included are black and white as well as green and orange. It is associated with nostalgic.

Table 2. Picture Clustering Summary

Cluster & Size	Name	Description	Picture Examples	
Cluster 1 13.73%	Grounded colours	Positive (6.258), Pleasant (6.104)		
Cluster 2 13.73%	Interaction	Interesting (12.720),		
Cluster 3 29.41%	High luminance	Interesting (9.874), Negative (9.612)		
Cluster 4 27.45%	Dark colours	Negative (10.033), Nostalgic (10.769)		
Cluster 5 15.69%	Few colours	Nostalgic (16.921)		

5.2 Text Study

In almost the same manner as the picture clustering, also the clustering of the text found the best choice with six clusters. In this case the wSSi comes to 0.54 and this solution shows a robustness of 0.99. Based on the prototypes table, further description of the clusters can be made (Table 3 shows a summary). Cluster 1 includes 5.34% of the words and is therefore the smallest cluster. Words like price, controlled or red are in this cluster which is associated with the emotion negative. The second cluster called *past* includes words which are related to previous times. Examples are the words steam-engine, archive or baroque concerts. It includes 13.33% of the words and is associated with nostalgic.

Table 3. Text Clustering Summary

Cluster & Size	Name	Description	Text Examples
Cluster 1 5.34%	Price	Negative (24.825)	Red [rot] Price [Preise]
Cluster 2 13.33%	Past	Nostalgic (16.987),	Steam-engine presentation [Dampfamaschinenvorführung] Historic [historisch]
Cluster 3 21.33%	Information	Interesting (22.731)	Information [Information] Exhibiton [Ausstellung]
Cluster 4 13.33%	Community	Happy (13.207)	Family [Familie] Welcome [willkommen]
Cluster 5 26.67%		Positive (19.110),	International Concentrated [konzentriert]
Cluster 6 20.00%	Extraordinary	Admirable (10.630), Fascinating (11.141)	Michaelangelo treasure [Schatz]

Cluster 3 is called *information* as most of the words included stand for a description of the museum like exhibition, collection or publications. It is associated with the emotion interesting and includes 21.33% of the words. Cluster 4 containing 13.33% is called *community* as most words are related to a collective. Examples are the words family, children or pre-Christmassy. Cluster 5, the biggest cluster, which is associated with the emotion positive contains a very wide spectrum of words. It starts with words like attractive and efficient and goes to concentrated and online. Cluster 6 including one fifth of the words is called *extraordinary*. In this cluster names of famous artists are very dominant. Moreover, the word experts, research team or treasure can be found in this cluster. Emotions associated are admirable and fascinating.

5.3 Combining Text and Picture

Based on the results of the cluster analysis a classification of the emotions to the clusters was unclear. Therefore, an additional sorting study was made to categorize Scherer's emotions into positive, negative and neutral. All emotions beside of nostalgic and negative were supposed to be positive whereas nostalgic was associated with neutral and negative with negative. These categories and the 11 clusters serve as input data for the following correspondence analysis. A two-dimensional solution of the correspondence analysis describes 100% of the model. The first dimension explains 57% whereas the second describes 43%. More information about the objects (clusters) and attributes (emotional categories) illustrated in Figure 1 is prepared in Table 4. Through the variable mass the importance of clusters and categories is described. In this analysis are all clusters nearly of the same importance (mass= 0.090 to 0.092) besides the category positive which shows a superior importance (mass= 0.742). The quality consisting of the variables cor1 and cor2 declares how detailed objects and attributes are described through the two illustrated dimensions. So, it is

not remarkable that the quality is 1 as the two-dimensional solution describing 100% is used. Figure 1 (below) represents the two-dimensional solution of the correspondence analysis. On the basis of the difference between the mean profile of all data which is illustrated in the coordinate system, and the object or attribute the homogeneity or heterogeneity is shown. Also the inertia is a variable regarding the difference. Especially, the clusters text1 (inertia= 0.439), text2 (inertia= 0.139) and picture5 (inertia= 0.139) as well as the categories positive (inertia= 0.126), negative (inertia = 0.499) and neutral (inertia = 0.356) show a high difference to the mean profile. In contrast, the cluster picture2 (inertia= 0.007) is located next to this point and is therefore not exactly characterized. Through which dimension an object or attribute is explained is pointed out in the variables ctr1 and ctr2. As Table 4 shows the category neutral (ctr2= 0.851) is described via dimension 2 whereas the category negative (ctr1= 0.871) is explained through dimension 1.

Table 4. Results of the Correspondence Analysis

Name	Mass	Inertia	k=1	Cor1	Ctr1	k=2	Cor2	Ctr2
Text1	0.09	0.439	-1.129	0.97	0.747	-0.198	0.03	0.03
Text2	0.091	0.139	0.162	0.063	0.016	0.623	0.937	0.304
Text3	0.091	0.023	0.171	0.42	0.017	-0.201	0.58	0.032
Text4	0.091	0.061	0.296	0.484	0.052	-0.305	0.516	0.073
Text5	0.09	0.073	0.285	0.373	0.048	-0.369	0.627	0.106
Text6	0.091	0.04	0.258	0.559	0.039	-0.229	0.441	0.041
Picture1	0.092	0.015	0.171	0.642	0.017	-0.128	0.358	0.013
Picture2	0.092	0.007	0.14	0.983	0.012	-0.018	0.017	0
Picture3	0.091	0.017	-0.197	0.755	0.023	-0.112	0.245	0.01
Picture4	0.092	0.046	-0.218	0.352	0.028	0.296	0.648	0.069
Picture5	0.09	0.139	0.052	0.007	0.002	0.644	0.993	0.322
Positive	0.742	0.126	0.15	0.494	0.109	-0.152	0.506	0.149
Neutral	0.129	0.377	0.152	0.029	0.019	0.874	0.971	0.851
Negative	0.129	0.497	-1.021	1	0.871	0.001	0	0

The category positive (ctr1= 0.109, ctr2= 0.149) is described in the two dimensions nearly equal. Based on these results it could be said that dimension 1 goes from negative (k1= -1.021) to positive (k1= 0.15) and the second dimension goes from positive (k2= -0.152) to neutral (k2= 0.874). That means, when looking at Figure 1, all objects which are located in the lower right box are perceived positive. Also the objects are described through a dimension. The cluster text1 (ctr1= 0.747) is explained by dimension 1. However, there is no high difference to dimension 2, the clusters picture1 (ctr1= 0.017), picture2 (ctr1= 0.012) and picture3 (ctr1= 0.023) are also described by dimension 1. Concerning the position along this dimension it can be said that clusters picture1 (k1= 0.171) and picture2 (k1= 0.14) are in the positive area

of the dimension which means that they are associated with the positive category. In contrast, text1 (k1= -1.129) and picture3 (k1= -0.197) are located on the negative side of dimension 1 which leads to a negative association. Dimension 2 explains the clusters text2 (ctr2= 0.304) and picture5 (ctr2= 0.322) which are perceived neutral. Moreover, text5 (ctr2= 0.106) is lightly explained in this dimension which is identified as positive. Looking at the resulting correspondence graph (Fig. 1) most text clusters (text3, text4, text5, text6) and also two picture clusters (picture1, picture2) are located in the lower right box and therefore are perceived positive. Whereas text1 and picture3 are identified as negative and text2, picture4, picture5 are more neutral. Through the positions of the relevant objects and attributes, relationships between clusters and emotional categories can be conveyed.

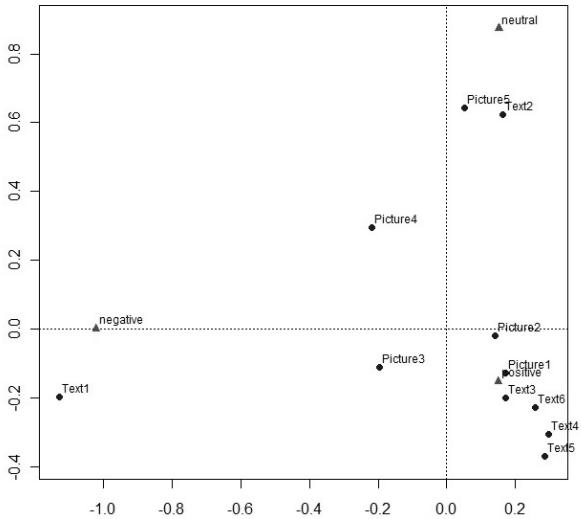


Fig. 1. Correspondence Analysis Graph

6 Conclusion

The website of the Albertina Museum offers text and pictures of artists and their paintings as well as pictures of the Albertina Museum itself and its history. Whereas the Technical Museum of Vienna provides pictures of technical objects and of interacting people. On the paintings of the artists the Albertina has little influence. However it might be possible to pick the best choice of available pictures.

Results show that most of the pictures used on the websites (43.14%) are associated with neutral emotion. Crucial for this allocation was the emotion nostalgic which was often used in context with pictures and is a neutral emotion. Only 27.46% of the pictures are perceived positive. And 29.41% of the analyzed pictures are viewed as negative. Having a deeper look at the clusters it could be said, that not only the object which is on the picture is crucial for the evoked emotion, as for example in cluster 3 and 4 heads of old men can be found or another example is that buildings are presented in all clusters. Therefore, the classification of the pictures might be laid

back on other variables such as low-level features. Pictures which are perceived negative have a high luminance. Neutral pictures show dark and only few colours. So marketers should avoid pictures with a high luminance and should use more pictures with grounded colours. In this case an adaption of the pictures towards a positive emotion might lead to a better visual appeal. As mentioned before, to pick the best choice of available pictures, is not only to find the best picture theme but also to have a look at the low-level features of a picture. Accordingly, further research regarding colour, brightness, orientation and so on would be necessary in order to get more detailed results and in succession to support marketers.

A circumstance that might influence the study negatively was the small size of the pictures in the card sorting study. Moreover, the psychographics of the respondents were not collected which might lead to different results between specialists in arts and technology or people who are not interested in art or technology at all. Concerning the text clusters, in this research the price cluster was associated with negative emotions which may not be surprising. Overall, it can be said that the text show a variety of emotions and nearly all are perceived positive. Therefore, the result shows a welcoming image, however prices are important and essential information which cannot be left out. As pointed out, verbal stimuli are used skilfully while a third of the pictorial stimuli are perceived negative. Therefore improvements are necessary to create a positive overall perception. Through a harmonious presence in the web the prospective tourists are impressed by the website which leads to a positive experience and induces trust and furthermore a relation to the offered touristic product.

7 References

Bagozzi, R., Gopinath, M. & Nyer, P. (1999). The Role of Emotions in Marketing. *Journal of Academy of Marketing Science* (27): 184-206.

Baloglu, S. & McCleary, K. (1999). A Model of Destination Image. *Annals of Tourism Research* 26(4): 868-897.

Bigné Alcaniz, E., Sánchez García, I. & Sanz Blas, S. (2009). The functional-psychological Continuum in the cognitive Image of a Destination: A confirmatory Analysis. *Tourism Management* (30): 715-723.

Choros, K. & Muskala, M. (2009). Block Map Technique for the Usability Evaluation of a Website. In N. Nguyen, R. Kowalczyk & S.-M. Chen (Ed.), *ICCCI*. Berlin Heidelberg: Springer-Verlag.

Dunn, J. R. & Schweitzer, M. E. (2005). Feeling and Believing: The Influence of Emotion on Trust. *Journal of Personality and Social Psychology* 88(5): 736-748.

Engel, J., Blackwell, R. & Miniard, P. (1995). *Consumer Behavior* (8 ed.). The Dryden Press.

Foscht, T. & Swoboda, B. (2007). *Käuferverhalten: Grundlagen- Perspektiven- Anwendungen* (3. ed.). Gabler-Verlag.

Ganglmair, A. & Wooliscroft, B. (2001). K-Means vs. Topology Representing Networks: Comparing Ease of Use Gaining Optimal Results with Reference to Data Input Order. In J. A. Mazanec, G. I. Crouch, J. B. Ritchie, and A. G. Woodside, *Consumer psychology of tourism, hospitality and leisure*. Wallingford, New York: CABI Publishing.

Grosspietsch, M. (2006). Perceived and projected Images of Rwanda: Visitor and international Tour Operator Perspectives. *Tourism Management* 27: 225-234.

Gunn, C. (1988). *Vacationscape: Designing tourist region.* New York: Van Nostrand Reinhold.

Haahti, A. & Komppula, R. (2006). Experience desing in tourism. In D. Buhalis & C. Costa (Eds.) *Toursim business frontiers*. Oxford: Elsevier.

Hong, S. & Kim, J. (2004). Architectural criteria for website evaluation - conceptural framework and empirical validation. *Behavoir & Information Technologie* 23(5): 337-357.

Izard, C. (1977). *Human Emotions.* New York: Plenum Press.

Kleinginna, P. R. & Kleinginna, A.M. (1981). A Categorized List of Emotion Definitions, with Suggestions for a Consensual Definition. *Motivation and Emotion* 5(4): 345-379

Lang, P., Bradley, M. & Cuthbert, B. (2008). International affective picture system (IAPS): Affective ratings of pictures and instructions manual. *Technical Report A-6*.

Lindgaard, G., Fernandes, G., Dudek, C. & Brown, J. (2006). Attention web designers: You have 50 milliseconds to make a good first impression! *Behavior & Information Technology* 25(2): 115-126.

Martinetz, T. & Schulten, K. (1991). A "Neural-Gas" Network Learns Typologies. In T. Kohonen, K. Makisara, O. Simula, and J. Kangas (Eds.), *Articial Neural Networks*. North-Holland: Amsterdam.

Mazanec, J. (2010). Tourism-Receiving Countries in Connotative Google Space. *Journal of Travel Research* 49: 501-512.

Meyer, W.-U., Reisenzein, R. & Schützwohl, A. (1993). *Einführung in die Emotionspsychologie* (Vol. I). Bern: Huber.

Moss, G., Gunn, R. & Heller, J. (2006). Some men like it black, some women like it pink: consumer implications of differences in male and female website design. *Journal of Consumer Behaviour* 5: 328-341.

Pan, B. & Fesenmaier, D. R. (2006). Online Information Search. *Annals of Tourism Research* 33(3): 809-832.

Pike, S. (2002). Destination Image Analysis - a review of 142 papers from 1973-2000. *Tourism Management* 23: 541-549.

Plutchik, R. (1962). *The Emotions: Facts, Theories and a new Model.* New York: Random House.

Roseman, I. J. (2001). A Model of Appraisal in the Emotion System: Integrating Theory, Research and Applications. In K.R. Scherer, A.`Schorr, and T. Johnstone (Eds), *Appraisal Processes in Emotion: Theory, Methods, Research*. Oxford University Press.

Ruso, B. & Mayer, M. (2006). Evolutionäres Webdesign. In *Raum - Heimat - fremde und vertraute Welt*. Wien: LIT Verlag.

Russell, J. & Mehrabian, A. (1977). Evidence for a Three-Factor-Theory of Emotions. *Journal of Research in Personality* 11: 273-294.

Scherer, K. R. (2003). Vocal communication of emotion: a review of research paradigms. *Speech Communication* 40: 227-256.

Scherer, K. (2005). What are emotions? And how can they be measured? *Social Science Information* 44(4): 695-729.

Schweiger, G. & Wiklicky, W. (1986). Entwicklung eines standardisierten Verfahrens der Nonverbalen Imagemessung bei Markenartikeln. *Werbeforschung & Praxis* 5: 175-180.

Sheldon, P. (1997). *Tourism Information Technology.* Wallingford: CAB International.

Strapparava, C. & Mihalcea, R. (2008). Learining to Identify Emotions in Text. In *SAC*. Ceará Brazil.

Ulich, D. & Mayring, P. (1992). *Psychologie der Emotionen.* Stuttgart: W. Kohlhammer GmbH.

Weideman, M. & Mgidana, M. (2004). Website Navigation Architectures and their Effect on Website Visibility: A Literature Survey. In *Proceedings of SAICSIT*.

Xiang, Z. & Pan, B. (2009). Travel Queries on Cities in the United States: Implications for Search Engine Marketing in Tourism. In *Information and Communication Technologies in Tourism 2009*. Vienna, Springer.

Ye, H. & Tussyadiah, I. P. (2011). Destination visual image and expectation of experience. *Journal of Travel & Tourism Marketing* 28: 12-144.

Zhang, L., Pan, B., Smith, W. & Li, X. (2009). Travelers' Use of Online Reviews and Recommendations: A Qualitative Study. *Journal of Information Technology and Tourism, 11*(2): 157-167.

eTourism: Multimedia Brochures and Communication Chains

Nicoletta Di Blas, and
Paolo Paolini

HOC-LAB, Department of Electronics and Information, Politecnico di Milano (Italy)
{nicoletta.diblas; paolo.paolini}@polimi.it

Abstract

This paper is about multimedia communication, for Tourism. Two original concepts are introduced: Multimedia Brochure and Communication Chain. A Multimedia Brochure is an agile and versatile interactive multimedia available for all kinds of devices and technologies: it can be used by any "actor" of tourism in order to make itself "visible" and "persuasive" over Internet. An innovative authoring environment and a streamlined production workflow allow effective production, combining good quality and low cost. The Communication Chain allows the interchange of (fragments of) multimedia brochures across different actors (e.g. cultural heritage, accommodation, food, shopping, sport, entertainment,…), avoiding (or at least minimizing) duplication, confusion and low quality, the occurrence which are frequent to be found today. The COMFIT project demonstrates, on a medium scale, the viability and the effectiveness of the approach.

Keywords: multimedia, multichannel, authoring tool, brochures, eTourism.

1 Introduction

The use of Internet has greatly affected the tourism industry, like many other fields. Roughly speaking, we can classify the application if ICT to tourism in 4 broad categories: (1) *Back-office* (accounting, management, and similar): this has been the first, but clearly the less relevant, application area for tourism. (2) *Operations* (booking, reservation, confirmation, etc.): they can be carried on directly or through intermediaries (which are currently dominating the field). This is probably the most important application today and it has greatly influenced and modified the industry. However, it is beyond the scope of this paper. (3) *Promotion/Marketing/Sale*: search engines, banners, links of all kinds are involved. Again, this is very relevant, but it is beyond the scope of this paper. (4) *Communication on site and on Internet* ("internet presence"): multimedia information (text, audio, images, videos…) for capturing the attention and persuading the users. These applications can be used via a PC, or an iPad (tablet), or a smart phone, an iPod, etc. or they can be delivered as interactive installations. They are the focus of this paper. Tourism communication can be about various subjects: a destination, a specific item of interest (a church, a monument, an archeological park…), an actor (a hotel, a restaurant, a shop…), etc. Multimedia tourism communication is quite widespread on the Internet today, but several problems persist:

- *Quality*: with a few exceptions (mostly regarding big operators), the quality of tourism communication is quite poor. Besides the operational information (maps, fees, reservations info, etc.), there is little room for meaningful, persuasive flow of information. The actual persuasion (i.e. convincing the tourist to visit a

destination, to book a hotel, to pay a visit to a shop, to take a specific tour…) is mainly left to traditional media, like magazines and guides. Minor destinations and actors, that can't afford to hire professional staff, sometimes perform naïve attempts that may even have a negative impact.

- *Costs*: good multimedia communication can be expensive. Even without considering the costs of taking pictures and videos, the sheer cost of professional web design and text-writing is very high and unaffordable for many small actors.
- *Scope:* Let us consider the case of a small town with a few natural attractions and monuments. The communication for this town should cover the general attractions, the excursions, the cultural sites, but, in order to be complete, it should also include the accommodations, the restaurants, the shops, etc. On the other hand, the communication of a specific hotel should include not only the hotel's description but also information about the territory, the attractions, etc. This is because an internet user may end up directly on this website. Therefore, in tourism, it is often the case that each actor has to cover a number of subjects, with duplications and overlaps. How many hotels of the same town describe the same monuments, the same attractions, the same history, the same fairs, etc.?
- *Technology*: technology continuously evolves. Beside "traditional" websites there are, today, tablets, smart-phones, podcasts, social spaces, YouTube, etc. Everything must be kept into account if an effective persuasion strategy, reaching all the potential targets, has to be implemented. Many tourism-related actors have problems into dealing with this continuous evolution.
- *Coordination:* the tourism field is quite fragmented and suffers from a lack of coordination among the actors. Even in those – few – cases in which the public administration takes upon itself the task of communicating the destination as a whole (typically via a "portal"), the problem persists. The portal is either a rigid frame that in the end discourages the users, or a "patchwork" of only poorly related pieces of content. There are a few good examples of "coordination", at international level (e.g. www.discoversouthcarolina.com), but they are exceptions (mostly to be found in North America).

2 Communication Chains

The notion of "communication chain" derives from the industrial notion of "supply-chain". A complex object like a car, for example, is the result of the assembly of several parts (an engine, a frame, an electric system, …). The car maker creates the overall design and directly manufactures a few crucial parts; specialized external manufacturers provide the others (e.g. the navigation lights). These external manufacturers, in the past, used to work on tailored design provided by car makers. The situation, today, is reverted: they work according to their own design, which is then 'forcedly' adopted by car makers. This is why many components of today's cars are the same across different brands and models: they are designed and manufactured by the same producer.

Let us now consider the "manufacturing" of multimedia communication for tourism. Let us assume that a cultural heritage institution creates content for a specific subject: e.g. the cathedral of a town. If culture per se is the only issue at stake, an accurate and persuasive description of the cathedral, in all its aspects, would be sufficient. If,

instead, the aim is *to attract* tourists, something more is needed: it would be advisable to add something about the town, as well as some historical notes, an explanation of the urban setting, information on relevant people of the past who were somehow related to the cathedral, etc. Since the history of the church is interwoven with the history of the town, this enlargement of scope is necessary in order to provide good quality communication. Let us now assume that the local administration wants to promote the historic center of this town. There will be an overall description, maps, itineraries, history, etc. Of course, the main cultural attractions like the cathedral will be included too, along with the best shops, cafés, hotels, etc. Let us now consider the communication of a hotel, in the same town, located near the cathedral. Besides illustrating the quality of the hotel, it will include the cathedral's description, the town's history, the attractions, and so on. *What can we learn from this example? That communication with a narrow scope cannot work and that each actor needs to cover information which is not its own: content should be shared and moved among several actors.*

Unfortunately, the current situation (in Italy, as in most other countries) is far from being optimal. First of all, there is a multiplication of efforts: each actor who needs, for example, a description of the town's history, will do her own work of research and writing. In addition, there is a lack of focus on core competences: hotel keepers are very good at communicating their establishments, shop keepers at describing their goods, cultural institutions at introducing their exhibits. They should focus their effort on the content for which they have specific expertise, instead of covering a variety of subjects.

The above described situation is well known to professionals working in eTourism. The two most popular solutions are: (1) coordinated portals and (2) links across sites. In the first case big, all-comprehensive portals are created, with slots for culture, leisure time, accommodations, etc. to be filled by the different actors (either directly or through intermediaries). This solution does not provide a satisfactory result for several reasons: standardized slots do not allow for strong, customized communication and thus each actor experiences a loss of identity, becoming "an item in a list". In addition, search engines do not work well with large portals: a hotel keeper wants to be well visible and clearly distinguishable when users look for "a hotel near the cathedral in town X"! As a matter of fact, big portals work well for very little destinations only, where strong actors requiring strong, personalized communication, are few. In other cases (like in Italy), portals turn out to be no more than directories of addresses and telephone numbers, useful for users who already know what to look for. Links across sites are another possible solution, but with evident drawbacks: users get disappointed at being bounced across different sites in order to get the whole picture, through an excessive, not manageable, number of links. Given all the above, we propose a different solution for eTourism based on the following main ingredients:

- *Multimedia brochure*: an agile kind of communication that complements traditional websites; an interactive multimedia narrative, combining different types of images, video, audio, text, etc. More than providing information or supporting operations (e.g. booking), it provides persuasion.

- *Costs and time containment*: a multimedia brochure, created with our authoring environment, can be generated at a fraction of costs of a traditional website, and in 1 or 2 weeks.
- *Narrative style*: the style is narrative and persuasion-oriented. It leads the user through a story, at the same time allowing engaging and effective interaction, skipping not interesting portions and focusing on the interesting subjects.
- *Multi-channel/multi-technology*: a multimedia brochure can be delivered through all devices (including PCs, smart-phones, iPad, iPod, tablets, MP3 readers, ordinary phones, multi-touch tables...) and all types of connectivity.
- *Brochure types*: different types of brochures can be created; "episodic" brochures, describing specific items (a church, an exhibition, a hotel, ...), "thematic" brochures, describing overall themes (e.g. sport, shopping, culture, ...),"itinerary" brochures, suggesting specific ways of traveling, and "recommendation" brochures, suggesting how to spend the day, what to visit in a city, etc.
- *Aggregators*: when several multimedia brochures are created for a destination, it is necessary to aggregate them in a manner suitable for the user. Several aggregators can be created with different styles: geographic (showing on a map where everything is), highlights (where a restricted number of brochures is presented), exploratory (allowing the users to interactively explore the material using different selection criteria), itinerary (suggesting some brochures in a specific order), emotional (combining different brochures in a suggestive manner), etc.
- *Communication chain*: in order to provide coordination among the different MM-brochures of a destination (or a region), they can be linked to each other (as ordinary websites do), or they can be assembled into mini-portals (the "aggregators"). But their peculiar and innovative feature is that they can be also "disassembled", allowing the reuse of specific parts, like in a building blocks' game, and "reassembled" to suit a specific purpose.
  Let us imagine a hotel keeper who wants to include in her communication something about the lake, the cathedral and the history of her town. Ordinarily she could either create links to other sites or generate new content to cover the additional subjects. With the communication chain approach, instead, she can extract from existing multimedia brochures, created by others, the relevant fragments of content, and "reuse" them in her own brochure. A small degree of adaptation is needed: for example, she could substitute generic pictures of the lake, with pictures where her hotel is clearly visible.

There are several advantages provided by the communication chain approach: quality (since the borrowed content is created by specialists), time and cost saving, communication strength. There is also an advantage for the destination as a whole: an overall coordinated communication (where a user can find the same "history of the town" across different sites) is stronger and more effective than a chaotic one. Conversely, when a general brochure needs a specific piece of content (e.g. a hotel's description) it can directly borrow the material from the corresponding multimedia brochure. In the next section we will describe how MM-brochures can be created: the format, the production workflow and the technology. We will then present an eTourism project where the "communication chain" approach is being exploited on a medium scale.

3 Format and Production

Multimedia brochures (fig. 1) are a lightweight form of communication, meant to generate interest and curiosity for a subject, rather than to "inform" in a traditional sense (Campione et al., 2011; Di Blas, Paolini, & Rubegni, 2010; Rubegni et al., 2010). They are composed by audio, visual communication (slideshow, videos, animations, 360° pictures…), and text. They are delivered via different devices and technologies (fig. 2).

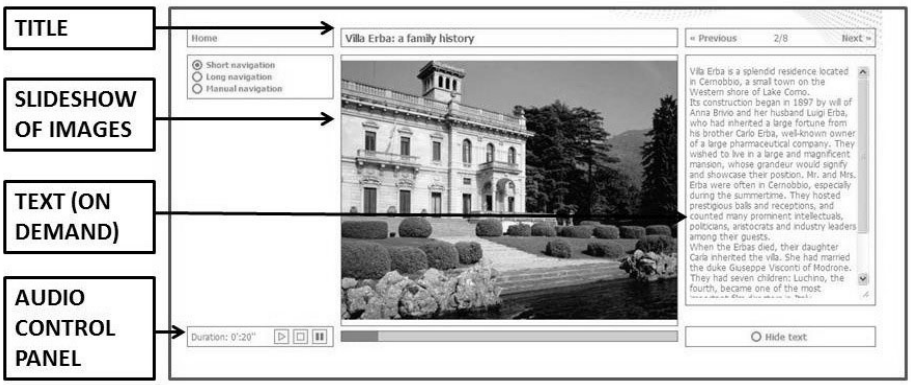


Fig. 1. A multimedia brochure about Villa Erba (Como, Italy); the web version.

Fig. 2. A 360° interactive picture, on iPad (www.manraylugano.ch).

We have developed a number of different "formats" for MM-brochures exploiting different information architectures (content organization) and different interaction

features (depending on the device). In addition, MM-brochures can be interconnected, i.e. a piece of content from a brochure can be linked to another brochure (fig. 3).

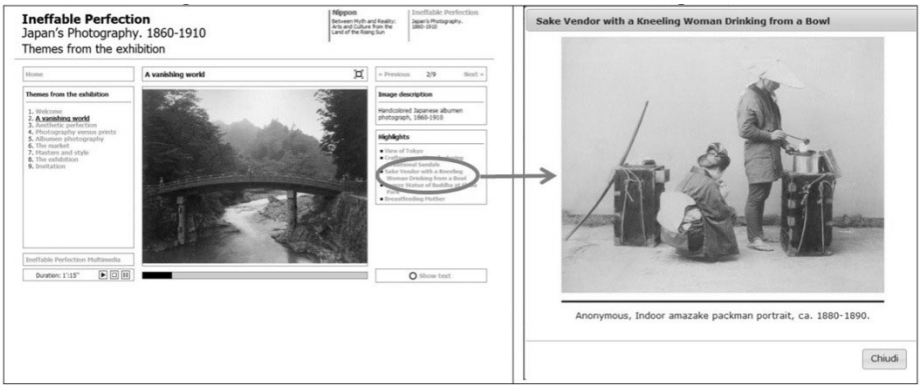


Fig. 3. The thematic brochure of an exhibition (left) is linked to a catalogue brochure (right). www.nipponlugano.ch.

When several brochures are available, they can be put together in what we call an "aggregator", i.e. an interface offering a coordinated access to large content. Aggregators can take different aspects: geographical maps, itineraries, catalogues, etc. They can be enhanced by the use of faceted search (Marchionini, 2006; Tunkelang, 2009; Spagnolo et al, 2010) and/or word clouds and tag clouds.

Fig. 4. "Nippon at a glance": an aggregator for 4 different exhibitions. A mosaic of thumbnails (linked to fragments of MM-brochures) with tag clouds and word clouds

To keep down the production costs, a streamlined production workflow is crucial. Our production workflow starts with (1) a structured interview to an expert on the subject, either via Skype/phone (more cost-effective) or in presence. After that, (2) an editorial plan of the brochure is sketched and relevant topics which emerged during the interview are organized according to a suitable format (i.e. linear sequence, hierarchical organization, linked brochures, etc.). When the editorial plan is defined, two actions run in parallel: (3.1) the interview's transcript is split into chunks according to the editorial plan; (3.2) the visual communication is prepared, looking for suitable images/videos/animations to illustrate the subject. If communication problems are detected (as the discovery of gaps, contradictions, etc.) an additional interview (4) with the expert may be necessary. Once the text for all fragments is ready, (5) a first version of the audios is created (in-house, typically using a voice synthesizer). All the above leads to (6) a first version of the brochure: the communication quality is checked and amendments are performed. Once the final text is approved, the recording (7) with a professional speaker is carried on. Steps 6 and 7 eventually lead to (8) the final version. The average production time for a brochure is a month, with variable costs – always low (from 500 Euros for tiny MM-brochures, to 4,000 Euros for large ones, with 1,000 Euros as average).

MM-brochures are created using the "1001stories" authoring environment, first developed in 2006 and then refined over the years (Campione et al., 2011). The current version is a highly sophisticated system that allows creating applications for a number of technologies: web for PC, web for mobile, podcasts, iPad (and general tablets), multi-touch tables, YouTube, standard phone (with audio-only content), etc.

Moving content across devices and channels requires more than adapting to different screen sizes: content adaptation and interaction tailoring is also needed. Our effort is to keep the process simple and the costs down. The core idea is to separate "authoring" in strict sense (text, images, audio creation) from the generation of specific applications, tuned for specific devices and/or specific user experiences. "1001stories" environment has 3 main components: (1) an authoring environment where the various pieces of content are authored; (2) a number of generation engines that generate the proper information architecture for a specific purpose; (3) a number of delivery engines, actually implementing the various interactive formats over various platforms. Information architectures are described via XML, according to a set of XML-schemas. The technical environment's complexity is all internal: the author finds it extremely easy to use (half-day training is enough).

4 Communication Chain at Work: the COMFIT Project

The COMFIT project is the first attempt, on a medium scale, to fully exploit the notions of multimedia brochures and communication chain, in the field of eTourism. The project is part of INTERREG, a cooperation program between Italy and Switzerland, fostering the development of innovative solutions for cooperating across the border. Project financing is a combination of funding from the European Commission, the participating regions and the partners. The regions are Lombardy (specifically, the Como area, near the Swiss border) and Ticino (the region of Lugano, near the Italian border). Partners for Italy are Politecnico di Milano, Como county and Centro Volta (a development nonprofit institution). Swiss partners are USI (the

University of Italian Switzerland) and T-Next (a leading ICT company of the area). The project started in November 2011 and will end by April 2012. In order to obtain a real impact on the tourism industry big efforts have been carried on to involve local administrations, associations, institutions and individual operators. During its lifetime, COMFIT will cover the full cost of a number of MM-brochures; after its completion, a service center will do the job, at moderate costs (for sustainability). The goal of COMFIT is to develop enough brochures and aggregators to demonstrate the feasibility of the communication-chain approach, its quality and value for the industry. Table 1 shows the plan and the development of multimedia brochures as at October 15th, 2011. In the first round each brochure is developed individually; in the second round (starting December 2011) we will systematically apply the communication-chain notion, "injecting" into each brochure parts of other brochures. Also, starting from December 2011, we will build aggregators, i.e. mini-portals offering combined visibility of groups of brochures (see figure 4).

Table 1. The COMFIT project's numbers (of multimedia brochures), at Sept. 2, 2011

Subject	**Total**	Completed	Under development	Additional goal
Total	**136**	**53**	**53**	**30**
Culture	**33**	15	18	0
Natural landscapes	**18**	2	11	5
Accommodation	**25**	10	5	10
Commerce/shopping	**21**	2	4	15
Restaurants/food	**7**	5	2	10
Services	**15**	7	8	0
Sport	**6**	2	2	2
Villas	**6**	4	2	0
General themes	**9**	6	0	3

Starting from February 2012, internet promotion for the brochures and the aggregators will conclude the project. The bet is twofold: (1) to demonstrate that the approach is effective; (2) to convince the tourism concerned actors of the two regions that the result is worth a small investment and that a service center, as legacy of the project, is needed. The service center will continue the job, using revenues from the eTourism actors themselves (in other words, they will pay for the service).

Table 2 shows the abstracts of a few multimedia brochures of different kinds. The reader is invited to figure out how content exchange (communication chain) will greatly improve the quality of each brochure.

Each MM-brochure, built as an interactive narrative, lasts between 6 to 15 minutes, with fragments lasting between 20 seconds and 1-2 minutes. Each brochure can be used by its owner as an independent website or as a companion for a traditional website. In addition, there will be a number of mini-portals (aggregators) clustering the brochures in meaningful ways for tourists (or users in a broader sense). The technology for the aggregators has already been created and the actual building of the mini-portals will be carried on in the last 4 months of the project. Users will locate the interesting brochures in different ways: by geographic locations, visual appeal, word

clouds, combinations of facets, etc. Exploration is the main paradigm, rather than search (De Caro et al., 2010; Spagnolo et al,, 2010).

Table 2. Abstracts of a number of COMFIT brochures about the Como area (Italy).

	TITLE*: Land flavors, Lake flavors* ACTOR: the Consortium "Land flavors, Lake flavors" ABSTRACT: the consortium wishes to preserve and promote to a larger audience the rich food traditions of the Como area.
	TITLE*: Two lakes* ACTOR: the two counties of Como and Ticino ABSTRACT: the area including Como and Lugano lakes is a wonderful trans-national destination. The beautiful landscapes, the high quality of prestigious accommodation and the quality of services, make this area especially suitable for large gatherings and events.
	TITLE*: Lugano and its Lake* ACTOR: Ticino County ABSTRACT: surrounded by mountains and on the shores of a beautiful lake, Lugano combines the quality of a modern city and of a small, traditional, town.
	TITLE*: LAC, the new cultural center of Lugano* ACTOR: the LAC foundation ABSTRACT: LAC will be the new cultural center of Lugano. It will host a museum, galleries, an auditorium, art exhibitions, and many other cultural attractions. It will organize events in cooperation with the cultural institutions of Lugano.
	TITLE*: Cultures Museum* ACTOR: the Museum ABSTRACT: the "Museo delle Culture" is devoted to promote the culture of extra-European continents. Its permanent collection, the nucleus of which was created by an artist who was also a collector, provides an engaging experience of ethnographic artistic expressions.
	TITLE: *Alessandro Volta* ACTOR: City of Como – Culture Department ABSTRACT: Alessandro Volta, born in Como, is credited with the invention and creation of the first prototype of battery. The brochure revives his biography and his relationship with the city.

	TITLE: *Comacina Island* ACTOR: "Comacina Island" cultural Association ABSTRACT: the Comacina Island is rich in history, art and traditions, since Roman times. It is associated to legends, myths and even spells. Today, it still attracts visitors, for the beauty of its landscapes, archeology and food.
	TITLE: *"Monti Lariani" trails* ACTOR: CAI-COMO, the local branch of the Italian Alpine Club ABSTRACT: "Monti Lariani" are middle-high mountain pastures, used for centuries as summer cattle pastures. CAI has restored trails and itineraries, allowing tourists to enjoy wonderful excursions with breathtaking views of the Como Lake.
	TITLE: *Belvedere Hotel - Bellagio* ACTOR: Hotel owner ABSTRACT: the Hotel Belvedere provides a familiar, relaxing and comfortable hospitality in Bellagio, one of the most charming locations on the Como lake. Owned by the same family since its foundation, the hotel provides today modern comfort and facilities.
	TITLE: *Palace Hotel - Como* ACTOR: Hotel keeper ABSTRACT: A charming, XIX century building, in liberty style, hosts this historical luxury accommodation in Como. The hotel is located in downtown Como, near the lake shore. Professional staff ensures top services for the guests on holiday or business trips.
	TITLE: *Comacina Island Inn* ACTOR: Inn owner ABSTRACT: the inn provides a unique food experience: a menu preserving the traditions and flavors of the island. The restaurant is the only one in this charming and mysterious place.
	TITLE: *The Wolf's "Crotto"* ACTOR: Owner ABSTRACT: "Crotti" were natural caves, used both in the Como and Lugano areas to preserve food (like salami), that later evolved into taverns. The Wolf's Crotto is a well renowned restaurant in the Como area, maintaining connections with the local food traditions.
	TITLE: *Como-Brunate Funicular* ACTOR: ATM, the city company that runs the service ABSTRACT: the funicular service, connecting Como to the nearby low mountain of Brunate, was inaugurated in 1894 using vapor energy. The modern funicular, based on an electric engine, provides an exciting and charming experience.

	TITLE: *Rowing Association "Lario"* ACTOR: The sport association "Società Canottieri Lario" ABSTRACT: the rowing association "Lario" (the ancient name of the Como lake), was founded in 1891. It is the home of 200 athletes, including national and international champions. It also provides sport possibilities for tourists and rehabilitation courses for people affected by disabilities.
	TITLE: *Art and Culture in the Como area* ACTOR: The Como county ABSTRACT: The Como area offers much more than just a lake or a landscape. Different cultures left their evidence, from prehistoric artifacts to "rational" architecture of the 20s, magnificent churches and stately villas.
	TITLE: *The "Lario" territory* ACTOR: The Como county ABSTRACT: Come lake is a landscape of breathtaking beauty. Since ancient time, men have been attracted by the lake shores. Even today, it is a preferred destination for international tourists looking for beauty, relax, food and fun.
	TITLE: *Villa Olmo* ACTOR: The City of Como ABSTRACT: Villa Olmo, with its beautiful building and charming garden, has been declared part of the World Heritage. In recent years it has become the location for the most important cultural events in Como.
	TITLE: *Villa Erba* ACTOR: The City of Como ABSTRACT: The magnificent Villa Erba is located near Cernobbio, on the western shore of the Como lake. Currently used for important events and gatherings, it was the family retreat of the Visconti family, descendants of the dynasty that ruled Milan in the late middle age. Luchino Visconti, the famous movie director (e.g. "The Gattopardo"), spent his infancy there, and many memories from that time and place can be detected in his movies.

5 Conclusions and Future Work

In our perception, the situation of ICT-based communication for tourism suffers from a number of problems, among which the most serious are the lack of coordination among actors (each one takes care of its own communication) and a (needless) multiplication of efforts. The result is a chaotic overall picture of a destination. In addition, there is a difficulty in keeping up with technology developments and in combining "permanent communication" (e.g. about a territory and its permanent features) with "temporary communication" (e.g. about an event, an exhibition, a congress, ...). Major actors (possibly, but not always) can cope with this situation and

afford the costs of multimedia visibility over Internet. Small actors (and sometimes whole destinations) are not keeping up. In this paper we put forth a possible solution that combines two innovative features: (1) "multimedia brochures" and (2) the "communication chain" approach. With our method and tools, MM-brochures can be created at low cost and quick time, by a staff of non-technical people. The result is an agile form of communication that can be delivered through *all* the relevant technological devices and channels. The communication chain approach is a way to effectively coordinate different pieces of content, borrowing them from different sources and recombining them in meaningful ways. We are adopting this approach in the COMFIT project that promotes a medium-sized area (Como and Ticino). When the project will be over, a permanent service centre will provide support to all the actors of the area, to keep the effort alive. We expect high visibility of the area on the internet, through a variety of technologies, and with limited costs (shared among the different protagonists), and relevant impact on the tourism market.

References

Campione, P., Di Blas, N., Franciolli, M., Negrini, M. & Paolini, P. (2011). A "Smart" Authoring and Delivery Tool for Multichannel Communication. In D. Bearman, and J. Trant. (Eds.), *Museums and the Web 2010*. Philadelphia, Pennsylvania, U.S.A.

De Caro, S., Di Blas, N. & Spagnolo, L. (2010). In Search of Novel Ways to Design Large Cultural Web Sites. In D. Bearman J. & Trant (Eds.), *Museums and the Web 2010.* Denver, Colorado, U.S.A.

Di Blas, N. & Paolini, P. (2010). Multimedia for Cultural Heritage Communication: Adapting Content to Context. In *Proceedings of the Esa Culture Bocconi 2010 conference*, Milano, October 7-9.

Di Blas, N., Paolini, P. & Rubegni, E. (2010). How to Build Multi-Media and Multi-Channel Corporate Brochures. In *Proceedings of the IEEE Professional Communication Conference*, July 7 - 9, 2010. Netherlands, University of Twente.

Marchionini, G. (2006). Exploratory search: from finding to understanding. *Commun. ACM* 49(4): 41-46.

Rubegni, E., Di Blas, N., Paolini, P. & Sabiescu, A. (2010). A Format to design Narrative Multimedia Applications for Cultural Heritage Communication. In *Proceedings of SAC'10*, March 22-26, 2010, Sierre, Switzerland.

Spagnolo, L., Bolchini, D., Paolini, P. & Di Blas, N. (2010). Beyond Findability: Search-Enhanced Information Architecture for Content-Intensive RIAs. *Journal of Information Architecture* 2(1): 19-36.

Tunkelang, D. (2009). *Faceted Search. Synthesis Lectures on Information Concepts, Re-trieval, and Services*. Morgan and Claypool Publishers.

Acknowledgements

This work is done in the frame of the COMFIT project, (ID Project 13737544; MISURA 2.1), an INTERREG Project Italy- Switzerland about tourism promotion in the Ticino and Como area.

An Analysis of the Perceived Value of Touristic Location Based Services

Barbara Neuhofer

School of Tourism
Bournemouth University, United Kingdom
bneuhofer@bournemouth.ac.uk

Abstract

Location based services have gained increasing importance for the tourism industry in recent years. However, a real breakthrough to the mass market has not been fully achieved yet, due to a number of potential reasons, inter alia, institutional circumstances, drawbacks in technology and missing understanding of the real values inherent in location based services. Preliminary literature still indicates a lack of research in terms of understanding consumers' perceptions towards location based services from a consumer-centric rather than a technology-focused perspective. Therefore, the aim of this particular study is to investigate the actual consumer perceived value of touristic location based services. An exploratory sequential strategy is applied to develop and test a six-dimensional measurement scale. The findings suggest the extension of the original model by two new value dimensions, namely informational and convenience value. Further need for research and managerial implications are provided.

Keywords: Mobile Services; Location Based Services; Perceived Value; Consumer-centric; Tourism;

1 Introduction

Over the last decade, the continuous advances in information and communication technologies caused a growing proliferation of mobile technologies. The gained mobility in combination with useful applications thereby implied increasing relevance of such services to the tourism industry. The benefits inherent in mobile technologies, such as the delivery of relevant content in real-time and on demand to the tourist consumer without geographical and temporal restrictions (Pihlström, 2008) have rendered mobile services a valuable information channel in tourism. In particular, location based services can be regarded as key services of the 21st century (Egger & Jooss, 2010), as they enable ubiquitous information retrieval (Balasubramanian et al., 2002) and the integration of GPS, allowing geographical positioning and access to location and context relevant information.

Although the high potentials of use are well estimated, a breakthrough to the consumer mass-market is still awaited. Apart from external barriers initially conditioning the adoption of mobile services, such as technological insufficiencies, underlying circumstances of the telecommunication industry or monetary hurdles (Göll et al., 2010), the poor adoption rates can also be explained by attitudinal barriers in terms of lacking understanding of the question of what in particular makes consumers adopt these new technologies. In other words, what are the decisive values that someone would use location based services? Hitherto, it appears that the question of what specific value location based services deliver is not fully explored yet.

Despite the increasing significance of mobile services for society and industry, to date only limited research has been conducted in terms of investigating mobile services from a consumer-centric perspective (Sweeny & Soutar, 2001). A number of authors (Dickinger, 2006; Rasinger et al., 2007; Pihlström, 2008) argue that most of the literature in the field has predominantly been technology-focused so far by applying prominent models, such as the TAM (Davis, 1989) to assess technology adoption. However, Pura (2005) claims that in the context of mobile services, there is great need for understanding end-users, their relative value perceptions and emotions derived from using a service. This is in line with a number of researchers (Komulainen et al., 2004; Lubbe & Louw, 2010) who highlight the necessity for identifying the values that trigger the intention to use mobile services. This knowledge is particularly essential for the actual creation of value. Once perceived values inherent in mobile services are understood, these can be specifically communicated to both, current users as well as non-users who might adopt mobile services in the future (Pura, 2005).

2 Theoretical Background

2.1 Location Based Services in Tourism

Location based services (LBS) have received increasing attention over the past years, mostly due to the recognition of the benefits gained by the geographical localisation and subsequent provision of relevant information (Pura, 2005). However, a number of recent studies (Pihlström, 2008; Wittmer et al., 2007) have witnessed a remaining breakthrough of location based services to the critical mass. As a matter of fact, market research affirms scarce adoption rates by revealing low numbers of people actually making use of mobile services. Although the mobile phone penetration rate for the EU-27 constitutes 122 % (RTR, 2010), the smart phone penetration rates are comparatively low among the EU5 (UK, France, Germany, Spain, Italy), with Italy having the highest penetration rate of 32 % and France the lowest with 15.2 % (Blanford, 2010). The major impediment thereby constitutes high costs (61.7 %), followed by the missing comprehension of the added value (36.1 %) (Statista, 2010). As a result, the relatively poor adoption rates of mobile services lead to the assumption that much more research is still needed to fully comprehend the real value that consumers gain from using these services.

However, Pedersen (2005) argues that existing models in current literature explain the adoption and use of mobile services only to an insufficient extent. Most studies conducted so far have focused on technological aspects of mobile services, such as the Technology Acceptance Model (Davis, 1989), the UTAUT (Venkatesh et al., 2003), the Technology Readiness (Davis et al., 1989) or the Innovation Diffusion Theory (Rogers, 2003) to name but a few.

However, despite the great attention in literature, these models have also received criticism in the course of the past 20 years (Bigné et al., 2010) and have often been subject of revision or extension with non-technology variables. As a result, instead of extending or reworking the TAM model yet another time in order to be applicable to the respective research context, the present study advocates the application of the perceived value construct instead. This is supported by Wittmer et al. (2007) who

claim that research must become less technology-focused and more end-user oriented, which is in line with Kleijnen et al. (2004) who confirm a huge lack of research focusing on the consumer perspective. Hence, the present lack of breakthrough of location based services coupled with researchers' demands for more consumer-centric work in order to understand the full complexity of the values inherent in location based services, provide the rationale for the present study to explore the perceived value of touristic location based services.

2.2 The Framework of Perceived Value

According to Bettman et al. (1998) perceived value can be described as the value that consumers receive by the use of a product or service. Perceived value originally derives from a number of earlier notions, such as perceived price, perceived service quality and perceptions (DeSarbo et al., 2001; Oh, 1999). Pihlström (2008) however argues that perceived value has substantially replaced most of these traditional notions to date.

In early literature, the concept of perceived value was regarded as a one or two-fold construct, whereby the emphasis was initially put on the estimation of the overall perceived value and later on the benefit-sacrifice continuum (Zeithaml, 1988). However, in recent years, the two-fold perspective of perceived value has come under increasing attack. Opponents to this simplistic approach argue that due to its dynamic nature, perceived value cannot be regarded as a compromise between getting and giving, but rather constitutes a consideration of irrational and subjective benefits based on affective emotional reactions, experienced by the individual (Sweeney & Soutar, 2001).

As a matter of fact, perceived value has undergone a transition from a simplistic model to a multi-dimensional construct (Sweeney and Soutar, 2001). Sheth et al. (1991a/b) probably provide one of the most seminal contributions to perceived value, by developing a five-dimensional measurement scale, consisting of *functional, social, emotional, epistemic* and *conditional value*. These original five dimensions have served as a basis for a number of studies in the past (e.g. Sweeney & Soutar, 2001; Pihlström, 2008, Wang et al., 2004). Sweeny and Soutar (2001) claim that Sheth et al.'s (1991b) model constitutes the currently best existing value construct applicable to different fields, however, indicate the potential for extension.

Pihlström (2008) for instance, applied Sheth et al.'s (1991b) original model and highlighted the need for adapting the theoretical framework to the context of mobile services, for the purpose of which a six-dimensional scale was created. Pihlström (2008) points out that for new electronic services a comprehensive conceptualisation is required, especially since mobile services may not be purely used for utilitarian but rather for experiential and hedonic reasons, which has been supported by various empirical studies (e.g. Pura, 2005; Yang & Jolly, 2008).

Thus, for the scope of the present study the development of an appropriate multi-dimensional scale that meets the requirements of touristic location based services is imperative. In order to incorporate all value dimensions which best support the investigation of touristic location based services, a new research model had to be created for this study, which was based on the theoretical foundations of two existing

models by Sheth et al. (1991a/b) and by Pihlström (2008). The conflation of both models allowed adopting a six-dimensional value scale by incorporating *functional, social, emotional, epistemic, conditional* and *monetary value.*

In addition to that, perceived value constructs often comprise so-called moderator and mediator variables. In this study, *attitude* as mediating variable to investigate users' attitudes towards using LBS as well as the outcome variable *behavioural intention* to determine users' predispositions towards using LBS, were included in the model. In due consideration of the exploratory nature of this research, a correlation between each single value dimension and attitude respectively behavioural intention was hypothesised, as well as a correlation between attitude as strong predictor of behavioural intention, leading to a total of 13 proposed hypotheses.

3 Methodology

The methodological approach of the present research consisted of two major steps. The preliminary research process allowed developing a measurement scale, which was applied in the second step to measure users' perceived values of touristic LBS. Complying with the requirements of researching mobile services, qualitative and quantitative data had to be gathered, a process which is commonly referred to as mixed methods approach (Creswell, 2003). Mixed methods were applied in the frame of a sequential exploratory strategy, which is commonly adopted when exploring new phenomena, testing elements or refining test instruments (Creswell, 2003).

As a first step, comprehensive secondary research of existing literature and primary research by means of a qualitative content analysis and a qualitative expert interview were conducted to generate the pool of items necessary for developing the six-dimensional measurement scale. The second part of the research investigated users' perceived values of location based services by conducting a quantitative survey.

With regard to sampling, the prior use of location based services constitutes a prerequisite for people participating in the survey. This is in line with the participation criteria of previous studies (Pura, 2005; Dickinger, 2006; Pihlström, 2008; Kumar & Lim, 2008) who argue that in the mobile service context, non-users do not know what to expect since they have never used these services and therefore are not able to validate them properly.

As a consequence, the target population is limited to LBS users only, which justifies the sampling method of a self-selection online survey. Respondents were asked to evaluate the pool of items of the six value dimensions as well as the variables attitude and behavioural intention. The items were measured by means of a five-point Likert-scale (Saunders et al., 2007) by asking for the extent of agreement, respectively disagreement, ranging from 1 (strongly agree) to 5 (strongly disagree). The online questionnaire was started 275 times, whereby an initial screening question ensured that only LBS users continued the survey, by means of which 101 people were filtered out. After data cleaning, a total number of 122 questionnaires could be retained for further analysis.

4 Findings

4.1 The Development of the Multi-Dimensional Measurement Scale

A step-by-step process was undertaken to develop the multi-dimensional scale for the present study. First of all, existing value scales and associated items were obtained by drawing from secondary literature research of mobile services and perceived value. Thereby, suitable items were adopted one-to-one, whilst others were slightly adapted in terms of wording in order to be more applicable to the context of touristic location based services. Items could be adopted from following studies (Sweeney & Soutar, 2001; Pura, 2005; Yu et al., 2005; Dickinger, 2006; Yang & Jolly, 2009; Kumar & Lim, 2008; Ruiz-Mafé et al., 2010).

Secondly, items were generated by means of a qualitative content analysis which proved to be crucial, as items from literature could be either partially confirmed or deemed irrelevant in the context of LBS. Thereby, frequently mentioned value aspects were incorporated in the corresponding value dimension of the construct. As a third step, the pool of items was judged by an expert interview to evaluate the semantic content of the items derived from both, literature and content analysis, which was crucial at this stage to ensure the validity of the scale. Eventually, a measurement scale with a total of 44 items divided into six value dimensions and two variables, *attitude* and *behavioural intention*, could be developed.

Previous studies (Pura, 2005; Dickinger, 2006; Yang & Jolly, 2008) suggest a separate analysis of the measurement model and the structural model in order to get a more comprehensive understanding of each single construct. On that account, firstly a factor analysis was performed to measure the proposed model in terms of items and corresponding dimensions. Secondly, a regression analysis was conducted to measure the hypothesised relationships between all perceived value dimensions and the dependent variables *attitude* respectively *behavioural intention*.

4.2 Results of the Factor Analysis

In order to test the factorability of the variable sets, a few indicators had to be considered. Kaiser-Meyer-Olkin Measure of Sampling Adequacy (MSA) was used, which indicated a value of 0.848 and hence exceeded the threshold of 0.8, being considered as meritorious (Backhaus et al., 2006). Bartlett's Test of Sphericity was significant 0.000. Communalities range between 0.691 and 0.880, whereby all communalities are higher than 0.40, which is considered good in social sciences. The factors show a high variance explained ranging from 3.570 % to 33.157 % with a total variance explained of 77.243 %. As a result, the adequacy of data for conducting a factor analysis can be assumed (Tabachnik & Fidell, 2001).

According to literature (Fabrigar et al., 1999), there exist a number of extraction procedures for the conduction of a factor analysis. In this particular study, a Principal Component Analysis (PCA) was applied.

Although there is disagreement in literature (e.g. Fabrigar et al., 1999; Costello & Osborne, 2005) about PCA being a true factor analysis method, since it constitutes a data reduction method by reducing variables to principal components accounting for

the maximum of the variance of observed variables (Suhr, 2005), it nevertheless is one of the most applied techniques in practice.

In terms of rotation, the most commonly applied method constitutes the orthogonal Varimax rotation, although literature strongly favours oblique rotation as it yields slightly better results (Costello & Osborne, 2005). The present study came to the conclusion to adopt a PCA with the oblique rotation method Promax (Kappa 4).

In terms of number of extracted factors, broad consensus suggests the extraction of factors with an Eigenvalue greater than 1.0 (Backhaus et al., 2006), whereby the number of factors revealing the clearest loading picture shall be retained (Costello & Osborne, 2005). Literature additionally proposes a scree plot test, which allows the researcher to manually search for the appropriate number of factors based on a graph of Eigenvalues, which reaffirmed eight extracted factors in this particular study.

The results of this study indicate that the same six value dimensions (*functional, social, emotional, epistemic, conditional, monetary*) as initially proposed in the research model (Sheth et al., 1991a/b; Pihlström, 2008) could be extracted from the data. In addition to that, two new dimensions could be extracted loading on separate constructs, which were named *convenience value* and *informational value*. The Pattern Matrix obtained by the PCA is depicted below in Table 1.

Table 1. Final Measurement Model

Nr. / Factor	Indicator	Load	V.E.	α
1 Functional	LBS are usually consistent and reliable	.970	11.237 %	.788
	LBS usually have good functions	.698		
	LBS usually fulfil my needs well	.807		
2 Convenience	LBS are usually easy to use	.886	3.570 %	.700
	LBS usually provide the service in a timely manner	.797		
3 Social	Using LBS helps me to feel accepted by others	.951	33.157 %	.895
	Using LBS improves the way I am perceived by others	.919		
	Using LBS makes a good impression on others	.667		
	Using LBS gives me social approval	.920		
4 Emotional	Using LBS is fun	.569	6.029 %	.867
	Using LBS is enjoyable	.844		
	Using LBS is interesting	.654		

Cont. Table 1

	Using LBS is a good pastime	.844		
5 Epistemic	I use LBS to experiment with new ways of doing things	.844	7.218 %	.819
	I use LBS to test new technologies	.796		
	I use LBS out of curiosity	.819		
6 Monetary	I value if LBS provide good service for the price	.942	3.687 %	.695
	LBS are usually reasonably priced	.866		
7 Conditional	I value the independence of place and time offered by using LBS	.647	7.782 %	.824
	I value the knowledge about places I get by using LBS	.652		
	I value the feeling that by using LBS I can orient myself even in unfamiliar environments	.675		
	I value the permanent availability of information by using LBS	.821		
8 Informational	I value the real time information that LBS offer	.772	4.563 %	.773
	I value the up-to-date information that LBS offer	.877		
Attitude	It is a good idea to use LBS			.905
	I am favourable about using LBS			
	I think using LBS is a positive experience			
Behavioural Intention	Given the chance, I intend to use LBS			.834
	I expect my use of LBS to continue in the future			
	I will use LBS more frequently in the future			
	I will recommend others to use LBS			
	Next time I need travel information, I intend to use LBS for this purpose			
Total Variance Explained			77.243 %	
Kaiser-Meyer-Olkin Measure of Sampling Adequacy			.848	
Bartlett's Test of Sphericity	Approx. Chi-Square		1680.493	
	df		276	
	Sig.		.000	

4.3 Results of the Regression Analysis

A regression analysis was performed to test the 13 initially hypothesised relationships as well as the four additional relationships between the newly developed dimensions (*informational value* and *convenience value*) and the dependent variables (Table 2). The coefficients indicate that all of the total 17 hypotheses, except one, could be supported. Whilst all values prove to have a significant positive correlation with behavioural intention, convenience value does not show a statistically important relationship but tends to, by slightly exceeding the significance level (ß=.176, p-value .053). All other perceived value dimensions namely *functional*, *social*, *emotional*, *epistemic* and *conditional value* as well as the newly extracted *informational value* have a positive correlation with the mediating variable *attitude* as well as with the outcome variable *behavioural intention*. Thereby, it has become evident that functional value has the strongest influence on attitude (ß=.658***), which was

followed by emotional value (ß=.587***) and conditional value (ß=.504***).The relationship between values and attitude is well established in literature and could also be confirmed in this particular study, by supporting all proposed hypotheses (H1., H2., H3., H4., H5., H6.). In terms of the relationship between values and behavioural intention, conditional value appeared to be the strongest indicator (ß=.538***), followed by epistemic value (ß=.522***) and emotional value (ß=.515***). All initially proposed hypotheses (H7., H8., H9., H10., H11., H12.) could be supported. The new dimension informational value shows a comparably weak relationship with attitude (ß=.348***), respectively behavioural intention (ß=.320***), likewise does convenience value with a weak effect on attitude (ß=.255**) and an insignificant effect on behavioural intention (ß=.176). The latter results can presumably be attributed to the fact that these value dimensions were not tested as a full construct but rather constitute split-offs from other value dimensions. However, the newly developed values can be considered for testing in future research. The relationship between attitude and behavioural intention showed a particularly strong effect, with a coefficient of (ß=.742***), which hence leads to the verification of H13. Table 2 depicts the results of the hypotheses test by outlining the correlation coefficient (Beta Coefficient), significance, t-value, R^2 and the indication of whether the hypotheses are supported or not.

Table 2. Hypotheses Test

	Hypotheses	**ß**	**t-value**	**Sign.**	**R^2**	**Supported**
H1.	Functional value -> Attitude	.658***	9.564	.000	.433	Yes
H2.	Social value -> Attitude	.322***	3.722	.000	.103	Yes
H3.	Emotional value -> Attitude	.587***	7.939	.000	.344	Yes
H4.	Epistemic value -> Attitude	.453***	5.559	.000	.205	Yes
H5.	Conditional value -> Attitude	.504***	6.389	.000	.254	Yes
H6.	Monetary value -> Attitude	.438***	5.333	.000	.192	Yes
new	Convenience value -> Attitude	.255**	2.885	.005	.065	(Yes)
new	Informational value -> Attitude	.348***	4.063	.000	.121	(Yes)
H7.	Functional value -> Behavioural Intention	.447***	5.478	.000	.200	Yes
H8.	Social value -> Behavioural Intention	.430***	5.223	.000	.185	Yes
H9.	Emotional value -> Behavioural Intention	.515***	6.583	.000	.265	Yes
H10.	Epistemic value -> Behavioural Intention	.522***	6.702	.000	.272	Yes
H11.	Conditional -> Behavioural Intention	.538***	6.984	.000	.289	Yes
H12.	Monetary value -> Behavioural Intention	.460***	5.681	.000	.212	Yes
new	Convenience Value-> Behavioural Intention	.176	1.954	.053	.031	(No)
new	Informational Value -> Behavioural Intention	.320***	3.695	.000	.102	(Yes)
H13.	Attitude -> Behavioural Intention	.742***	12.111	.000	.550	Yes

5 Discussion

The findings gathered by the factor analysis did not only confirm the initially proposed six-dimensional scale but also revealed two additional factors, which led to the suggestion of an extended model by proposing two new dimensions labelled informational value and convenience value. The conceptual idea of the latter value is consistent with findings by Pura (2005), who, in the context of electronic services split Sheth et al.'s (1991a) original functional dimension into convenience and monetary value, by acknowledging the importance of convenience in the field of mobile service use. The items of the informational value dimension were predominantly drawn from the qualitative user content analysis, indicating a strong need for multiple types and sources of information when using LBS. Due to the fact that these items loaded on a distinct factor and were not associable to the conditional value dimension, as proposed earlier, the actual need for an informational value factor when dealing with LBS has evidently been confirmed.

The results gathered from the regression analysis of the structural model demonstrate that, apart from one, all hypothesised paths between the value dimensions and the dependent variables, could be supported. This leads to the argumentation that all values positively affect the user's attitude and behavioural intention to use touristic location based services. However, the findings also reveal that the effects of value dimensions on the former and latter variable do not necessarily correspond. As a matter of fact, it is evident that functional value of LBS, for instance, constitutes the strongest determinant of attitude towards using LBS, whereas conditional value is the strongest predictor of behavioural intention. In this sense, results indicate that in terms of attitude, functional value could be revealed to have the strongest correlation, followed by emotional value and conditional value. With regard to behavioural intention, conditional value proved to be the strongest indicator, followed by epistemic value and emotional value.

6 Conclusion

Location based services are predicted to find wide application within the tourism industry in the coming years. However, to date, LBS have not found yet the foretold adoption by the mass market, arguably due to generally missing understanding of the real values inherent in location based services. As a result, the aim of the present study was to investigate the perceived value of touristic location based services from a user-centric perspective. For this purpose, the study adopted the framework of perceived value rather than traditional technology-oriented theories, which have been numerously applied in the past. This allow to gain a more comprehensive understanding in order to explore the full complexity of the underlying value perceptions of location based services. The study developed a six-dimensional research model for the particular context of location based services by conflating elements of the original models developed by Sheth (1991a/b) and Pihlström (2008).

The adoption of a sequential exploratory strategy by implementing qualitative and quantitative methods in a sequential manner ensured the development of a valid measurement scale with dimensions and items appropriate for the context of LBS. The subsequent conduction of a quantitative survey had the aim to explore the value

perceptions of location based services. The statistical analysis by means of a factor analysis affirmed all six initially proposed value dimensions from literature and extended the original model by two new value dimensions, namely informational value and convenience value. A consequent regression analysis tested the structural model in terms of the hypothesised correlations between all values and the mediating, respectively outcome variable, the majority of which could be supported.

7 Implications and Future Research

The application of the framework of perceived value contributed to the understanding of technological innovations by applying the consumer-oriented perceived value construct rather than purely technology-adoption oriented models, as numerously used in past research. Thereby, the demand of current literature for adopting a multi-dimensional approach could be taken into account.

Acknowledging the limitations of this study in terms of the small sample size, the consequent limited generalisability as well as the inclusion of LBS users only; further research in the field of touristic LBS is needed to validate the preliminary results of this exploratory study. Moreover, the replication of the newly developed eight-dimensional construct in further studies, particularly in the context of mobile services, is highly encouraged. Although single items of the newly extracted dimensions need further improvement, there is great potential that this model could serve as a valid framework for further research in an information technology context.

Given the increasing proliferation of LBS, the results of the present study have crucial implications for various industry sectors. Given the high consumer expectations towards mobile services on the one end and the relatively low willingness to adopt these on the other end, the industry has to respond by both integrating and actively communicating the distinct values of these services to end-users. From a managerial standpoint, the findings provide critical knowledge in terms of the real values of using location based services. Industry players such as marketers, developers or service providers can profit from the gathered findings by implementing the identified values in the programming process or by communicating specific values more effectively. Moreover, the knowledge about which values particularly affect the user's attitude, respectively the intention to use LBS entails essential implications for practice. Finally, given the assumption that non-users do not know what to expect from LBS, the revealed values can be emphasised in the encouragement of non-users to adopt location based services in the future.

References

Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (2006). *Multivariate Analysemethoden: Eine anwendungsorientierte Einführung.* 11th edn. Berlin: Springer.

Balasubramanian, S., Peterson, R.A. & Jarvenpaa, S.L. (2002). Exploring the implications of m-commerce for markets and marketing. *Journal of the Academy of Marketing Science* 30(4): 348-361.

Bettman, J., Luce, M. & Payne, J. (1998). Constructive consumer choice processes. *Journal of Consumer Reserach* 25(3): 87-96.

Bigné, E., Sanz, S., Ruiz, C. & Joaquín, A. (2010). Why Some Internet Users Don't Buy Air Tickets Online. In U. Gretzel, R. Law & M. Fuchs (Eds.), *Information and communication technologies in tourism*. Wien, Springer.

Blanford, R. (2010). *comScore data shows smartphone situation in Europe*. [Online]. Retrieved August 13, 2010, from http://www.allaboutsymbian.com/news/item/11342_comScore_data_shows_smartphone.php

Costello, A.B. & Osborne, J.W. (2005). Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From Your Analysis. *Practical Assessment, Research & Evaluation* 10(7): 1-9.

Creswell, J.W. (2003). *Research Design: Quantitative, Qualitative, and Mixed Methods Approaches.* Thousand Oaks: Sage Publications, Inc.

Davis, F.D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly* 13(3): 318-340.

Davis, F.D., Bagozzi, R. & Warshaw, P. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science* 35(8): 982-1003.

DeSarbo, W., Jedidi, K. & Sinha, I. (2001). Customer value analysis in a heterogeneous market. *Strategic Management Journal* 22: 845-857.

Dickinger, A. (2006). *Mobile Service Usage: Antecedents and Behavioral Consequences of Perceived Quality in User Segments.* Vienna: Vienna University of Economics and Business Administration, Institute for Tourism and Leisure Studies.

Egger, R. & Jooss, M. (2010). Die Zukunft im mTourism: Ausblick auf Technologie- und Dienstentwicklung, In R. Egger & M. Jooss (Eds.), *mTourism: Mobile Dienste im Tourismus.* Wiesbaden, Gabler Verlag.

Fabrigar, L.R., Wegener, D.T., MacCallum, R.C. & Strahan, E.J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4: 272-299.

Göll, N., Lassnig, M. & Rehrl, K. (2010). Location-Based Services im mTourismus: Quo Vadis? In R. Egger & M. Jooss (Eds.), *mTourism Mobile Dienste im Tourismus*. Wiesbaden, Gabler Verlag.

Kim, H., Chan, C.H. & Gupta, S. (2005). Value-based adoption of mobile internet: an empirical investigation. *Decision Support Systems* 43: 111-126.

Kleijnen, M., Wetzels, M. & De Ruyter, K. (2004). Consumer Acceptance of Wireless Finance. *Journal of Financial Services Marketing* 8(3): 206-218.

Komulainen, H., Mainela, T., Tähtinen, J. & Ulkuniemi, P. (2004). Exploring customer perceived value in a technology intensive service innovation. In *Proceedings of the 20th IMP Conference in Copenhagen*, Denmark.

Kumar, A. & Lim, H. (2008). Age differences in mobile service perceptions: comparison of Generation Y and baby boomers. *Journal of Services Marketing* 22(7): 568-577.

Lubbe, B.A. & Louw, L. (2010). The perceived value of mobile devices to passengers across the airline travel activity chain. *Journal of Air Transport Management* 16: 12-15.

Oh, H. (1999). Service quality, customer satisfaction and customer value: A holistic perspective. *International Journal of Hospitality Management* 18: 67-82.

Pedersen, P.E. (2005). Adoption of mobile internet services: an exploratory study of mobile commerce early adopters. *Journal of Organizational Computing and Electronic Commerce* 15(3): 203.

Pihlström, M. (2008). *Perceived value of mobile service use and its consequences.* Helsinki: Edita Prima Ltd.

Pura, M. (2005). Linking perceived value and loyalty in location-based mobile services. *Managing Service Quality* 15(6): 509-539.

Rasinger, J., Fuchs, M. & Höpken, W. (2007). Information search with mobile tourist guides: a survey of usage intention. *Information Technology & Tourism* 9(3/4): 177-194.

Rogers, E.M. (2003). *Diffusion of innovations*. 5th edn. New York: Free Press.

RTR (2010). *RTR Telecom Monitor*. [Online]. Retrieved August 13, 2010, from http://www.rtr.at/en/komp/TKMonitor_2_2010/TM2-2010-e.pdf

Ruiz-Mafé, C., Sanz-Blas, S. & Tavera-Mesías, J.F. (2010). A comparative study of mobile messaging services acceptance to participate in television programmes. *Journal of Service Management* 21(1): 69-102.

Saunders, M., Lewis, P. & Thornhill, A. (2007). *Research Methods for Business Students.* 4th edn. Essex: Pearson Education Limited.

Sheth, J.N., Newman, B.I. & Gross, B.L. (1991a). *Consumption values and market choices, theory and applications.* Cincinnati: South Western Publishing Co.

Sheth, J.N., Newman, B.I. & Gross, B.L. (1991b). Why We Buy What We Buy: A Theory of Consumption Values. *Journal of Business Research* 22: 159-170.

Statista (2010). *Warum nutzen Sie das mobile Internet nicht?* [Online]. Retrieved September 19, 2010, from http://de.statista.com/statistik/daten/studie/156723/umfrage/gruende-fuer-die-nichtnutzung-des-mobilen-internets/

Suhr, D.D. (2005). Principal Component Analysis vs. Factor Analysis. In *SAS SUGI 30 Proceedings*, *Statistics and Data Analysis Section.* Cary, NC, SAS Institute Inc.

Sweeney, J. C. & Soutar, G. N. (2001). Consumer Perceived Value: The Development of a Multiple Item Scale. *Journal of Retailing* 77(2): 203-220.

Tabachnick, B. G. & Fidell, L. S. (2001). *Using Multivariate Statistics.* 4th edn. Boston: Allyn and Bacon.

Venkatesh, V., Morris, M.G., Davis, G.B., & Davis, F.D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly* 27(3): 425-478.

Wang, Y., Lo, H.P., Chi, R. & Yang, Y. (2004). An integrated framework for customer value and customer-relationship-management performance: a customer-based perspective from China. *Managing Service Quality* 14(2/3): 169-182.

Wittmer, A., Riklin, T. & Beritelli, P. (2007). *Do Location Based Services in Tourism Really get Assessed?: A Critical Analysis of Adopted Evaluation Methods and Techniques in Research Projects.* International ENTER Conference. [Online]. Retrieved August 15, 2010, from www.alexandria.unisg.ch/export/PDF/publications/Pietro_Beritelli.pdf

Yang, K. & Jolly, L.D. (2008). Age cohort analysis in adoption of mobile data services: gen Xers versus baby boomers. *Journal of Consumer Marketing* 25(5): 272-280.

Yu, J., Ha, I., Choi, M. & Rho, J. (2005). Extending the TAM for a t-commerce. *Information & Management* 42(7): 965-976.

Zeithaml, V.A. (1988). Consumer Perceptions of Price, Quality and Value: A Means-End Model and Synthesis of Evidence. *Journal of Marketing 52* (July): 2-22.

Acknowledgements

The author likes to gratefully acknowledge the support provided by the University of Derby and especially thank Dr. Tim Heap for his supervision.

A Typology of Travellers based on their Propensity to go online before, during and after the Trip

Anne-Mette Hjalager[a], and
Jan Møller Jensen[b]

[a] Department of Business and Environmental Economy
University of Southern Denmark
hjalager@sam.sdu.dk

[b] Department of Marketing & Managemetn
University of Southern Denmark
jmj@sam.sdu.dk

Abstract

This article develops a typology of tourists based on the self-reported importance of Internet access before, during and after their trip. Survey results suggest the prevalence of five distinct tourist segments: Offliners (marginal Internet users), online planners (mainly going online before the trip), online explorers (who access online sources continuously before and during the trip), online keepsakers (who collect information and share before and after the trip), and online equilibrists (active online in all travel phases). Further, results support the importance of the internet in searching for general travel information. A learning curve is evident, as the most experienced travellers are also the most devoted online information searchers. The study implies that future Internet travel applications may more efficient if planned to transcend the borders between promotion, marketing, destination services and loyalty content.

Key words: Traveller typology. Online information search. Socio-demography.

1 Introduction

Over the past two decades, the Internet has facilitated profound changes in society, with tourism as a phenomenon, business and recreational activity being widely affected (Buhalis & Law, 2008). Travel related information can be increasingly retrieved on the Internet (Pan et al, 2011; Xiang & Gretzel, 2010), and Internet searches have become the dominant information resource in connection with travel planning (Grønflaten, 2009). Understanding to what extent and how tourists use the Internet before, during and after their trips is essential for the design of effective tourism information systems and for the implementation of successful marketing campaigns (Parra-López et al., 2011; Xiang et al., 2008). This study seeks to provide an improved insight into the importance to travellers of Internet access for acquiring and sharing of travel related information before, during and after their trip. It provides a consistent and explanatory typology of travellers.

2 Theory

Providers of tourism services, the tourism business and destination marketers spend considerable effort to invent and develop appealing and functional interfaces with tourists in order to sustain market shares and gain access to new market segments (Buhalis & Law, 2008; Pan et al, 2011). There has been quite some focus on mapping and comprehension of customers' information search (Jensen, 2011) prior to holidays

in the research (Fodness & Murray, 1997; Fodness & Murray, 1999), and the derived results are, of course, of immense value for any agency responsible for tourism marketing, from single enterprises to destinations and national entities, for whom an appropriate mix of channels and communication forms is dependent upon updated understanding of search behaviour and preferences. While the Internet rapidly became a source of information *prior* to travel, present developments indicate that this has also been extended to use *during* holidays. In recognition of the importance of the "word of mouth" to the reputation of a tourism facility or destination, tourism literature also addresses what happens *after* the holiday. DiPietro el al (2007) for example, demonstrates that the recommendations from others are decisive, and particularly backpack travellers seem to rely on recommendations from other travellers, and see others as an audience for relived experiences (Hanlan & Kelly, 2005). Showing photographs and videos, and talking about holiday experiences is a common and classic social pastime, and in this, the Internet plays an increasingly important role. Blogs, photo and video sharing facilities etc. are places for storing and sharing memories of holidays and good times (Tussyadiah & Fesenmaier, 2009, Bronner & De Hoog, 2011). Sites such as, for example, Tripadvisor provide facilities for compiling a systematic rating of facilities following a visit as a counterbalance the commercial marketing.

The research tends to address the actions of tourists' in searching and disseminating information related to their tourism activities, either before, during or after the holiday, while there is very limited consistent evidence that encompasses the entire cycle of activities during a trip from the first considerations to the dissemination of subsequent information in its aftermath (DiPietro et al, 2007). Likewise, despite a growing focus on travellers' use of the Internet for travel planning, little research has investigated how important online access is perceived to be by travellers during and after travel (Ip et al, 2010). Grønflaten (2009) has, in his comprehensive review of information sources, not found any studies that directly compare the information search strategies by the same respondents in different travel phases. This study takes up this major research challenge.

The nature and extent of tourism related information search (Fodness & Murray, 1999), and most specifically online search, depends upon a range of external conditions and on the situation and characteristics of the individual. Numerous studies identify socio-demographic variation, such as generation (Jones & Fox, 2009), education, income (Ip et al, 2010), where the younger, more affluent and better educated segments are consistently found to be first-movers, although the Internet has caught on in wider segments more rapidly than expected by many (OECD, 2008).

The effect of product expertise, or knowledge about the amount of information acquired during a purchase process, and the consideration that the consumer gave the decision process has received considerable attention in the literature (Alba & Hutchinson, 1987; Selnes & Troye, 1989). However, the relationship is unclear. Researchers supporting a positive relationship between prior experience and information search argue that consumers with greater product experience find and process supplementary information more easily, thereby reducing the average cost of their search. While others who claim a negative relationship explain their findings by the fact that persons with prior knowledge need less additional information. Finally,

researchers supporting an inverted-U relationship argue that consumers with limited experience will be less motivated to search for and process information compared with moderately experienced consumers who, in turn, experience a greater need for information search than those with a high degree of prior purchase experience (Brucks, 1985).

The sociology of tourism also reveals a clear personal learning curve. The intensity and direction of information search prior to travel depends not only on the travel experience, but also on familiarity with the specific destination (Fodness & Murray, 1997; Hyde, 2006; Tan & Cheng, 2011). In addition, the more experienced the travellers, the more likely they are to adopt adventurous behaviour (Alvarez & Asugment, 2006; Dey & Sarma, 2010), which in turn accounts for information search preferences and activities (Fodness & Murray, 1999). Studies of online travel search suggest that there are distinct groups of frontrunners and later adapters of new technological opportunities (Ryan & Rao, 2008). Someone is beating the ICT track for others to follow. Studies of advanced tourists, who are geared to be online throughout all of the phases of a trip, might provide predictions for the adoption of these activities by a larger majority of the tourism population in due course.

3 Research Questions and Methodology

3.1. Research Questions

Combining information search before, during and after trips is crucial to move theory along in the field of tourist's information search reflecting on the knowledge gaps presented above, three sets of research questions were formulated:

- Q1a. To what extent do tourists perceive online access as important before, during and after their trip?
- Q1b. Is it possible to identify tourist segments which reflect the self-reported importance of online access before, during and after a trip?
- Q1c. How do these segments differ in terms of socio-demographic profiles?
- Q2a. Can a learning curve be deduced and understood as a relationship between travel experience and the importance of online access?
- Q2b. Can a learning curve be deduced and understood as a relationship between online experience and the importance of online access?
- Q3a. What purposes do tourists use the Internet for before, during and after their trip?
- Q3b. How do segments differ with respect to their online activities before, during and after their trip?

3.2 Measurement Instrument and Data Collection

A web-based questionnaire was designed to collect information about tourists' perceptions of the importance of online access and their actual use of the Internet before, during and after a trip. In the first section, respondents were asked to provide information on various socio-demographic variables including gender, age, education and income. The second section was concerned with respondent's online experience and included questions about how respondents perceived their own online capabilities and how frequently they were online. The third section was concerned with

respondents travel experience and asked respondents to indicate how frequently they travel and how many trips (excluding business travel) they have made during the last three years. In the last section, respondents were asked to rate the perceived importance of having access to the Internet before, during and after their trip and to indicate how many trips they had made in the period, and whether they had made use of the Internet before, during and after these trips and what purposes.

The sampling frame comprised a representative panel of Danish online users owned by a private marketing research institute, and made available for this study. A total of 532 responses were received, of which 515 were adequately completed. The gender of respondents was evenly distributed (males 49.9% and females 50.1%). Six out of ten (58%) were aged from 40 to 60, 19.7% were 39 years or younger, and only 12.2% were 60 years or older. Three out of four (73.8%) were married or co-habiting and about one forth (26.2%) were singles. Nearly half (45.4%) of the respondents had a higher education of three years or longer, while only a few had a high school degree. About two-thirds (64.1%) had an annual household income of DKK 500,000 or more. Reflecting the sampling method, 94% report being online seven days a week.

3.3. Constructing a Typology

An analysis of the overall sample shows that 80% of the 515 respondents reports online access as important or very important in the pre-trip planning period. This clearly indicates the vital role played by the Internet in travel planning (Q1a). Although occurring in relatively lower proportions (20% and 29%, respectively) the results further reveal that for many travellers, the Internet is also an important source for seeking and sharing travel related information during and after the trip.

Having assessed the importance of online access before, during and after the trip, the next step was to determine whether it was possible to identify distinct groups of tourists, in terms of their perceived importance of access to the Internet before, during and after their vacation (Q1b). To achieve this aim a hierarchical cluster analysis was conducted using the three variables; importance of access to the Internet before, during and after the trip. Distances between the clusters were calculated with the squared Euclidean distance measure, and the aggregation of clusters was performed with Ward's procedure to minimize the within-cluster differences and to maximize the between-cluster differences. Increases in the agglomeration coefficients (indicating the dissimilarity between the clusters to be merged at each clustering stage) were inspected in order to assess the number of clusters. Based on the agglomeration coefficients and interpretation of cluster profiles a 5-cluster solution was chosen. The mean of the three cluster variables, for each cluster are shown in Table 1 along with the results of an ANOVA test. As expected, significant differences were found between clusters across the three clustering variables.

- Cluster 1 "offliners" includes 15.4% of the respondents, and was labelled offliners since this group rated the importance of online access low in all three phases of the trip process.
- Cluster 2 "online planners", representing 25.1% of the respondents, placed high importance of online access before the trip while online access in the other two phases was rated relatively low.

- Clusters 3 “online explorers” represents 26.9% of the respondents with high mean scores for online access before and during the trip, and low mean score for online access after the trip.
- Cluster 4, “online keepsakers” includes 14.6% of the respondents with high mean scores for online access before and after the trip and low mean score for online access during the trip.
- Cluster 5, “online equilibrists” represents the remaining 17.9% of the respondents who rate the importance of online access highly in all three phases of the trip.

Table 1. Travellers’ rated importance of access to the Internet before, during and after the trip. Overall mean and cluster means

	Total n=515	Off-liners n=81	Online plan-ners n=129	Online explo-rers n=138	Online keep-sakers n=75	Online equili-brists n=92	F Value	Sig.
Cluster size	100%	15.4%	25.1%	26.9%	14.6%	17.9%		
Importance of online access before the trip	4.19	2.40[a]	4.53[b]	4.41[b]	4.57[b]	4.65[b]	192.167	.000
Importance of online access during the trip	2.53	1.69[a]	1.60[a]	3.51[b]	1.73[a]	3.76[b]	293.588	.000
Importance of online access after the trip	2.61	1.65[a]	1.64[a]	2.15[b]	3.99[c]	4.39[d]	364.127	.000

Note: Mean values with the same superscript are not significantly ($p < 0.05$) different from one another, based on an ANOVA post-hoc test.

The constructed typologies are found to be appropriate for the purpose of this article, thus indicating that different online search strategies exist in parallel.

4 Socio-Demographic Profile of the Typology

In this section cross-tabulations and the chi-square statistic are employed to determine statistical differences between the clusters with respect to other variables included in the research questions (Q1c). Results are displayed in Table 2.

The chi-square tests clearly confirm that gender, age, household composition, education and household incomes are significantly different among the five clusters.

While offline travellers and online planners are equally distributed amongst males and females, males seem to dominate the segment of explorers who are online more intensively before and during the trip as opposed to the two segments of online keepsakers and online equilibrists; the latter groups emphasise sharing travel information after the trip, which segments are dominated by females.

Table 2: Cluster profiles across socio-demographic variables

Socio-demographics	Total	Off-liners n=81	Online plan-ners n=129	Online explo-rers n=138	Online keep-sakers n=75	Online equili-brists n=92	Chi-square (df)	Sig.
Gender							12.577 (4)	.014
Male	49.9	51.9	51.9	59.4	38.7	40.2		
Female	50.1	48.1	48.1	40.6	61.3	59.8		
Household composition							11.841 (4)	.019
Single	26.2	31.2	22.7	17.5	34.7	32.6		
Couple	73.8	68.8	77.3	82.5	65.3	67.4		
Age							64.454(16)	.001
18-29	13.2	7.4	9.3	6.5	30.7	19.6		
30-39	16.5	12.3	11.6	15.9	22.7	22.8		
40-49	28.7	19.8	31.8	36.2	25.3	23.9		
50-59	29.3	35.8	32.6	31.2	17.3	26.1		
60+	12.2	24.7	14.7	10.1	4.0	7.6		
Education							28.792(12)	.004
Primary school	7.2	11.1	8.5	6.5	4.0	5.4		
High school	47.4	55.6	49.6	38.4	52.0	46.7		
Higher: 3-4 years	29.3	23.5	20.9	34.8	25.3	41.3		
Higher: 5+ years	16.1	9.9	20.9	20.3	18.7	6.5		
Household Income DKK							35.796(16)	.003
< 300 000	15.5	19.1	17.7	7.6	20.7	17.1		
300.000-499.999	20.6	25.0	22.1	14.4	20.7	23.7		
500.000-699.999	24.2	33.8	22.1	24.6	19.0	22.4		
700.000-899.999	23.3	19.1	24.8	31.4	24.1	11.8		
900.000+	16.4	2.9	13.3	22.0	15.5	25.0		

Similarly to previous studies on travellers' use of the Internet (for example Kim et al, 2011), the group of off-line travellers has an overrepresentation of travellers in the age groups of 50 years or higher (60.5%), whereas the online traveller segment included greater numbers from the age groups below 50 (66.3%). Online keepsakers and equilibrists are generally young. Online planners and online explorers include a higher proportion of couples than the other groups. Not surprisingly, the group of offliners is also, on average, characterized by shorter educations and lower incomes. They are contrasted particularly by the higher educated and more affluent online explorers. The equilibrists are an interesting group from a socio-demographic

perspective; they are young, but also high average income earners. They have both the means and knowledge to be front-runners in terms of access online, both before, during and after the trip. It is also worthwhile to note the domination of females amongst online keepsakers, who are generally young and not (yet) high income earners. This segment seems to correspond to the lifestyle and information and communication habits of interpersonal relations of emerging adults (Subrahmanyam et al, 2008).

5 Travel and Internet Experience

Research question 2 concerns the relationship between online experience and travel experience on the one hand and the perceived importance of online access on the other hand. To examine this relationship a series of ANOVA tests, coupled with post hoc multiple comparisons, were conducted to compare online experience and travel experience across the five clusters.

In Table 3, the results in the upper section show significant differences across the five clusters with respect to online experience, here measured by respondents' perception of own online capabilities and estimated minutes of online use pr. day. Offliners perceive their own online capabilities to be significantly lower compared to all other groups. Travellers with low online confidence may find it more costly (time and effort) and risky to use online facilities and therefore prefer offline services. In contrast, online equilibrists, rating their own online capabilities significantly higher than all other groups, may perceive less cost and more benefits from using online services in all phases of their travel behaviour. Similar patterns are found with respect to how many minutes pr. day cluster groups are online, although only mean scores for online keepsakers and equilibrists are significantly higher when compared with other groups.

The lower part of Table 3 shows that the five clusters differ significantly, and in a similar pattern, with respect the respondents' travel frequency and number of trips during the past three years. The offliners are clearly less frequent travellers than the explorers and the equilibrists. Online explorers and equilibrists travel significantly more than all other groups. Accordingly, those who travel intensively also find the access to the Internet more important in all phases of the trip. These results are consistent with Jensen (2009, 2011).

This finding supports the positive learning curve thesis, and it stands in contrast to Vogt et al (1998), who more than a decade ago found that skilled and experienced travellers are far more likely than other tourists to rely on their own personally compiled travel files for information about destinations they have not visited before. Thus, the prevalence of the Internet both in daily routines and during holidays tends to have changed the value of what is previously learned, and where online access is available. This knowledge seems to be utilized and increasingly considered a standard prerequisite. This study can also clearly demonstrate a spill-over from work and leisure routines resulting in expectations to online access before, during and after a trip.

Table 3. Cluster differences in relation to travel experience and online experience

	Total	Off-liners n=81	Online plan-ners n=129	Online explo-rers n=138	Online keep-sakers n=75	Online equili-brists n=92	F Value	Sig.
Online experience								
Perception of own capabilities online[2)]	3.15	2.98[a]	3.40[b]	3.63[b]	3.73[b]	3.98[c]	15.229	.000
Online minutes pr. day [3)]	6.12	5.57[a]	5.60[a]	5.90[a]	6.75[b]	7.14[b]	11.720	.000
Travel experience								
Travel frequency[1)]	1.56	1.17[A]	1.55 [B,C]	1.87[C]	1.25[A,B]	1.71[C]	12.962	.000
Number of trips during the last three years	4.62	3.00[a]	4.40[b,c]	5.62[c]	3.89[a,b]	5.45[c]	12.703	.000

Note: Mean values with the same superscript are not significantly ($p < 0.05$) different from one another, based on an ANOVA post-hoc test. Capital letters based on Scheffe and small letters based a Dunnets T3 test. Scales: [1)] Scale: 0=never, 0.5=every other year or less, 1=about once a year, 2=approximately 2 times per year, and 3= 3 or more times a year. [2)] Scale: 1=I am definitely a new beginner, 2= I am at a low level, 3=I am at a medium level, 4=I am at a high level, 5= I am a super user. [3)] Scale: 1=never use the Internet, 2=1-10 minutes, 3=11-20 minutes, 4= 21-30 minutes, 5=31-40 minutes, 6=41-50 minutes, 7= 51-60 minutes, and 8=more than an hour a day

Table 3 particularly leads the attention to the online planners who tend to be very rational in their general online behaviour, and for them the social dimensions and interrelations play a minor role. Quite the reverse is seen in the groups of online keepsakers and equilibrists, for whom staying in touch with friends and family is of crucial importance in daily life. During a holiday, keepsakers change their behaviour, and online access becomes less important for them. The explanation for this might be related to the steep roaming costs of Internet access particularly when travelling abroad (O'Reagan (2008).

6 Online Activities Before, During and After the Trip

With data from the survey it is possible to provide a more detailed picture of the online activities and purposes of being online before, during and after a trip (Q3a and Q3b). A series of ANOVA tests, coupled with post hoc multiple comparisons, were conducted to compare online activities before, during and after the trip across the five clusters.

The results in Table 4 show that pre-trip online activity level is significant higher among online planners, online explorers and equilibrists compared with that of off-liners and online keepsakers. These three groups are more likely to go online to seek inspiration for choice of destination, search for other types of information, visit travel agency homepages, use travel portals and to book their trip online.

Table 4. Average number of trips with the respective activities before the trip

	Total	Off-liners n=81	Online plan-ners n=129	Online explo-rers n=138	Online keep-sakers n=75	Online equilibris ts n=92	F Value	Sig.
Pre-trip online activities								
Inspiration for choice of destination	2.66	1.36^{a}	2.70^{b}	3.27^{c}	2.20^{b}	3.24^{c}	18.248	.000
Other information about the travel	3.05	1.78^{a}	3.13^{b}	3.72^{c}	2.59^{a}	3.46^{c}	20.441	.000
Booking	2.75	1.49^{a}	$2.87^{b,c}$	3.30^{c}	2.44^{b}	$3.12^{b,c}$	16.038	.000
Search travel agencies' websites	2.20	1.44^{A}	2.29^{B}	2.43^{B}	$1.92^{A,B}$	2.60^{B}	5.319	.000
Information on travel sites for example LonelyPlanet or Xpedia	.84	$.33^{A}$	$.61^{A,B}$	$1.04^{B,C}$	$.80^{A,B,C}$	1.34^{C}	6.580	.000
Information search on social media for example Twitter or Facebook	.16	$.05^{A}$	$.04^{A}$	$.10^{A}$	$.15^{A}$	$.54^{B}$	10.191	.000

Note: Mean values with the same superscript are not significantly ($p < 0.05$) different from one another, based on an ANOVA post-hoc test. Capital letters based on Scheffe and small letters based a Dunnets T3 test

As expected, the results in the upper part of Table 5 show that the online activities of online planners' are significantly lower during and post-trip. Online explorers and equilibrists more often bring their own computer and/or smartphone along and are more likely to use them to search information at the destination.

As observed in the lower part of Table 5, online keepsakers and equilibrists are significantly more likely to share pictures and inform others about their trip through Facebook and Twitter after returning home. Respondents rarely place travel reviews on travel portals, this does not differ significantly among the five clusters.

Table 5: Average number of trips with the respective activities during and after the trip

	Total n=515	Off-liners n=81	Online plan-ners n=129	Online explo-rers n=138	Online keepsak ers n=75	Online equili-brists n=92	F Value	Sig.
During trip online activities								
Carried a laptop on the trip	1.35	$.40^{A}$	$.41^{A}$	2.49^{B}	$.65^{A}$	2.38^{B}	48.937	.000
Carried a smartphone on the trip	3.02	1.93^{a}	2.84^{b}	3.62^{c}	$2.65^{a,b}$	3.61^{c}	13.806	.000
Was online during the trip	1.96	$.64^{A}$	1.09^{A}	3.13^{B}	1.17^{A}	3.20^{B}	58.717	.000
Was online on smartphone during the trip	.66	$.19^{A}$	$.29^{A}$	1.11^{B}	$.25^{A}$	1.25^{B}	15.338	.000
Uploaded pictures on social media during the trip	.27	$.06^{A}$	$.08^{A}$	$.21^{A}$	$.36^{A}$	$.76^{B}$	10.454	.000
Stayed in touch with SMS, MMS	2.71	1.41^{A}	2.60^{B}	3.30^{B}	2.67^{B}	3.16^{B}	15.117	.000
Post-trip online activities								
Uploaded pictures on social media after trip	.82	$.16^{A}$	$.33^{A}$	$.49^{A}$	1.36^{B}	2.13^{C}	36.750	.000
Rated the trip on travel portals, for example LonelyPlanet or Xpedia	.08	$.05^{a}$	$.06^{a}$	$.06^{a}$	$.07^{a}$	$.18^{a}$	1.407	.230

Note: Mean values with the same superscript are not significantly ($p < 0.05$) different from one another, based on an ANOVA post-hoc test. Capital letters based on Scheffe and small letters based a Dunnets T3 test

7 Discussion and Conclusion

This study confirms the importance of online search, particularly during the preparation phase of a holiday. It contributes with a new typology of search behaviour, demonstrates that behaviour varies across the five cluster categories and shows that there are systematic differences between the socio-demographic characteristics of these groups. Being online – particularly during and after the trip – is characteristic of the relatively more privileged groups who are travel frontrunners.

Taking the rapid adoption of Internet services into consideration, subsequent adoption by other groups can be expected with a fairly short delay, costs of online activities possibly being the modifying factor.

Online access tends to be so widespread that travellers increasingly enter a state of online "dependency", in the sense that they never finish searching for information and communicating it further. The learning curve seems to show perpetual progression among those who set the standards of behaviour. Contrarily, genuine "information relaxation" during a holiday seem to be a decreasing phenomenon, as continuous, self-imposed and always informed decision making seem to become the norm. Further, and not least qualitative, research and repeated quantitative studies are necessary to understand the nature and implications of this potential online dependency during travel.

This study contains issues of importance for tourism information and marketing. Evidence about tourist information-search preferences are valuable to tourism service providers and destinations who seek to combine an effective mix of media through which to reach target tourists. It is particularly remarkable that 65% of the respondents find online access important in more than one phase of their trips. This suggests a considerable potential for Internet access concepts that transcend the normal categories of travel service and destination services. This is not only a matter of simple web design strategies, but also of an innovative re-conceptualisation of travel information services (Guo & Wu, 2009) and space structuring (Molz, 2011) that efficiently and attractively combines destination promoting tourist information with communication, shopping and entertainment.

As with any study of this nature, there are some limitations that could be addressed in future studies. Respondents were recruited from an online panel which may have biased the proportion of travellers that find online access important. It would be of value to conduct similar studies on other populations with less online experience.

References

Alvarez, M. & Asugman, G. (2006). Explorers versus planners. *Annals of Tourism Research* 33(2): 319-338.

Bronner, F. & De Hoog, R. (2011). Vacationers and eVOM: Who posts, and why, where, and what? *Journal of Travel Research* 50(1): 15-26.

Brucks M. (1985). The Effect of Product Class Knowledge on Information Search Behavior. *Journal of Consumer Research* 12: 1-15.

Buhalis, D. & Law, R. (2008). Progress in tourism management: twenty years on and ten years after the internet: the state of eTourism research. *Tourism Managemen*, 29(4): 609-623.

Chiou, W.-B., Wan, C.-S. & Lee, H.-Y. (2008).Virtual experiences vs brochures in the advertisement of scenic spots: How cognitive preferences and order effects *in*fluence advertising effects on consumers. *Tourism Management* 29(1): 146–150.

Dey, B. & Sarma, MK. (2010). Information source usage among motive-based segments of travelers to newly emerging tourist destinations. *Tourism Management* 31: 341-344.

DiPietro, R.B., Wang, Y., Romph, R. & Sievert, D. (2007). At-destination visitor information search and venue decision strategies, *International Journal of Tourism Research* 9(3): 175-188.

Fodness, D. & Murray, B. (1997). Tourist information search. *Annals of Tourism Research 24*(3): 503-523.

Fodness, D. & Murray, B. (1999). A model of tourist information search behavior. *Journal of Travel Research 37*(3): 220-230.

Grønflaten, Ø. (2009). The tourist information matrix – differentiating between sources and channels in the assessment of travellers' information search. *Scandinavian Journal of Hospitality and Tourism* 9(1): 39-64.

Guo, T., J, S. & Wu, B. (2009). Location-based push service system of tourism information. *Journal of Computer Applications* 29(12): 3453-3457.

Hanlan, J. & Kelly, S. (2005). Image formation, information sources and an iconic Australian tourist destination, *Journal of Vacation Marketing* 11(2): 163–177.

Hyde, K. (2006). Contemporary information search strategies of destination-naïve international vacationers. *Journal of Travel & Tourism Marketing* 21(2/3): 63–76.

Ip, C., Lee, H. & Law, R. (2010). Profiling the users of travel websites for planning and online experience sharing. *Journal of Hospitality and Tourism Research*, in print

Jensen, J.M. (2009). Travellers' Intentions to Purchase Travel Products Online: The Role of Shopping Orientation. In M. Matias, P. Nijkamp, and M. Sarmento (Eds.), *Advances in Tourism Economics: New Developments*. Heidelberg, Physica-Verlag.

Jensen, J.M. (2011). Shopping orientation and online travel shopping. The role of the travel experience. *International Journal of Tourism Research*, in print.

Jones, A & Fox. S. (2009). *Generations online 2009*. Pew Internet & American Life Project.

Kim, M., Chung,N. & Lee, C. (2011). The effects of perceived trust on electronic commerce: Shopping online for tourism products and services in South Korea. *Tourism Management* 32(2): 256-265.

Migas, N., Anastasiadou, C. & Stirling, A. (2008). Individualized Tourism Brochures as a Novel Approach to Mass Customization, *Journal of Hospitality Marketing & Management* 17(1): 237-257.

Molz, J. (2011).Connectivity, collaboration, search. In M. Büscher, J. Urry, and K. Witchger (Eds.), *Mobile methods*. London: Routledge.

OECD (2008). *Information technology outlook 2008*. Paris: OECD.

Pan, B., Xiang, Z., Law, R. & Fesenmaier, D. (2011). The Dynamics of Search Engine Marketing for Tourist Destinations. *Journal of Travel Research* 50(4): 365-377.

Parra-López, E., Bulchand-Gidumal, F., Gutierrez-Tano, D. & Diaz-Armas, R. (2011). Intentions to use social media in organizing and taking vacation trips. *Computers in Human Behaviour* 27: 640-654.

Selnes F. & Troye S.V. (1989) Buying Expertise, Information Search, and Problem Solving. *Journal of Economic Psychology* 10: 411-428.

Subrahmanyam, K., Reich, S.M., Waechter, N. & Espinoza, G. (2008). Online and offline social networks: Use of social networking sites by emerging adults, *Journal of Applied Developmental Psycholog* 29(6): 420-433.

Tan, W. & Chen, T. (2011). The usage of online tourist information sources in tourism information search: an exploratory study. *The Service Industries Journal*, in print

Tussyadiah, I. & Fesenmaier, D. R. (2009). Mediating tourist experiences: access to places via shared videos. *Annals of Tourism Research* 36(1), 24-40.

Vogt, C., Stewart, S. & Fesenmaier, D. (1998). Communication strategies to reach first-time visitors. *Journal of Travel & Tourism Marketing* 7(2): 69–89.

Yoo, K. Lee, Y., Gretzel, U. & Fesenmaier, D. (2009). Trust in travel-related consumer generated media. *Information and Communication Technologies in Tourism* 2: 49-59.

Xiang, Z., Wober, K. & Fesenmaier, D.R. (2008). Representation of Tourism in Search Engines. Journal of Travel Research 47(2): 137-150.

Xiang, D. & Gretzel, U. (2010). Role of social media in online travel information search, *Tourism Management* 31(2): 179-188.

A Methodology for Building Microformats

Wilhelm Loibl, and
Dora Tüz

Institute for Tourism and Leisure Studies
Vienna University of Economics and Business, Austria
{wloibl, dora.tuez}@wu.ac.at

Abstract

As tourism is a very information intense industry, many prospective customers resort to well-known generic search engines like Google to find information about touristic offers. Often the results are not completely satisfactory. One possibility to enhance search engine results is through the use of microformats, XHTML-code which is inserted directly into the webpage and which gives the information contained within well defined meaning. To date several such microformats exist, all of which are for general purposes and not specifically for the tourism domain. In order to encourage the creation of tourism-related microformats the authors propose a methodology for microformats engineering based on existing ontology engineering methodologies. An example in the domain of events exemplifies the usage.

Keywords: microformats, tourism, engineering methodology, semantic web

1 Introduction

Travel and tourism is a very information-intense industry (Werthner & Klein, 1999) consisting mostly of small and medium-sized businesses (Dell'Erba, Fodor, Höpken, & Werthner, 2005). Along with the growing importance of the World Wide Web as a source of information and the decreasing cost of hard- and software many of these businesses set up their own website to provide information and communicate directly with prospective customers (Gratzer, Werthner, & Winiwartner, 2004). With such a huge wealth of information from many different sources available, online travellers and tourism professionals alike struggle to gather information of high quality (Scharl, Dickinger, & Weichselbraun, 2008). In order to find the information necessary for their trip planning many tourists resort to generic search engines like Google. These search engines act as "gateways" to travel-related information on the web and are important traffic generators for tourism-related websites (Xiang & Gretzel, 2010). The usefulness of current search engine technology is limited though. On the one hand this is because of the inability of many users to form precise queries. On the other hand most search engines still rely on matching the terms of the user's query against the (weighted) terms of the individual documents without considering the content itself (Berry & Browne, 1999).

A more thorough discussion about the shortcomings of search engines can be found in Henzinger (2007). What is needed in order to improve search results are so called Semantic Searches (Maedche & Staab, 2002). Web pages are enriched with additional metadata describing the content of the page (Davis, Studer, & Warren, 2006). The data taken from these documents can now be combined with information gathered from other documents to answer more complex queries. Constructing a web where machines can read and interpret information is the basic idea behind the "Semantic Web" (Antoniou & van Harmelen, 2008). Over the course of the last decade several

initiatives have created such systems of metadata for the tourism industry. These range from full-fledged ontologies, describing the whole domain of discourse in greater detail, to embeddable semantics which are inserted directly into the HTML-code of a webpage. Although the technology already exists, uptake by the industry has been rather slow (McDowell et al. 2003, Haustein & Pleumann, 2002). The reason lies in the missing motivation because much technical sophistication and substantial effort on the part of the website designer is necessary to structure the data. In order to justify structuring the data it must be rewarding and as easy as possible to do so. Microformats satisfy these conditions.

In order to encourage the creation of microformats for the tourism industry, the authors propose a new microformats engineering methodology building on existing methodologies for ontology engineering. This greatly facilitates the engineering process using well known tools and reduces the costs. In the remainder of section 1 the basic terminology used throughout the paper will be explained followed by a discussion about ontologies and microformats in the domain of tourism. Section 2 contains a discussion about the state of the art in ontology as well as microformats engineering and thereby laying the foundation for the research methodology (section 3) and the new microformats engineering methodology presented in section 4. The example given in section 5 helps to understand how the methodology is used in practice. The managerial implications given in section 6 outline how tourism businesses may benefit from this work while section 7 points out which steps must still be taken for the industry to fully profit from these technologies.

1.1 Basic Terminology

Some terms like domain or ontology have already been mentioned without further explanation of their meaning. The following section gives a short overview of important Semantic Web terminology used in this paper.

- Syntax vs. Semantics. The syntax is grammatical rules whereas semantics denote the meaning of an expression. In a web context semantics can be added to a webpage by including metadata (i.e. data about data) to an online document describing the content or the document itself (e.g. the author or creation date).
- Domain. A domain is a collection of all entities about a specific subject (Hjorland & Albrechtsen, 1995). The domain of tourism therefore consists of all technological elements and information entities related to travel (Xiang, Wöber, & Fesenmaier, 2008).
- Ontology. Ontologies provide a formal description of a certain domain (a finite list of terms and the relationships between these). Terms are classes of objects (concepts) which are important in the domain of discourse. The relationships are defined using ontology languages (Antoniou & van Harmelen, 2008). The process of manually constructing domain ontologies may become rather complicated according to the size and complexity of the domain.
- Embeddable Semantics is an umbrella term for all of those formats which use (X)HTML tags to integrate metadata into the code of a webpage itself. Two important such formats are RDFa and microformats.
- RDFa only defines the syntax and is interoperable with RDF. All vocabularies (the terms allowed in the document) are defined independently and are freely

intermixable (Adida, Birbeck, McCarron, & Pemberton, 2008[Sept. 3, 2011]). To extend RDFa in such a way, the RDF schema mechanism can be used (Adida & Birbeck, 2008[Sept. 3, 2011]).

- Microformats define the syntax as well as the vocabulary. They are generally designed for human consumption first, machine readability and interpretability comes only second. As they can only be used for describing explicit information more sophisticated knowledge representation like inference is not possible (Khare & Celik, 2006).
- A schema either defines the structure of a document (XML) or a set of terms (a vocabulary) which can be used in a data model (RDF(S)) (Antoniou and van Harmelen, 2008). In the latter case a schema adds additional meaning to data, providing guidelines for its interpretation like certain relationships or constraints (Allemang & Hendler, 2008).

1.2 Ontologies and Microformats in the Tourism Domain

Despite the difficulties inherent in constructing domain ontologies, several projects for the domain of tourism exist. An overview of these ontologies is given by Feilmayr and Pröll (2009). Although RDFa offers more expressivity (Adida, 2008), microformats are currently deployed in greater numbers (Lewis, 2010). Unfortunately, overall usage numbers are still quite limited. As Loibl (2011) shows for the Austrian market only about 1.2 % of all websites of tourism-related businesses contain semantic markup. Most of this markup is comprised of microformats giving general information about the address e.g. in the form of a vcard microformat. An example of such a format, taken from the website of an Austrian golf resort at http://www.golfresort.at is given below.

```
<address id="c-footer" class="vcard">
  <span class="org fn">
    <span class="organization-name">Golfresort
    Haugschlag</span>
  </span>,
  <span class="adr">
    <span class="postal-code">A-3874</span>
    <span class="locality none">Haugschlag</span>
    <span class="street-address">Haugschlag 160</span>
    <span class="country-name none">Austria</span>
  </span>, T:
  <span class="tel">
    <span class="type none">home</span>
    <span class="value">+43 (0)2865/8441-0</span>
  </span>
</address>
```

As illustrated by this example of a virtual business card above, microformats are added to web pages by inserting class-attributes with specified values (Lewis, 2010). For the human user this HTML code is rendered as a normal web page with address data. Any program capable of reading and interpreting the additional metadata the information contained in the page can be extracted and used for further purposes like

exporting the data into the address book of an e-mail program. In the tourism domain microformats could be used to automatise communication between companies. For example restaurants have to manually update data kept on restaurant search engines. Using microformats the changes must only be made on the company website and all partner businesses can automatically extract the new data from it. This is especially helpful for data that changes more often like weekly menu cards, room availability with hotels or air fares.

According to Lewis (2010) only 8 stable microformats exist with another 15 being in a draft status. None of these explicitly support the description of information specifically for tourists. In order to encourage the creation of microformats for the use in tourism businesses this paper provides an easy to follow step-by-step engineering approach.

2 State of the Art

As creating microformats is to some extent quite similar to the process of ontology engineering, the following section provides an overview of prevalent methodologies in this area.

2.1 State of the Art in Ontology Engineering

The ontology engineering methodology used depends on what type of ontology is to be created. Representation ontologies do not make claims about the world but instead provide a representational framework. The primitives provided form the basis for the description used in domain and generic ontologies. Domain ontologies describe concepts which are specific for a certain domain, e.g. accommodations while generic or upper level ontologies define concepts that are generic across many different domains. Application ontologies model all knowledge required for a particular application including method- and task-specific extensions. Such ontologies are not reusable themselves (van Heijst, Schreiber, & Wielinga, 1997).

Although different methodologies for building different types of ontologies exist, a common sequence of steps can be distinguished (Pinto & Martins, 2004). During the specification phase the purpose of the ontology is identified together with its intended areas of application. In the conceptual phase a conceptual model is formulated describing the concepts in the domain of discourse together with the relationships among them. Often groups of highly interlinked concepts can be partitioned into sub-ontologies. This conceptual model is then formalized using axioms which restrict possible interpretations of the meaning of the concepts contained in the ontology (formalization stage). The formal model is then implemented using a representation language and the implementation is committed to updates and corrections until the end of its life cycle. In addition to these sequential steps, other activities have to be performed throughout the life cycle. One of the most important activities for ontology engineering is the acquisition of knowledge. This can be accomplished through expert interviews, brainstorming, text analysis or any number of other techniques.

Evaluating the quality of prototypes according to the criteria defined in the specification phase helps guiding the evolutionary process. An extensive documentation regarding all decisions taken during the development process

facilitates maintenance as well as possible reuses of the ontology (Pinto & Martins, 2004). Generic ontologies are mostly built from scratch or by merging other generic ontologies. Several methodologies like TOVE (Grüninger & Fox, 1995) and METHONTOLOGY (Fernandez, Gomez-Perez, & Juristo, 1997), designed for building ontologies from scratch, gained a lot of attention in the relevant literature. According to Suarez-Figueroa (2010) even newer developments like DILIGENT (Pinto, Staab, & Tempich, 2004) and On-To-Knowledge (Staab, Schnurr, Studer, & Sure, 2001) do not provide detailed guidelines for reusing existing ontologies. The focus of these methodologies lies on collaborative engineering. Domain ontologies can be built from scratch as well or by reusing modules from other ontologies.

The methodology proposed by Pinto and Martins (2001) recommends that the process of composition should already start in the conceptualization phase and consists of seven consecutive steps. First the general integration possibility must be assessed. This depends on the framework that is being used for building the ontology. The individual building blocks (sub-ontologies) making up the future ontology together with the knowledge represented in each module and the assumptions related to the modules, have to be identified. Next possible candidate ontologies, representing the main concepts, have to be identified. If necessary, these ontologies must be reengineered before they can be evaluated. After the assessment, the most adequate source ontologies can be chosen and integrated into the resulting ontology. This integration process may involve adapting the information contained in a module, specialization (resulting in a more specific domain ontology) or inserting more general information (augmentation). The last step involves the evaluation of the resulting ontology according to specific evaluation criteria (Pinto & Martins, 2001).

2.2 State of the Art in the Creation of Microformats

For the design of microformats only guidelines and design principles exist. Allsopp (2007) even discourages the premature creation of a new microformat for a specific task. Instead all other possibilities for solving the problem at hand must be taken into consideration and only if these prove to be insufficient a new microformat may be created. Reuse is the first of three design principles defined by Khare (2006). An example would be the hCalender microformat which is a 1:1 implementation of the iCalendar standard defined in RFC2445 (Celik & Suda, 2011[Sept. 3, 2011]). Another important design principle is to reduce complexity. A microformat should always focus on solving a specific problem by employing the simplest solution.

The last design principle is recyclability. Microformats should be modular and easily embeddable. Based on these principles, Allsopp (2007) describes a series of tasks as a guideline for creating a microformat. The starting point is always a precise problem and not some hypothetical issue. This is important because it means that some solution in the form of a web page already exists. Wherever data is provided for human consumption first and machines second, there is an ideal field of application for microformats. If the problem at hand is generalisable for a greater number of users this constitutes an ideal candidate for a microformat. In order to determine generalisability, the engineer may turn to the community for advice. The

microformats community provides a mailing list for such discussions[1]. The next step involves researching current behaviour. A new format should contain only a minimal set of properties that is really needed and leave all excess data. To get an overview of the data the microformat has to structure, a large sample of real-world web pages of the respective domain ought to be investigated. The microformats community again helps by providing a wiki where the research process can be documented[2]. When a consensus has been reached as to the data that should be structured it is beneficial to search for existing standards already describing this data. Not only does reusing existing standards save a lot of time and effort it also improves the quality of the solution and interoperability with existing applications. At this stage the resulting format can be specified as a draft schema using an XHTML Meta Data Profile (Celik, 2003[Sept. 3, 2011]) and is open for discussion. An important indicator for maturity is usage. When a microformat is widely used it can be seen as mature. Such a pattern is characteristic for an evolving prototyping life cycle.

3 Research Methodology

A methodology in the context of this work is an integrated series of stages or phases that is used to structure a class of thought intensive work. A stage consists of a series of activities. Each activity is composed of tasks, which are the smallest units of work. These tasks are allocated to individual project members who are held accountable for their execution (Suarez Figueroa, 2010). The research methodology used for this work can be divided into two steps. In a first step, methodologies for software and ontology engineering as well as available design principles for microformats were contrasted with each other. This enabled the authors to recognize the focal point in the microformats engineering process. In addition, a preliminary methodology based on ontology engineering methodologies was created. In a next step this preliminary microformat engineering methodology was utilized to create a microformat for the tourism domain. The lessons learned from this practical application were then used to refine the initial engineering methodology.

4 Proposition of a Methodology for Creating Microformats

Ontology engineering methodologies stress the importance of knowledge acquisition (Pinto & Martins, 2004). This is especially true for microformats. As has already been noted, a microformat should be based on existing standards. Furthermore, it has to contain only the data necessary to solve a specific problem. Because of this importance of knowledge acquisition it has a very prominent position in the methodology and has to be explained more thoroughly. Design activities are often differentiated into conceptualization and formalization phases (Pinto & Martins, 2004). For building microformats such a distinction is not absolutely necessary as microformats are not as strictly formalized as ontologies are. Fig. 1 gives an overview of the methodology proposed in this work. It is based on ontology engineering methodologies and puts special emphasis on knowledge discovery. The remainder of section 4 describes the phases and individual steps of the methodology.

[1] See: http://microformats.org/mailman/listinfo/microformats-discuss
[2] See: http://microformats.org/wiki

In the specification phase the first step is problem definition. A problem that could be solved using microformats usually stems from some shortcoming of the website. Capturing motivating scenarios as defined by Grüninger & Fox (1995) help to uncover examples where a problem was not adequately addressed by current technology. These scenarios already hold the core of a possible solution. The next step defines the field of application. Useful tools are UML use cases showing certain scenarios how users interact with a system (Fowler, 2003) or informal competency questions (Fernandez-Lopez & Gomez-Perez, 2002). Any suitable microformat must be usable for answering these questions. In a web context a good starting point for formulating these competency questions lies in the search behaviour of the users of the website. Another use case would be the automatic extraction of a weekly menu card by gastronomy search engines. Here the main question would certainly be what type of information these companies present on their search platforms. The answer may be found on the administration web page of the respective search engines or by analysing the search web-interface.

Knowledge acquisition is definitely the most important. First, the information provided on the own website must be analyzed by extracting all information that is to be structured. For a restaurant this may be the online menu card. This information is collected into classes and provided with a label. In the remainder of this work these labels are referred to as terms. If someone wanted to create a microformat to structure an online menu card, the first source to investigate would most certainly be his or her own website. The information there likely would inform the user about different types of dishes like soup or desserts. Under the “soup”-heading the user would find different objects of type “soup” like French onion soup together with additional information about that soup like price or ingredients. Consequently “French onion soup” would be an object of type “dish” (class label) whereas the heading grouping all objects of type “soup” together would be generalized into the term “dishtype”.

The resulting microformat would contain a class “dishtype” containing one or several objects of type “dish”. On a web page these more general objects like “soup” would certainly be marked up using certain tags like header-tags or using tables. A resulting DOM-subtree may be used for structuring the emerging microformat. The Document Object Model offers a system-independent convention for representing and interacting with objects found in HTML-documents. Therefore a HTML-document may be represented as a tree (or more than one tree) where a subtree is a smaller part of that structure. As microformats are primary for human consumption and only secondary for machines retaining the structure discovered on domain specific websites is mandatory.

At this stage though it is still not determined that a new microformat has to be created at all. Possible alternatives may take the form of existing microformats, ontologies, XML schemas or some other way to structure online data like e. g. semantic HTML. Only if no suitable alternatives (as compared to the competency questions) emerge, the creation of a microformat is warranted. Based on the terms extracted from the own website, other relevant sources of information are examined.

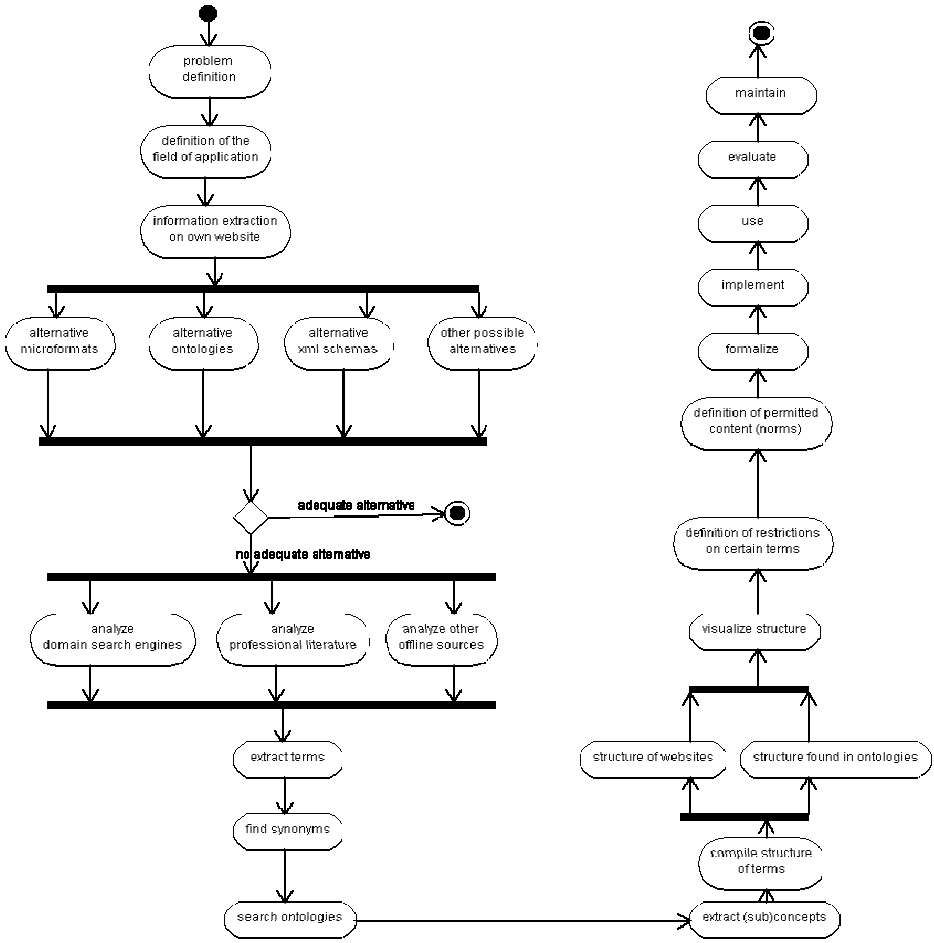


Fig. 3 Microformats Engineering Methodology

These sources are found offline, like catalogues or offline menu cards. Another source of inspiration is the search interfaces of domain search engines and professional literature. As with all external sources the legal terms must be considered. They regulate how if and how these sources may be used. All terms as well as their synonyms are then used for searching ontologies. A synonym for “beverage” would be “drink”. At this stage a semantic search engine is used to see if the concepts of “beverage” or “drink” are already described somewhere else. These descriptions may be utilized for defining the structure of the microformat (conceptualization phase).

The fundamental structure of the microformat is taken from the websites mimicking the visible structure of the web pages. If the visible structuring of the data is constant over all investigated websites the microformat should be grounded on that structure (e. g. most menu cards found on restaurant websites are similarly structured). Another important aspect is the reduction of the terms. Only terms that appeared in most

sources and can be seen as absolutely necessary to answer the competency questions should be incorporated into the microformat. A “dish”-element may provide information that a diabetic may eat it. But if no application uses that information it may be left out. The final step in the conceptualization phase is to visualize the resulting structure using either trees or graphs.

In the formalization phase certain restrictions on the terms may be defined. It must be noted that microformats are considerably less formalized than ontologies. Therefore this step is not mandatory. Much more important is the definition of the permitted content. When describing the price for the French onion soup it may be necessary to give details about the currency the price is in. For describing currency the norm ISO 4217 exists (http://www.currency-iso.org/iso_index/iso_tables/iso_tables_a1.htm). The results of the last two steps are summarized using an XHTML Meta Data Profile Description (Celik, 2003[Sept. 3, 2011]). The authors regard the description in the profile as a prototype that may be utilized for the uses defined during specification phase. All following phases aim at enhancing the prototype until suitable stability has been reached.

5 Example in the Tourism Domain

In order to learn more about the applicability of the preliminary methodology, a microformat for describing information about touristic events was created. The focus was on structuring information about events as well as any necessary additional information like tickets. The requirements were defined based on use cases. First a search for existing microformats and ontologies yielded the hCalendar microformat which is suitable for describing events as well as the ebSemantics event ontology (ebSemantics, n. d.[March 3, 2011]). These were found to be inadequate because they did not support the mark up of ticketing information. Consequently the existing microformat was to be extended. Next, a sample of online event calendars found on community websites as well as a sample of specialized event websites was examined manually. The majority of Austrian community websites provide an event calendar for local events aimed at locals and visitors alike.

A list of 2328 links to websites was compiled and a sample of 30 pages was chosen randomly. For each of these online event calendars the information presented on them was extracted manually and categorized into classes and labelled. The same was done for online event calendars. The results from the first result page of the search engine Google, using the keywords “event” and “österreich”, were analyzed. These terms together with their synonyms (gathered using WordNet[1]) were used as input into the semantic search engine Swoogle[2]. The results were analyzed manually to find any sub-ontologies that could be used for further describing the data discovered on the websites. In addition, a literature review was done to find ontologies not yet analyzed. The common structure discovered on the websites was enriched with suitable sub-ontologies. Afterwards, the core structure, necessary to accomplish the requirements defined at the beginning, was extracted and visualized using graphs.

[1] See: http://wordnetweb.princeton.edu/perl/webwn
[2] See: http://swoogle.umbc.edu/

At the end two XMD profiles, one for pricing the other for ticket-information, were defined. These two microformats are modular, the ticketing-microformats uses the pricing-microformat to mark up information about prices.

6 Managerial Implications

Most tourism businesses nowadays depend to a large extent on internet communication services for contacting their customers. Unfortunately the majority of companies is of small or medium size and often lacks the financial resources or the knowledge to build sophisticated websites. Microformats may help these firms solve some problems like poor visibility on the web. With this work the authors pursue several goals. First and foremost they want to spread the idea of using microformats. With rising usage numbers even more tourism businesses or web agencies may be persuaded to use them creating a snowball effect. Of course this vision can only become true if microformats exist which are useful for tourism businesses. To this end, a methodology for creating such formats is proposed in the work at hand which is precise enough to be used out of the box but sufficiently easy to be used by businessmen with less experience in microformats engineering. For the most part, the methodology was prepared for academics though and should be seen as an invitation to provide the industry with ready to use microformats.

7 Conclusions and Further Research

This work is supposed to provide a methodology that is easy to use and precise enough to create microformats for immediate use. The first issue, easy usage, is tackled by providing a clear sequence of steps that have to be followed when creating a new microformat. As the methodology is still in a very early stage of development, two main areas of discussion remain. Further research may discover that some important steps are still missing in the methodology. The second issue is automation. The work for compiling the example presented in section 5 was mostly done manually. Better tool support eases usability for less experienced microformats. This also applies to the second issue, precision. For each step shown in Fig. 1 an assortment of adequate methods has to be presented. These methods must be described in great detail so the user can decide which method to use. The two issues raised above show quite clearly that the current methodology is only a first step. Further research has to improve this methodology by employing it for the creation of microformats. All experience gained through application can then be fed back into a next revision. Other approaches like the stage-gate approach may further enhance the methodology (Cooper, Edgett, & Kleinschmidt, 2002)

Another problem is the slow uptake of the technology. In order to uncover barriers which impede the usage of microformats the authors will analyse tourism related websites in Austria according to their use of semantic metadata. Based on the insights gained through this analysis users and non-users of semantic formats will be questioned as to their motivations.

References

Adida. B. (2008). hGRDDL: Bridging microformats and RDFa. Web Semantics. *Science, Services and Agents on the World Wide Web* 6(1): 54-60.

Adida, B. & Birbeck, M. (2008). *RDFa Primer.* Retrieved from http://www.w3.org/TR/xhtml-rdfa-primer/

Adida, B., Birbeck, M., McCarron, S. & Pemberton, S. (2008). *RDFA in XHTML: Syntax and Processing*. Retrieved from http://www.w3.org/TR/2008/REC-rdfa-syntax-20081014

Allemang, D. & Hendler, J. (2008). *Semantic Web for the Working Ontologist modelling in RDF, RDFS, and OWL*. Amsterdam: Morgan Kaufmann.

Allsopp, J. (2007). *Microformats. Empowering Your Markup for Web 2.0*. Berkeley. Friends of ED.

Antoniou, G. & van Harmelen, F. (2008). *A Semantic Web Primer, second edition*. Cambridge: The MIT Press.

Berry, M. & Browne, M. (1999). *Understanding Search Engines. Mathematical Modeling and Text Retrieval*. Philadelphia: Society for Industrial and Applied Mathematics (SIAM).

Celik, T. (2003). *XMDP: Introduction and Format Description.* Retrieved from http://gmpg.org/xmdp/description

Celik, T. & Suda, B. (2011). *hCalendar 1.0.* Retrieved from http://microformats.org/wiki/hcalendar

Cooper, R.J., Edgett, S.J. & Kleinschmidt, E.J. (2002). Optimizing the Stage-Gate Process: What Best-practice Companies Do-I. *Research Technology Management* 45(5): 21-27.

Davis, J., Studer, R. & Warren, P. (2006). *Semantic Web Technologies: Trends and Reseach in Ontology-based Systems*. Chichester: Wiley.

Dell'Erba, M., Fodor, O., Höpken, W. & Werthner, H. (2005). Exploiting Semantic Web Technologies for Harmonizing E-Markets. *Information Technology & Tourism* 7: 201-219.

ebSemantics (n. d.). Retrieved from: http://www.ebsemantics.net/doc

Feilmayr, C. & Pröll, B. (2009). Ontologiebasierte Informationsextraktion im eTourismus. In M., Lassnig and S., Reich (Eds.), *eTourismus, Praxis der Wirtschaftsinformatik, Heft 270*. Heidelberg: Dpunkt.verlag.

Fernandez-Lopez, M. & Gomez-Perez, A. (2002). Overview and Analysis of Methodologies for Building Ontologies. *The Knowledge Engineering Review* 17(2): 129-156.

Fernandez-Lopez, M., Gomez-Perez, A. & Juristo, N. (1997). METHONTOLOGY: from ontological art towards ontological engineering. In *Proceedings of AAAI97 spring symposium series, workshop on ontological engineering*. Stanford, CA.

Fowler, M. (2004). *UML Distilled. Third Edition. A Brief Guide to the Standard Object Modeling Language*. Amsterdam: Addison Wesley.

Gratzer, M., Werthner, H. & Winiwarter, W. (2004). Electronic business in tourism. *International Journal of Electronic Business* 2(5): 450-459.

Grüninger, M. & Fox, M. (1995). Methodology for the design and evaluation of ontologies. In *Proceedings of IJCAI95's workshop on basic ontological issues in knowledge sharing*. Montreal, Canada

Haustein, S. & Pleumann, J. (2002). Is Participation in the Semantic Web Too Difficult? In I. Horrocks & J. Hendler (Eds.), *ISWC 2002, LNCS 2342*.

Henzinger, M. (2007). Search Technologies for the Internet. *Science* 317(5837): 468-471.

Hjorland, B. & Albrechtsen, H. (1995). Toward a New Horizon in Information Science: Domain-analysis. *Journal of the American Society for Information Science* 46(6): 400-425.

Khare, R. & Celik, T. (2006). Microformats: a pragmatic path to the semantic web. In *Proceedings of the 15th international conference on World Wide Web*. Edinburgh, Scotland: ACM Press.

Khare, R. (2006). Microformats: The Next (Small) Thing on the Semantic Web? *IEEE Internet Computing* 10 (1): 68-75.

Lewis E. (2010). *Microformats made simple*. Berkeley: New Riders.

Loibl, W. (2011). An Overview of the Use of Semantic Markup on Tourism Websites in Austria. In *TTRA Europe 2011 Conference Proceedings. Creativity and Innovation in Tourism*. Archamps, France.

Maedche, A. & Staab S. (2002). Applying Semantic Web Technologies for Tourism Information Systems. *Proceedings from ENTER 2002*. Innsbruck, Springer.

McDowell, L., Etzioni, O., Gribble, S., Halevy, A., Levy, H., Pentney, W. & Vlasseva, S. (2003). *Evolving the Semantic Web with Mangrove*. Technical Report UW-CSE-03-02-01, February 2003.

Pinto, H. S. & Martins, J. P. (2001). A methodology for ontology integration. In *Proceedings of the 1st international conference on knowledge capture*. New York: ACM Press.

Pinto, H. S. & Martins, J. P. (2004). Ontologies: How can They be Built? *Knowledge and Information Systems* 6. London: Springer.

Pinto, H. S., Staab, S. & Tempich, C. (2004). DILIGENT: Towards a fine-grained methodology for Distributed, Loosely-controlled and evolving Engineering of oNTologies. In R. L. Mantaras & L. Saitta. *Proceedings of the 16th European Conference on Artificial Intelligence (ECAI 2004)*. Valencia, Spain.

Scharl, A., Dickinger, A. & Weichselbraun, A. (2008) Analyzing News Media Coverage to acquire and structure Tourism Knowledge. *Information Technology & Tourism* 10: 3-17.

Simperl, E. P. B. & Mochol, M. (2006). Cost Estimation for Ontology Development. In W. Abramowicz (Ed.). Proceedings of BIS 2006. Poznan, Poland.

Staab, S., Schnurr, H. P., Studer, R. & Sure, Y. (2001). Knowledge Processes and Ontologies. *IEEE Intelligent Systems* 16(1): 26-34.

Suarez Figueroa, M. C. (2010). NeOn Methodology for Building Ontology Networks: Specification, Scheduling and Reuse. (Unpublished doctoral dissertation). Universidad Politecnica de Madrid.

Van Heijst, G., Schreiber, A. T. & Wielinga, B., J. (1997). Using explicit ontologies in KBS development. *Int. J. Human-Computer Studies* 45: 183-292.

Werthner, H. & Klein, S. (1999). *Information Technology and Tourism - A Challenging Relationship*. Wien, New York: Springer.

Xiang, Z. & Gretzel, U. (2010) Role of social media in online travel information search. *Tourism Management* 31: 179-188.

Xiang Z., Wöber K. & Fesenmaier D. (2008). Representation of the Online Tourism Domain in Search Engines. *Journal of Travel Research* 47(2): 137-150.

TripAdvisor.com vs. NYCGO.com: Evaluation of Functional Components of Generalist and Specialist Tourism Websites

Young Shin Lim[a] , and
Eunice Eunjung Yoo[b]

[a] School of Communication
The Ohio State University, USA
lim.956@osu.edu

[b] Department of Recreation, Park and Tourism Management
The Pennsylvania State University, USA
eyoo@psu.edu

Abstract

While tourism websites are becoming increasingly pervasive, there is a lack of empirical research that captures and documents the quality of different types of tourism websites. In this regard, this study compares the effectiveness of generalist and specialist tourism websites in terms of the functional components they employ to communicate travel information. The findings of the study suggest that there indeed exist variations in the functional features adopted by websites to communicate travel information between generalist and specialist tourism websites, and the differences fundamentally come from the scale and technical assets of the websites. The findings provide valuable illustration of the characteristics and limitations generalist and specialist online media can typically possess. In addition, there are implications for tourism information providers and tourism website users in terms of better system design and more effective travel planning.

Keywords: tourism websites; TripAdvisor; destination marketing; generalist vs. specialist media

1 Introduction

How do people get travel information when they plan a trip somewhere? Nowadays most people browse tourism websites that provide information regarding their potential destinations. The Internet enables tourists to access reliable and accurate information as well as to undertake reservations in a fraction of time, cost, and inconvenience required by conventional methods (O'Connor & Murphy, 2004). Therefore, increasing numbers of people are using the Internet for travel information search, and many Internet websites provide more in-depth materials and richer content compared with conventional promotional agents (Govers & Go, 2003).

As a result, tourism websites are gradually replacing the role played by traditional paper-based travel guide, and there exist overwhelming numbers of tourism websites on the Web. Now, the key is not to get access to many different tourism websites, but to discern tourism websites which can provide better support for travel planning. However, while tourism websites are becoming increasingly important, there is a lack of empirical research to describe and explain differences in quality and effectiveness of tourism websites. Therefore, this study aims at investigating the effectiveness of tourism websites through the lens of specialist vs. generalist online media.

When travellers search for travel information, they tend to browse through multiple websites. These websites fall into two categories: the websites that host information on multiple destinations and require a user to specifically search for a particular destination (i.e. generalist website such as TripAdvisor.com), and those websites that are dedicated to information on one very specific destination (i.e. specialist website, such as NYCgo.com). Choosing between the two is the typical decision tourists make when they search for travel information, yet the comparison between the effectiveness of generalist and specialist websites has been somewhat ignored.

This paper tries to examine the effectiveness of specialist and generalist tourism websites with emphasis on their characteristics as online information source. On one hand, tourism websites are electronic versions of paper-based travel guides, but on the other, the range and format of information communicated through tourism websites greatly differ from the information delivered by their traditional counterpart. As for online media, we assume that the quality of information and the overall effectiveness of tourism websites are largely influenced by the website features they adopt to communicate travel information. For example, there are many travel websites which employ user review functions. By adopting the feature, the websites are enabling communication among the users and provide a wider range of travel information from multiple perspectives. In this regard, we aim to compare the effectiveness of generalist and specialist tourism websites in terms of functional components employed by them. Further, we try to document the current trends in tourism website design, and provide better knowledge on quality of tourism websites to potential travellers. We believe the results can also provide valuable implications regarding characteristics and limitation of overall specialist and generalist online media.

2 Specialist Media vs. Generalist Media

Specialization of media content has shown an upward trend ever since the advent of cable television in the United States. In many cases, the word specialization implies notion of expertise in the specific domain, and similarly, media users may now expect specialization in content beyond television channels as it affords powerful psychological connotations regarding excellence in quality of content offerings (Koh & Sundar, 2007; Koh & Sundar, 2010). Researchers have previously demonstrated the relationship between specialist and generalist television sets and the viewer's perception of the content provided by this medium (Nass, Steuer, Henriksen, & Dryer, 1996). There was also a study on the effects of specialization in web agents on viewers' perceptions of the information delivered by the media (Koh & Sundar, 2010).

However, most researches focused on how receivers perceive information from specialist and generalist media differently, and no claims have been made about the actual effectiveness of specialist websites versus generalist websites. Therefore, the goal of this research is to evaluate actual quality of specialist and generalist tourism websites by investigating the functional components they are adopting to communicate travel information.

3 Tourism Websites and Model of DMS Website Evaluation

Due to its intangible nature, tourism is an information-intense industry, and communication of tourism information is the key objective in travel marketing. Therefore, it is critical to understand the nature of tourism websites, i.e., the composition of online tourist-related information potentially available to tourists, provides an important stepping-stone for the development of successful marketing programs and better information systems in tourism (Fesenmaier, Wober, & Werthner, 2006; Xiang, Wober & Fesenmaier, 2008). Moreover, tourism websites directly influence the perceived image of the destination and creates a virtual experience for tourists. The way a website presents information graphics, and photos in many ways influences how these messages are received by "viewers" (Rosen & Puriton, 2004). Therefore, well-designed tourism websites are of great importance to both travellers and tourism marketers.

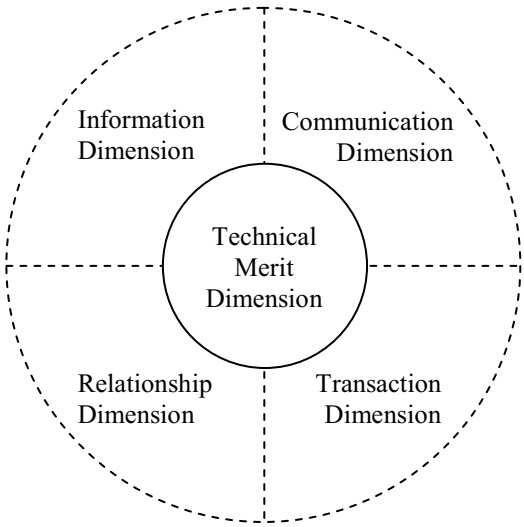


Fig. 1. A conceptual model of destination marketing organizations' website evaluation (Source: adapted from Li & Wang 2010)

In this regard, there has been increased concern on evaluation of tourism websites. There have been studies on different travel related websites, such as for airlines (Chu, 2001) and hotels (Wong & Law, 2005) from varying perspectives. In this paper, to evaluate the quality of generalist vs. specialist tourism websites, we employ Li and Wang's (2010) conceptual model of destination marketing organization's website evaluation. The model extends Wang and Russo's Information, Communication, Transaction, and Relationship (ICTR) model (2007) by employing Technical merit as a fifth dimension (Figure 1). The major strength of the model is that it highlights the trend that tourism websites are evolving from static, mere transfer of traditional travel guide to dynamic sites with increasing levels of interactivity and rich functionality (Li & Wang, 2010). The model considers destination marketing sites to be composed of

five interrelated dimensions 1) information: up-to-date, accurate delivery of information, 2) communication: effective and sustainable communication with the users, 3) transaction: easy and reliable transaction, 4) relationship: sustainable relationship building with the users, and 5) technical merit: features previously did not exist, but enabled by the Internet such as adoption of different modalities (Li and Wang, 2010; Wang and Russo, 2007). According to the model, every item that exists on tourism websites falls into one of five component categories and influences the effectiveness of the website. In this regard, this study aims to compare specialist and generalist tourism websites based on Li and Wang's framework. Further illustrations of the five dimensions are presented on the method section.

4 Research Question & Hypotheses

Based on our literature review, the research question and hypotheses are derived as follow:

RQ: For tourism websites, what is the relationship between the level of specialization and the website's effectiveness originated from the adopted website's functional components?

The main hypothesis is:

Hypothesis: There exist differences between generalist and specialist tourism websites' effectiveness originated from the adopted functional components.

Based on this, the main hypothesis can be further refined into sub-hypotheses:

H1: There exist differences between generalist and specialist tourism websites' effectiveness originated from the adopted information components.

H2: There exist differences between generalist and specialist tourism websites' effectiveness originated from the adopted communication components.

H3: There exist differences between generalist and specialist tourism websites' effectiveness originated from the adopted transaction components.

H4: There exist differences between generalist and specialist tourism websites' effectiveness originated from the adopted relationship components.

H5: There exist differences between generalist and specialist tourism websites' effectiveness originated from the adopted technical merit components.

5 Methods

5.1 Content Analysis

In order to test the hypotheses, content analysis of tourism websites of top ten U.S. travel destinations in terms of cities was conducted. The unit of analysis was an entire website for specialist websites and each city pages for generalist websites, and different layers within the websites and the city pages were examined. The independent variable was the level of specialization. It was measured as either generalist or specialist. While the former deals with multiple destinations (generally they cover international destinations), the latter focuses on one city. Therefore, the independent variable was consistently measured at the nominal level. The dependent variable covers the five functional dimensions, information, communication, transaction, relationship, and technical merit. Every item existing on tourism websites was coded as one of the dimension. The codes for content analysis were mainly employed from Li and Wang's (2010) classification of websites' items for the model of DMO website evaluation, and there were some inductive changes based on the features of the sampled tourism websites.

Here are the details about the five dimensions and the lists of codes.

Information dimension: Information dimension does not deviate much from the role of traditional paper-based travel guides. Information dimension is measured by the range of information provided by tourism websites and the features which help websites' visitors find the information they need (Wang, 2008; Wang & Russo, 2007). The codes (i.e. website features) belonging to this variable were content classification codes considering features such as *attraction information, maps & directions, destination background information & travel tips, themed menu categories, transportation information, event calendar, event information, restaurant information, printable travel guide/brochure, accommodation information, travel package, local weather information, shopping information, trip/vacation planner, link to nearby cities, neighborhood guide, pop-up information, visitor center information,* and *featured highlights*.

Communication dimension: Communication dimension is measured by features which enabled dialogue between the users and the website as well as among users (Wang & Russo, 2007). *Search function, online forum, user gallery, user review, frequently asked questions, email newsletter, brochure request capabilities, whether the website has social media accounts,* and *whether they provide some features which enable the users to connect to their social media accounts* were included in the codes for the communication dimension.

Transaction dimension: With transaction dimension we tried to capture the features which enable visitors to actually make purchases for their upcoming travel directly on the website or advertisements which promote transaction on other websites (Wang, 2008; Wang & Russo, 2007). *Transportation reservation, accommodation reservation, restaurant reservation, attraction ticket reservation, event ticket reservation, merchandizing (such as selling paper-based travel guide on the site), classified ads,* and *pop-up ads* were coded for this variable.

Relationship dimension: The relationship dimension captured those website features which help to solidify the relationship between the website and the users (Wang & Russo, 2007). Accordingly, the codes belonging to this variable were *membership, personalization, contact information, best deal/special offer, virtual tours,* and *web seal certification.*

Technical merit dimension: For the technical merit dimension we looked at the features which deliver format changes in light of the Internet (Li & Wang, 2010). The codes for the technical merit dimension were *multimedia, slideshow,* and *multiple languages*.

Most codes were measured at the nominal level, which means we focused on whether or not the websites were adopting the coded features to provide information to potential travellers. Exceptionally, five codes in transaction dimension *transportation reservation, accommodation reservation, restaurant reservation, attraction ticket reservation, event ticket reservation,* were coded at the ordinal level. Those codes were analyzed with three levels based on whether the websites provide 1) on site reservation, 2) links to reservation website(s), and 3) no features for reservation.

5.2 Sampling and Data Collection Procedure

Overall, 40 tourism websites across ten US cities (two generalists and two specialists per city) were chosen for the analysis. The ten destinations were selected based on the "2008 Overseas Visitation Estimates for U.S.", for which data was provided by the US Department of Commerce, ITA, Office of Travel and Tourism Industries. The ten cities were chosen based on the ranking. Exceptionally, for the case of Honolulu (ranked No. 7), most tourism websites were either about Hawaii or Oahu rather than about the city of Honolulu. Therefore, it was replaced with Philadelphia, which was ranked No. 11. The final ten cities were New York City, Los Angeles, San Francisco, Miami, Orlando, Las Vegas, Washington DC, Chicago and Boston.

For each of the destinations, we selected two generalist websites and two specialist websites. For generalist websites, we navigated to the city overview page of the specific city within the two generalist tourism websites - TripAdvisor (http://www.tripadvisor.com) and Lonely Planet (http://www.lonelyplanet.com). The two generalist websites were selected based on Alexa's ranking of travel guides and directories websites. The rationale for the selection is that we assume the website's popularity as the most important criteria when potential travellers choose a generalist travel website. TripAdvisor and Lonely Planet ranked No. 1 and No. 3, respectively, and the No. 2 website Yahoo! Travel was excluded from the sampling, because the website (Yahoo!) is not primarily pertaining to the travel domain. In sum, ten city information pages were sampled from both TripAdvisor and Lonely Planet. Specialist websites were chosen based on Google search results. We searched the city name, tourism, travel, and guide altogether on Google, and selected the first two relevant websites with dedicated and complete tourism information. Evidently, this eliminated travel websites (i.e. that are usually aggregators for reservation functions), rating websites (that rate different attractions within a particular city) and city websites (that hosted non-tourism information such as job-search details, real estate etc). Table 1 presents the list of specialist websites.

Table 1. Lists of Specialist Tourism Websites

City	Specialist Websites	
New York City	NYC The Official Guide http://www.nycgo.com/	NYC Tourist http://www.nyctourist.com/
Los Angeles	Los Angeles The Official Guide http://discoverlosangeles.com/	LA Tourist http://www.latourist.com/
San Francisco	Only in San Francisco http://www.onlyinsanfrancisco.com/	San Francisco Guide http://www.sfguide.com/
Miami	Miami – Express Yourself http://www.miamiandbeaches.com/	Miami Florida http://www.miami-florida.com/
Orlando	Visit Orlando http://www.orlandoinfo.com/	Orlando Travel Guide http://orlandotouristinformationbureau.com/
Las Vegas	Las Vegas http://www.visitlasvegas.com/	The Official VEGAS Travel Site http://www.vegas.com/
Washington DC	Washington DC http://washington.org/	The District - The Tourist's Guide to Washington, DC http://www.thedistrict.com/
Chicago	Choose Chicago http://www.choosechicago.com/	Explore Chicago http://www.explorechicago.org/
Boston	Boston Travel and Tourism Destination Guide http://www.startthereboston.com/	Boston Guide http://www.bostonusa.com/
Philadelphia	Philadelphia and The Countryside http://www.visitphilly.com/	Philadelphia http://www.philadelphiausa.travel/

All sampled websites were coded to completely accommodate all the elements on the entire websites and the city pages. The coding was conducted by single independent coder who underwent intensive training and discourse sessions. Since websites are dynamic and constantly correcting in most cases to remain up-to-date, the period of coding is extremely important. The analysis of the websites was done within a one week timeframe – from the 3rd through the 10th of October, 2010.

6 Results

To examine the relationship between level of specialization and functional dimensions of tourism websites, we conducted Chi-square tests when the dependent variable was nominal and t-tests when the dependent variable was ordinal. Basically, each code was tested separately, since even though the codes were grouped with similar function, there can be differing requirements for the adoption of features. In this regard, we did not make a quantitative aggregation of each dimension, but rather we tried to draw the meaning by qualitatively summarizing the results for each dimension.

In terms of information dimension, specialist websites reported better adoption of several website features. More precisely, for *themed menu category* (χ^2 (1, N = 40) = 23.02, p = .00), *event calendar* (χ^2 (1, N = 40) = 24.00, p = .00), *event information* (χ^2 (1, N = 40) = 19.80, p = .00), *trip planner* (χ^2 (1, N = 40) = 5.71, p = .02), and *visitor center information* (χ^2 (1, N = 40) = 13.33, p = .00), specialist websites were more

likely to using the features to deliver travel information. Generalist websites were better only for the use of *destination background information & travel tips* feature (χ^2 (1, N = 40) = 10.16, p = .00). There were five features which all sampled websites had, and they were *attraction information, transportation information, restaurant information,* and *accommodation information*. In opposite, none of the websites had *non-ad pop-up information*. Moreover, there was no significant difference in the adoption of features *maps & directions, printable travel guides/brochures, travel package, local weather information, link to nearby cities,* and *neighborhood guide* between generalist websites and specialist websites.

By contrast, generalist websites excelled in the communication dimension. Significantly more generalist websites adopted the features *search function* (χ^2 (1, N = 40) = 7.06, p = .01)*, online forum* (χ^2 (1, N = 40) = 40.00, p = .00)*, user gallery* (χ^2 (1, N = 40) = 13.33, p = .00)*, user review* (χ^2 (1, N = 40) = 32.73, p = .00)*, email newsletter* (χ^2 (1, N = 40) = 10.00, p = .00)*, whether the website has social media accounts* (χ^2 (1, N = 40) = 4.44, p = .04)*,* and *whether they provide some features which enable the users to connect to their social media accounts* (χ^2 (1, N = 40) = 17.14, p = .00) than the specialist websites did. There was no significant difference in the use of *frequently asked questions*, and specialist websites were more likely to utilizing *paper-based brochure request capabilities* than the generalist websites (χ^2 (1, N = 40) = 8.49, p = .00).

For the transaction dimension, generalist websites displayed more instances for *transportation reservation* (t(38) = 4.25, p = .00)*, accommodation reservation* (t(38) = 2.37, p = .02)*,* and *classified ads* (χ^2 (1, N = 40) = 4.44, p = .04)*,* while specialist websites were more likely to present *attraction ticket reservation* (t(38) = -5.20, p = .00) and *event ticket reservation* (t(38) = -3.12, p = .00)*.* There was no significant difference in terms of *merchandizing, pop-up ads* and *restaurant reservation*.

For the relationship dimension, there was one website feature all sampled websites had and another feature that none of the websites had, they were *contact information* and *virtual tour*, respectively. There was no significant difference in terms of *personalization* in between generalist and specialist websites. However, for *membership* (χ^2 (1, N = 40) = 21.54, p = .00) *and web seal* (χ^2 (1, N = 40) = 10.00, p = .00) generalist websites were more likely to using the features, and specialist websites were better only for the use of *best deals/special offer* feature (χ^2 (1, N = 40) = 7.62, p = .01).

Finally, for the technical merit dimension, more generalist websites were using *multimedia* (χ^2 (1, N = 40) = 15.17, p = .00) and *multiple languages* (χ^2 (1, N = 40) = 13.33, p = .00) features, and there was no significant difference in terms of the adoption of *slideshow*.

In sum, our results demonstrated differences in effectiveness between specialist and generalist websites. Our sub-hypotheses are supported as indicated below:

> H1: There exist differences between generalist and specialist tourism websites' effectiveness originated from the adopted information components

Supported. While specialist websites were more likely to report *themed menu category, event calendar, event information, trip planner* and *visitor center*

information, generalist websites were more likely to report *destination background information & travel tips* feature.

H2: There exist differences between generalist and specialist tourism websites' effectiveness originated from the adopted communication components.

Supported. Generalist websites were more likely to report most of the features covering the communication dimension *search function, online forum, user gallery, user review, email newsletter, websites' social media accounts,* and *connectivity to users' social media accounts,* while specialist websites were more likely to report *paper-based brochure request capabilities*.

H3: There exist differences between generalist and specialist tourism websites' effectiveness originated from the adopted transaction components.

Supported. While specialist websites were more likely to report *attraction ticket reservation* and *event ticket reservation*, generalist websites were more likely to report *transportation reservation, accommodation reservation* and *classified ads*.

H4: There exist differences between generalist and specialist tourism websites' effectiveness originated from the adopted relationship components.

Supported. While generalist sites were more likely to report *membership* and *web seal*, specialist websites were more likely to display *best deals/special offer* features.

H5: There exist differences between generalist and specialist tourism websites' effectiveness originated from the adopted technical merit components.

Supported. Generalist websites were more likely to report *multimedia* and *multiple languages* features.

Overall, the main hypothesis was supported and therefore, the specific research question was answered. There were significant differences between generalist and specialist websites. A further discussion of results is provided below.

7 Discussion

This study examined the effect of the level of specialization in tourism websites on the effectiveness based on the adoption of different functional website components. The results indicated significant differences in employment of website features in different functional dimensions between specialist and generalist tourism websites. Since both specialist and generalist websites had certain website features they are significantly better in terms of the utilization, we were unable to conclusively establish one as being better than another. However, we were able to ascertain those features that cause these differences in websites' function, and their close relationship with the characteristics of specialist and generalist websites.

For the information dimension, both specialist and generalist websites provided basic fixed content information, such as transportation and accommodation. However, for information or planning tools which require in depth knowledge about the destination or timely updates, specialist websites showed better performance. There can be two explanations for this. First of all, for most cases of specialist websites, the agents were

located in the specific cities. Therefore, we assume that specialist websites had better access to timely information such as event dates, and could have more connection to local tourism service providers. Moreover, since generalist websites deal with multiple destinations, collecting in depth information and building related tools for all the locations might not be cost effective for them.

For the adoption of communication component, generalist websites reported far more cases. Generalist websites overwhelmed specialists regarding the adoption of website features which make users actively engage in communication, such as *online forum, user gallery, user review, websites' social media accounts,* and *connectivity to users' social media accounts,* while specialist websites were only more likely to report *paper-based brochure request capabilities*. We would like to attribute this to the scales of the websites. Generalist websites have more visitors, and visitors are more likely to sign up for membership. Therefore, generalist websites are in better positions to utilize the 'wisdom of crowds', and in doing so they show more interactivity features for user generated content, such as forums and photo galleries. Also those features require higher level of technology sophistication; therefore, it might be unfeasible for specialist websites to build such system for themselves.

For the utilization of transaction components, generalist websites were better in terms of supporting transportation and accommodation reservation, while specialist were better supporting reservation of event tickets and attraction tickets. We assume this is also related to the scale and purpose of the websites. For many generalist websites, taking the role of intermediary between travellers and airlines or hotel chains is the major part of their business, and they have assets to build the kind of system. However, the purpose of many specialist websites is more likely to be the introducing the city to potential visitors, and they rather tend to intermediate the transactions between visitors and local service providers.

The results were similar for the relationship dimension and the technical merit dimension. One noticeable point about the relationship dimension is that generalist websites were significantly more likely to employ membership system. On the one hand, this is because generalist websites utilize more features which require membership such as forum, user reviews and reservation functions. But on the other hand, it is also related to the fact that generalist websites provide information on multiple destinations. In the case of specialist websites, travellers do not visit the website after they traveled the destination. However, travellers usually visit generalist websites repeatedly, when planning for their next trip, especially, when sustainable customer relationships were already built. In this regard, relationship building is more difficult for specialist websites, and it also affect the utilization of other website features which require membership.

Lastly, generalist websites were significantly better in terms of utilization of technical merit. This accords with other results related to technical sophistication, and high visitor traffic would have made technology adoption more cost effective.

To conclude, all of our hypotheses were supported, and we found significant differences between generalist and specialist tourism websites across all five functional dimensions documented in this study. The practical implications of these distinctions are manifold. Successfully integrating information systems, such as

websites, into the fabric of tourism organizations and their marketing strategies is critical for the survival/success of Destination Marketing Organizations (DMOs) as promoters of destinations and brokers of information between tourism suppliers and consumers (Buhalis, 2000; Gretzel et al. 2006; Wang & Fesenmaier, 2006). Also, effective web-based marketing requires taking full advantage of websites as a communication medium (Brownlie et al., 1994; Cronin, 1995). The findings of the current study suggest that specialist websites (the category DMO websites fall into) have some strength over generalist websites content-wise. However, they were weak in those points which demand scale and technical sophistication. This suggests important points of improvement for both, specialist and generalist websites and how the two can complement each other.

We believe the findings of the current study are valuable in two aspects. First, the effectiveness of a website provides insights for tourism marketers towards a better organization of travel related information by designing effective tourism information systems, and for implementing successful marketing campaigns. Second, as potential tourists, website users can get a richer and deeper understanding of which website to consider when they are looking for information for their preferred tourism destination. Also, from a atheoretical aspect we could examine the difference between specialist and generalist online media, and the limitations they possess in terms of adopting certain website features.

8 Limitations and Future Research

Primarily, the biggest study limitation was the fact that we studied only 40 websites across ten US destinations. We assume that with our choice of American destinations, and American websites, our external validity may be hampered. Further, we chose to study only two generalist websites as opposed to two distinct specialist websites for each of the 10 destinations, which hinders our ability to draw certain conclusions about generalist sites (e.g. if or not generalist websites assume a standardized layout). Therefore, future studies with a larger sample covering more generalist websites and more destinations are crucial for further generalization of the findings from this study.

References

Alexa Internet, Inc. (2010). Alexa ranking of travel guides and directories websites. Retrieved January 10, 2010, from http://www.alexa.com/topsites/category/Top/Recreation/Travel/Guides_and_Directories

Buhalis, D. (2000). Marketing the competitive destination of the future. *Tourism Management 21*(1): 97-116.

Choi, S., Lehto, X. Y. & Morrison, A. M. (2007). Destination image representation on the web: Content analysis of Macau travel related websites. *Tourism Management 28*(1): 118-129.

Chu, R. (2001). What online Hong Kong travelers look for on airline/travel websites? *International Journal of Hospitality Management 20*(1): 95-100.

Cronin, M. J. (1995). *Doing more business on the internet: How the electronic highway is transforming American companies.* New York: Wiley.

Fesenmaier, D. R., Wöber, K. W. & Werthner, H. (2006). *Destination recommendation systems: Behavioral foundations and applications.* Boston: CABI.

Govers, R. & Go, F. M. (2005). Projected destination image online: Website content analysis of pictures and text. *Information Technology & Tourism 7*(2): 73-89.

Koh, Y. J. & Sundar, S. S. (2007, May). *Costco.com or Wine.com? Effects of specialization in web agents, web sites and computers on e-commerce trust.* Paper presented at the 57th annual conference of the International Communication Association, San Francisco, CA.

Koh, Y. J. & Sundar, S. (2010). Heuristic Versus Systematic Processing of Specialist Versus Generalist Sources in Online Media. *Human Communication Research 36*(2). 103-124.

Li, X. & Wang, Y. (2010). Evaluating the effectiveness of destination marketing organisations' websites: evidence from China. *International Journal of Tourism Research 12*(5): 536-549.

Nass, C., Steuer, J., Henriksen, L. & Dryer, C. (1994). Machines, social attributions, and ethopoeia: Performance assessments of computers subsequent to "self-" or "other-" evaluations. *International Journal of Human-Computer Studies 40*(3): 543-559.

O'Connor, P. & Murphy, J. (2004). Research on information technology in the hospitality industry. *International Journal of Hospitality Management 23*(5): 473-484.

Rosen, D. E. & Purinton, E. (2004). Website design: Viewing the web as a cognitive landscape. *Journal of Business Research 57*(7): 787-794.

Sundar, S. S., Xu, Q. & Oeldorf-Hirsch, A. (2009). Authority vs. peer: How interface cues influence users. In *Proceedings of the 27th International Conference Extended Abstracts on Human Factors in Computing Systems*, 4231-4236. Boston, MA.

U.S. Department of Commerce. (2009). *Overseas visitation estimates for U.S. states, cities, and census regions: 2008*. Washington, DC.

Wang Y. (2008). Web-based destination marketing systems: Assessing the critical factors for management and implementation. *International Journal of Tourism Research 10*(1): 55-70.

Wang, Y. & Fesenmaier, D. R. (2006). Identifying the success factors of web-based marketing strategy: An investigation of convention and visitors bureaus in the united states. *Journal of Travel Research 44*(3): 239-254.

Wang, Y. & Russo, S. (2007). Conceptualizing and evaluating the functions of destination marketing systems. *Journal of Vacation Marketing 13*(3): 187.

Wong, J. & Law, R. (2005). Analysing the intention to purchase on hotel websites: a study of travellers to Hong Kong. *International Journal of Hospitality Management 24*(3): 311-329.

Xiang, Z., Wober, K. & Fesenmaier, D. R. (2008). *Representation of the online tourism.*

Hotel Information Exposure in Cyberspace: The Case of Hong Kong

Rosanna Leung, and
Rob Law

[a]School of Hotel and Tourism Management
The Hong Kong Polytechnic University, Hong Kong
rosanna.leung@connect.polyu.hk
hmroblaw@polyu.edu.hk

Abstract

Search engines are an everyday tool for Internet surfing. They are also a critical factor that affects e-business performance. This study compares the quantity of hotel web pages and pages indexed by search engines. Values of index ratio indicated that international chain hotels were not well indexed by domestic search engines. Furthermore, the web servers of local chain and independent hotels of Hong Kong contain a large amount of outdated or unlinked materials. Though how search engine index pages cannot be controlled, hotel managers could control what should be indexed. Web masters were recommended to relocate these materials to independent folders and use robots exclusion protocol to restrict the access of web crawlers and robots.

Keywords: index ratio, robots exclusion protocol, search engine, web content, hotel website

1 Introduction

No matter how rich the information provided on a website (functionality) is and how user-friendly the website is (usability), if a website cannot be found from a search engine, it remains unknown to potential customers and business partners. Web surfing currently becomes a major information search channel. As such, tourism organizations have adopted the Internet as one of the primary marketing channels for attracting customers (Buhalis & Law, 2008). Nowadays, many companies have spent plenty of resources to create and maintain their websites. However, a large amount of investments on the web does not necessarily mean more business in return (Hernández, Jiménez, & Martín, 2009). Most Internet users could only memorize few popular web addresses (Uniform Resource Locator, URL). They simply search their desired information by typing the keywords and locate related result from search engine results. In 2001, 28 million websites were available; whereas a decade later, it has ten-fold increase to 298 million websites (Netcraft, 2011).

Apparently, the number of websites is growing in million sites each month. Researchers found that over 80% of web surfers use search engines to locate their desired information (Fox, 2002, Lawrence & Giles, 2000), and 78% of Internet users search online product information before they buy (Pew, 2011). There has been evidence indicating an increasing number of tourism enterprises using the Internet as a marketing tool, and more travel planning being conducted via the Internet (Buhalis & Law, 2008, Pan & Fesenmaier, 2006). As accommodation is one of the most important components for travellers, hotel websites will be a main search target for travellers, and for this reason, hotel websites are examined in this study.

Currently, the most popular international search engines were Google, Yahoo, Bing, Ask, and AOL (Hitwise, 2011). Making a website visible to search engines is one of the critical success factors on hotel marketing. However, most hotels only focus on maintaining their websites' content up-to-date and they are not aware of the visibility and freshness of the information of their websites on search engines. Furthermore, the indexing power of search engine crawlers and robots was enhanced significantly in the past decade. As such, many web contents incorporated Flash, and PDF can now be searched by visitors via search engines. However, web masters may not want search engine spiders to collect the whole web server's content. Therefore they should setup instructions that bound spiders fetching the specific web contents.

In view of the arising challenge, this research study is:

- to analyze the web information richness in search engine databases of various hotel categories;
- to compare the crawling result of the top global and domestic search engines; and
- to examine the search engine index ratio among hotel categories.

2 Research Background

2.1 Information on the Internet

The Internet is one of the most cost effective and efficient information and marketing channels that has no geographical and time barriers (Kasavana, Knutson, & Polonowski, 1997, Schmidt, Cantallops, & dos Santos, 2008, Tjostheim, Tussyadiah, & Hoem, 2007). According to Peterson, Balasubramanian, and Bronnenberg (1997), the Internet provides the capability of inexpensively storing a vast amount of information in different virtual locations. With the growing importance of search engines, the information quality that search engines provide should match customers' need. Prior studies have confirmed good quality of information provided by search engines must be up-to-date, available and visible, complete, popular, and with an adequate amount of information (Klein, 2002, Knight & Burn, 2005, Zhu & Gauch, 2000). Due to the huge volume of web pages being updated on each day, search engines delete the removed page information in around 10 and 26 days for Google and Yahoo (Lewandowski, 2008). However, if a page is out-dated but remains on the web server, search engines will retain the records in their databases. Furthermore, the popularity could affect influence the search engine ranking (Alexa, 2011).To increase the popularity of a website, more external links pointing to the URL are always associated with a larger amount of drive in traffic. Prior studies have indicated the importance of search engine marketing (Paraskevas, Katsogridakis, Law, & Buhalis, 2011, Xiang & Pan, 2009). However, the prerequisite of search engine marketing is that web contents must be stored in search engine databases; therefore the content indexing should not be overlooked.

2.2 Search Engine Indexing

Search engines collect web contents by sending out crawlers and robots (commonly known as spiders), which are the software agents that search engines employed to collect content for their databases (Kobayashi & Takeda, 2000). The basic web components for displaying a web page on the web browser included Hypertext

Markup Language (HTML) / Extended Markup Language (XML) codes, photos, animated graphics, and videos. In addition, to better control the web content, web programming codes such as JavaScript, Flash, PHP, ASP, CGI are included (Lecky-Thompson, 2008). With the increasing amount of broadband penetration, web surfers can enjoy high quality video and audio entertainments. In order to create more interactive and animated websites, Flash becomes one of the most popular web development tools. In 2001, many web contents generated by Flash, ZIP, and PDF were unable to be indexed by search engines. As such, it was not recommended that web designers to use these tools for web design (Sherman & Price, 2001). With the advancement of the search engines' crawling power, most spiders are now able to retrieve information from these file formats so web designers could have more choices to make attractive websites without worrying about search engine indexing and ranking issues.

To ensure a hotel website is listed by major search engines, a simple way is to register in a search engine's directory. All search engines encourage web masters to register their URLs onto their databases so their spiders can visit and index the content (Baidu, 2011; Bing, 2011a; Google, 2011a; Yahoo, 2011a). No search engine guarantees all pages are indexed, and they limit crawling frequency and the depth so the new websites will remain "invisible" until the next update crawl (Wouters, Reddy, & Aguillo, 2006). To increase crawling depth and indexing coverage, search engine engineers suggest web masters to submit sitemaps (Bing, 2011b, Google, 2011b), and enhancing web design style (Google, 2011b, Yahoo, 2011b). This is to ensure that their spiders could collect more accurate information and ensure all pages could be crawled without any missing pages. However, not all the pages inside a web server should be visible to visitors. As a result, the robot exclusion protocol was introduced in 1994 for controlling the access right of web crawlers and robots (Zittrain, 2008). Findings of a prior study discovered that 8% of the web content crawled by the search engine spiders were outdated or inactive (Al-Masri & Mahmoud, 2008) but the study did not include those contents that should not be visible to visitors such as system messages and software application program files. This hints some web masters did not manage their website contents well. Therefore, this study is to examine the current situation of content management of Hong Kong hotel web masters.

3 Research Methodology

This study examined the index ratio of Hong Kong hotel websites by counting the number of hotel pages and the pages indexed by search engines. The index ratio will be further explained in Section 4.3. In total, there were 150 hotels listed on Hong Kong Tourism Board's website (HKTB, 2011) and their names and URLs were retrieved. Two sets of data were collected in July 2011 from the Internet. The first batch of data was collected via sitemap generation software. By providing the hotel URL, the software scanned the whole tree structure from the home page within the domain and generated the sitemap of the hotel website. Because the programming languages adopted by each hotel website were different (e.g. HTML, PHP, ASP, and Flash), and the crawling methodologies of the sitemap generation software were different, this study adopted two sitemap generation applications. The sitemaps created by both sitemap generators were manually reviewed by the authors. After that,

the number of information pages and number of languages available were recorded into the database. Several chain hotel websites shared the information among hotels. So it is impossible to separate the pages by individual hotels. Therefore, these websites were excluded in this study, resulting in 127 hotel websites included for data analysis. The second batch of data was collected from search engines. According to Hitwise (2011), the top three search engines by search volume were Google (65.45%), Yahoo (15.66%) and Bing (12.97%). With around 60% of the tourists visiting Hong Kong were from mainland China (hereafter known as China), the website ranking in China's search engines could thus affect a hotel's business performance. In China, the leading search engine Baidu ranked six in Alexa (2011) which took up 65% of the search volume. As a result, these four search engines were included in this study. The command "site:*domain name*" was used to list out the pages crawled by the search engines. The total numbers of pages of each search engine were also recorded in the database.

4 Findings and Discussions

The hotel websites analyzed in this study were classified into four categories. An international chain refers to a hotel chain which has sister hotels in multiple continents. An Asia-based hotel chain refers to the chain that manages hotels within Asia. A local chain relates to the hotel network in Hong Kong, and independent refers to those that are not affiliated with any hotel networks. In this study, out of 127 hotels, 22 were affiliated with 10 international chains, 15 were affiliated with six Asia-based chains, 48 were part of 17 local chains, and 42 hotels were operated independently.

4.1 Content Pages on the Hotel Websites

In this study, a majority of the hotel websites provided content with more than one language, but six websites only had English pages. Among these six hotels, four of them were independent hotels and two of them were international chain hotels. Many chain hotels had more than one website. One of the websites was hosted by the chain, in which the domain name is the name of the chain. The other website's domain is the name of the hotel, which is normally owned by the hotel and maintained by hotel staff. Table 1 illustrates the average number of web pages in each category. The result indicated that the international chain, on average, offered 6.45 languages which were significantly more than the other three categories. For local chain and independent hotels, they normally offered English and Chinese (traditional and/or simplified) versions only. Asia-based chain websites additionally offered Japanese and Korean pages. For international chain hotels, they further provided pages in Spanish, French, German, and/or Portugal.

When the total number of web pages was examined, international chain hotels provided 370 pages and Asia-based chain hotels offered 380 pages. Both of them were two and three times more than local chain and independent hotels. Apparently, the number of available languages directly affects the total number of web pages. When the total number of pages on a hotel website is divided by number of available languages, the result shows a different view. The average number of pages per language for international chain hotels was 64 pages and the corresponding number for Asia-based chain hotels was 95 pages. Local chain hotels and independent hotels

were quite similar with 39 and 33 pages per language. As a result, the amount of web content information that Asia-based chain hotels provided was three times richer than local chain and independent hotels, and they were 30% richer than international chain hotels.

Table 1. Number of Hotel Web Pages by Hotel Categories

	International Chain	Asia-based Chain	Local Chain	Independent Hotel	Total	χ^2	Sig.
N	22 (17.32%)	15 (11.81%)	48 (37.80%)	42 (33.07%)	127 (100%)	17.18	.001*
No. of chains	10	6	17	-			
No. of Languages Provided						F	Sig.
Mean	6.45[1,2,3]	4.20[1]	3.38[2]	2.68[3]	3.80	18.31	.000*
Std.	3.83	1.86	1.34	0.93	2.37		
Total No. of Pages							
Mean	369.05[4,5]	380.13[6,7]	148.76[4,6]	93.59[5,7]	197.98	8.19	.000*
Std.	533.61	242.86	147.45	107.39	285.16		
Average No. of Pages per Language							
Mean	64.30	94.87[8,9]	39.60[8]	32.64[9]	48.44	8.06	.000*
Std.	81.08	42.76	30.86	34.89	50.07		

[1, 4, 6] *The mean difference for these two sectors is significant at the 0.05 level: p =0.005, 0.010, and 0.020*

[2, 3, 5, 7-9] *The mean difference for these two sectors is significant at the 0.05 level: p =0.000*

** Significant at p<0.05*

4.2 Pages Indexed by Search Engines

The crawling power of Google's spider seems stronger than the other three. In this study, among all hotel categories, Google indexed almost twice the number of pages than other search engines. Asia-based chain ranked the first with 288 pages indexed, local chain ranked second with 192 pages, and international chain and independent hotels had 171 and 151 pages crawled by Google's spider. The indexing power for both Yahoo and Bing were quite similar (Table 2). These three search engines indexed all the hotels listed on HKTB's website. The domestic search engine, Baidu, indexed the least number of hotel web pages and was unable to index all listed hotels. Three local chain and six independent hotel websites were not found in Baidu's database. For international chain hotels, Baidu listed out all the web pages by the domain name.

However, it was unable to filter out the pages that belong to designated hotels. Therefore, the total number of pages was unable to retrieve these four hotels. Among all search engines, the numbers of pages indexed for Asia-based chain and local chain were higher than international chain hotels. For Baidu, the total number of pages crawled for all hotel categories was 30% less than the Google's record. Particularly for international hotels, Baidu only indexed 15% of Google's contents.

Table 2. Number of Pages Crawled by Search Engines by Categories

	Google	Yahoo	Bing	Baidu	Total	F	Sig.
International Chain							
N	22	22	22	17		10.396	.000*
Mean	171.05[1,2,3]	90.14[1]	65.19[2]	26.07[3]	93.83		
Std.	137.29	64.09	44.68	24.20	97.29		
Min	33	25	5	0			
Max	503	331	156	75			
Asia-based Chain							
N	15	15	15	15		17.391	.000*
Mean	288.53[4,5,6]	138.73[4]	153.80[5]	90.64[6]	169.24		
Std.	97.19	81.04	81.28	37.23	106.04		
Min	12	8	8	5			
Max	395	835	353	149			
Local Chain							
N	48	48	48	45		12.900	.000*
Mean	191.78[7,8,9]	80.44[7]	81.42[8]	38.40[9]	99.02		
Std.	212.99	76.46	75.38	34.58	133.21		
Min	12	2	2	1			
Max	897	327	324	158			
Independent Hotel							
N	42	42	42	37		6.137	.001*
Mean	151.39[10,11,12]	64.15[10]	67.98[11]	45.78[12]	83.47		
Std.	208.00	68.29	72.91	54.75	125.79		
Min	1	1	1	1			
Max	867	362	361	223			

[1-9] *The mean difference for these two sectors is significant at the 0.05 level: p =0.000*
[10-12] *The mean difference for these two sectors is significant at the 0.05 level: p =0.007, 0.011, and 0.001*
** Significant at p<0.05*

4.3 Index Ratio

This study further examined the search engine index ratio. This ratio was calculated by the following formula:

$$\textit{Index Ratio} = \frac{\textit{Total number of pages indexed by a search engine}}{\textit{Total number of content pages on a hotel website}} \qquad (1)$$

When the ratio is less than one, that means a hotel website is not fully indexed by the search engine. If the ratio is greater than one, which implies the search engine over indexed some web contents such as outdated promotional materials or system messages which should be invisible to visitors. Fig 1 shows a contrasting index ratio pattern for Google and Baidu. There were a large number of the hotel websites (90%) which Google has indexed at least 50% of the web pages. However, for Baidu, less than 30% of the hotel websites has over 50% index ratio. For both Yahoo and Bing, their index ratios were quite similar with around 60% of the websites received at least 50% index ratio. However, the over index ratio for Google was also very high. Over 60% of the hotels received over 100% indexing, and 21% of them received over 200% indexing. For Yahoo and Bing, only 23% of hotel websites were over indexed. For Baidu, only 7% was over indexed. This hints that visitors could have higher opportunities to obtain dated information from Google search result or could easily access some system function pages which could cause system errors.

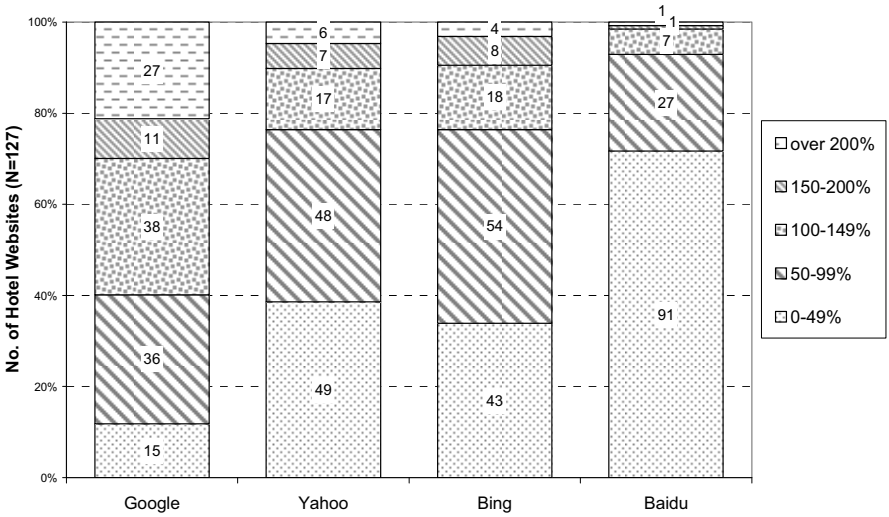


Fig 1. Frequency Diagram of Hotel Websites Search Engine Index Ratio

When the result was further divided into four hotel categories, more interesting findings were received. Although the average total number of web pages for independent hotels was only 94, Google has crawled an average of 208 pages. This gave independent hotels the highest index ratio with an average of 176% indexed by Google and 51% indexed by Baidu. In this study, Google even indexed one independent hotel with 10 times of its total number of web pages. The excess pages were mainly dated events information in 2010 and 235 pages of online reservation templates. Local chain hotels also have a high index ratio with 42% for Baidu to 170% for Google.

After examining the content of these sitemaps, several websites' application function or system message pages were indexed by Google, and plenty of the unlinked content was outdated and inactive web content was stored on the web server. If the web

master did not control the search engine crawling behavior, the search engine spiders could crawl the whole web server and list out unwanted information to visitors.

Table 3. Hotel Website Index Ratio by Categories

	Google	Yahoo	Bing	Baidu	Total	F	Sig.
International Chain							
N	22	22	22	17		5.46	.000*
Mean %	100.43[1,2]	66.10	54.04[1]	25.34[2]	64.73		
Std.	72.65	61.65	48.48	28.70	61.93		
Min	5.52	3.02	1.13	0.00			
Max	257.00	269.12	151.28	76.92			
Asia-based Chain							
N	15	15	15	15		5.16	.000*
Mean %	92.70[3,4]	45.37[3]	50.63	31.24[4]	55.39		
Std.	63.49	37.74	38.33	31.35	49.24		
Min	13.33	5.14	37.55	11.89			
Max	280.51	159.32	159.32	126.27			
Local Chain							
N	48	48	48	45		20.29	.000*
Mean %	170.49[5,6,7]	77.87[5]	76.44[6]	41.53[7]	92.43		
Std.	141.91	53.49	42.48	37.31	93.92		
Min	14.83	6.51	3.21	0.66			
Max	715.63	287.50	177.45	136.17			
Independent Hotel							
N	42	42	42	37		9.12	.000*
Mean %	175.91[8,9,10]	87.84[8]	90.65[9]	50.74[10]	102.87		
Std.	188.27	68.12	68.02	51.05	118.31		
Min	6.25	3.48	3.48	0.00			
Max	1070.37	324.00	324.00	275.31			

[5-7, 10] *The mean difference for these two sectors is significant at the 0.05 level: p =0.000*
[1-4, 8, 9] *The mean difference for these two sectors is significant at the 0.05 level: p =0.047, 0.001, 0.027, and 0.003*
**Significant at p<0.05*

5 Conclusions

This study examined the Hong Kong hotel website information richness and the page index ratio among search engines. The results indicated Asia-based chain hotels provided richer content than international chain hotels. Several international chain hotels not only disseminate information via their chain websites but also via the domain under hotel names. The content in hotels' own domain is richer than the corresponding one on the chain's website but the number of available languages was far less than the chain websites. Two of them only provided English pages. The average number of languages available on international chain hotel websites was 6.45 languages. The other hotel categories mainly provided English, together with traditional and simplified Chinese. Google's spiders got the highest crawling ability which enables the majority of hotel web pages indexed and become visible to Google

visitors. The crawling power of Baidu was the weakest. On international chain hotel websites, the pages indexed was only 15% of that from Google's. Though the indexing and crawling power of Google was the highest among all search engines, around 60% of the hotel websites were over indexed. This could make visitors locate the dated information or accidentally access part of the system which might affect the database accuracy. When comparing the index ratio among hotel categories, the local chain and independent hotels got the highest index ratio. Such a high ratio could be caused by dated web pages resided on the web server and the application program files. If the web master of these two categories did not manage their web content properly, the dated content and program files could be crawled by search engines. Finally, international chain hotels performed best in link popularity. This could help the hotel websites improve search engine ranking and driving in business (Alexa, 2011).

Hotel web masters should not only focus on the website indexing of international search engines, the performance of domestic search engines is also important. For instance, over 60% of tourists to Hong Kong were from China and over 65% of Chinese use Baidu as their primary search engine. As such, web masters should spend more effort on improving the indexing and ranking position of websites on Baidu. To prevent search engines from over indexing, it is recommended that web masters should implement the robots exclusion protocol to control search engines.

In case web masters prefer to keep the dated content for future reference, they should relocate the pages to separate folders. To prevent web spiders from crawling these specific folders, they must place a text file call robot.txt in the root of the website hierarchy. This file provides instructions allowing or limiting the search engine spiders fetching specific folders' content. If this file does not exist, the spiders will assume there are no restrictions on spider crawling of the web server.

6 Limitations and Future Research

This study has several limitations. First, the website hierarchy of international chain hotel websites was huge, and for this reason, the sitemap generation software was unable to fetch the complete web hierarchy in each case. Second, some websites used server-side programming for information retrieval but the sitemap generation software did not function properly so that pages using ASP or ASPX might be omitted. Furthermore, this study follows the hotel URL listed on HKTB's website. However, several hotels, especially chain hotels, had more than one domain name. Some of the hotels listed the chain URL but some listed the hotel domain.

Therefore, the results were not standardized. Finally, the tree structure of the website information arrangement could affect the fetch result. Some websites use languages as the name of the sub-folder where some websites use program variables to customize language settings. This could affect the search engine spider's performance. As mentioned by the top two search engines, enhancing the web design style could affect the performance of indexing and ranking (Google, 2011b, Yahoo, 2011b). Therefore, future research can further examine the relationship between index ratio and matching of search engines' preferred design style.

References

Alexa. (2011). *Alexa Top 500 Global Sites*. Retrieved August 26, 2011, from http://www.alexa.com/topsites

Al-Masri, E. & Mahmoud, Q. H. (2008). Investigating web services on the world wide web. In *Proceeding of the 17th International Conference on World Wide Web - WWW'08* (pp. 795-804). Presented at the 17th International Conference on World Wide Web, Beijing, China.

Baidu. (2011). *Baidu Search Help Center-Site Registration*. Retrieved August 29, 2011, from http://www.baidu.com/search/url_submit.html

Bing. (2011a). *URL Submission - Bing*. Retrieved August 28, 2011, from http://www.bing.com/webmaster/SubmitSitePage.aspx

Bing. (2011b). *Getting started checklist*. Retrieved August 29, 2011, from http://onlinehelp.microsoft.com/en-us/bing/hh204497.aspx

Buhalis, D. & 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* 29(4): 609-623.

Fox, S. (2002). *Pew Internet & American Life Project*. *Search Engines*. Retrieved September 25, 2009, from http://www.pewinternet.org/Reports/2002/Search-Engines.aspx

Google. (2011a). *Add your URL to Google*. Retrieved August 29, 2011, from http://www.google.com/addurl/

Google. (2011b). *Search Engine Optimization (SEO) - Webmaster Tools Help*. Retrieved August 29, 2011, from http://www.google.com/support/webmasters/bin/answer.py?answer=35291

Hitwise. (2011). *Search Engine Analysis*. *Hitwise United States - the power of competitive intelligence*. Retrieved August 26, 2011, from http://www.hitwise.com/us/datacenter/main/dashboard-23984.html

HKTB. (2011). *Search Hong Kong Hotels - Hong Kong Tourism Board*. Retrieved August 26, 2011, from http://www.discoverhongkong.com/eng/jsp/hotel/search-result.jsp

Kasavana, M. L., Knutson, B. J. & Polonowski, S. J. (1997). Netlurking: The Future of Hospitality Internet Marketing. *Journal of Hospitality & Leisure Marketing 5*: 31–44.

Klein, B. (2002). When Do Users Detect Information Quality Problems on the World Wide Web? In *Proceedings of the 8th Americas Conference on Information Systems (AMCIS 2002)*. Dallas, Texas.

Knight, S. & Burn, J. (2005). Developing a Framework for Assessing Information Quality on the World Wide Web. *Informing Science 8*: 159-172.

Lawrence, S. & Giles, C. L. (2000). Accessibility of Information on the Web. *Intelligence* 11(1): 32-39.

Lewandowski, D. (2008). A three-year study on the freshness of web search engine databases. *Journal of Information Science* 34(6): 817 -831.

Murphy, H. C. & Kielgast, C. D. (2008). Do small and medium-sized hotels exploit search engine marketing? *International Journal of Contemporary Hospitality Management 20*(1): 90-97.

Netcraft. (2011). *March 2011 Web Server Survey | Netcraft*. Retrieved April 4, 2011, from http://news.netcraft.com/archives/2011/03/09/march-2011-web-server-survey.html

Pan, B. & Fesenmaier, D. R. (2006). Online Information Search: Vacation Planning Process. *Annals of Tourism Research* 33(3): 809-832.

Paraskevas, A., Katsogridakis, I., Law, R. & Buhalis, D. (2011). Search Engine Marketing: Transforming Search Engines into Hotel Distribution Channels. Cornell Hospitality Quarterly 52(2): 200 -208.

Peterson, R., Balasubramanian, S. & Bronnenberg, B. (1997). Exploring the Implications of the Internet for Consumer Marketing. *Journal of the Academy of Marketing Science 25*(4): 329-346.

Pew. (2011, May). *What Internet Users Do Online | Pew Research Center's Internet & American Life Project*. Retrieved August 27, 2011, from http://www.pewinternet.org/Trend-Data/Online-Activites-Total.aspx

Rosen, D. E. & Purinton, E. (2004). Website Design: Viewing the Web as a Cognitive Landscape. *Journal of Business Research 57*(7): 787-794.

Schmidt, S., Cantallops, A. S. & dos Santos, C. P. (2008). The Characteristics of Hotel Websites and Their Implications for Website Effectiveness. *International Journal of Hospitality Management 27*(4): 504-516.

Sherman, C. & Price, G. (2001). *The invisible Web: uncovering information sources search engines can't see*. Information Today, Inc.

Tjostheim, I., Tussyadiah, I. P. & Hoem, S. O. (2007). Combination of Information Sources in Travel Planning a Cross-national Study. In *Information and Communication Technologies in Tourism 2007*. Wien: Springer-Verlag.

Wouters, P., Reddy, C. & Aguillo, I. (2006). On the visibility of information on the Web: an exploratory experimental approach. *Research Evaluation 15*(2): 107-115.

Xiang, Z. & Pan, B. (2009). Travel Queries on Cities in the United States: Implications for Search Engine Marketing in Tourism. In *Information and Communication Technologies in Tourism 2009*. Amsterdam, The Netherlands: Spring-Verlag.

Yahoo. (2011a). *Yahoo! Directory Suggest a Site*. Retrieved August 29, 2011, from http://add.yahoo.com/fast/add

Yahoo. (2011b). *SEO basics | Yahoo! Style Guide*. Retrieved August 29, 2011, from http://styleguide.yahoo.com/resources/optimize-search-engines/seo-basics

Zhu, X. & Gauch, S. (2000). Incorporating quality metrics in centralized/distributed information retrieval on the World Wide Web. In *Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval - SIGIR '00*, Athens, Greece.

Drivers of E-commerce Adoption in Egyptian Travel Agents

Mohamed Abou-Shouk[a], and
Wai Mun Lim[b]

[a] Fayoum University, Egypt and Plymouth University, UK.
mohamed.aboushouk@plymouth.ac.uk

[b] Plymouth University, UK.
wai.mun.lim@plymouth.ac.uk

Abstract

Benefits gained from e-commerce adoption, drivers pushing agents to adopt it, and inhibitors hindering the adoption are examples of factors positively or negatively affecting e-commerce adoption. Drivers of adoption could be internal or external pressures on travel agents to adopt technology in order to support their future survival in the travel and tourism global market. Mixed method approach is used in this study to investigate the drivers of e-commerce adoption in the Egyptian travel agents. Findings revealed that adapting to technology changes is the strongest driver of to adopt e-commerce by travel agents.

Keywords: e-commerce; drivers; travel agencies; Egypt; developing countries

1 Introduction

Environmental factors as cited in literature (Kuan & Chau, 2001; Premkumar & Roberts, 1999; Thong, 1999; Wiertz, 2001) are among those factors significantly affecting the adoption of technologies in small and medium-sized enterprises (SMEs). Environmental factors refer to the pressures that come from the environment surrounding the business (Kuan & Chau, 2001).

Additionally, environmental factors have been included among other factors in literature based models which investigate the factors affecting the adoption of information and communication technologies (ICTs) in SMEs. After reviewing literature models and articles relating to the factors affecting ICTs adoption in SMEs, it is found that environmental factors have been cited among the significant factors affecting the adoption decision.

According literature, environmental factors could be external pressures (Premkumar & Roberts, 1999; Voges & Pulakanam, 2011), industry pressures (Andreu, Aldas, Bigne, & Mattila, 2010; Grandon & Pearson, 2004), pressures of trading partners (Lacovou, Benbasat, & Dexter, 1995), employees' pressures (Mehrtens, Cragg, & Mills, 2001), and customers' pressures (Andreu, et al., 2010; Wiertz, 2001). It is worth to mention that travel agents are typically classified as SMEs (Gammack, Molinar, Chu, & Chanpayom, 2004; Karanasios, 2008; Liu & Arnett, 2000; Standing, Borbely, & Vasudavan, 1999). Furthermore, the word 'drivers' is used in this study to refer to the the forces which push travel agents to adopt e-commerce in order to enhance their competitive and survival positions in the global travel and tourism market.

The study contributes to the existing body of knowledge by investigating drivers of adoption in Egypt as a developing country, where only few studies have examined these factors in developing countries (Thomas, Shaw, & Page, 2011). It also supports the viewpoint that generalizing the findings of studies conducted in developed countries to developing countries lacks rigor without empirical evidence (MacGregor & Kartiwi, 2010). Additionally, the study investigates these drivers of adoption in the travel sector, and particularly for travel agents, which have not been investigated and documented thoroughly up until now (Hung, Yang, Yang, & Chuang, 2011).

2 Literature Review

Most of literature models and articles have cited drivers of adoption as a broad concept, such as environmental factors or environmental characteristics. Therefore, the literature review of this study lists theses drivers of adoption cited in literature studies. As a result, little information can be found in the academic literature on drivers to adopt e-commerce in SMEs, while more information can be found on broader aspects, like Internet and ICTs, and Internet and e-commerce which are complex and interrelated disciplines (Simpson & Docherty, 2004). Additionally, the drivers relating to ICTs, Internet and information systems are also applicable to e-commerce adoption.

Furthermore, there is a wide agreement among researchers that drivers of ICTs adoption among SMEs are seen as external pressures (Voges & Pulakanam, 2011) derived from customers, suppliers, competitors and business partners (Beekhuyzen, Hellens, & Siedle, 2005; Poon & Joseph, 2001; Simpson & Docherty, 2004). Drivers of adoption have been summarized in table 1.

Table 1. Review of drivers of technology adoption in SMEs

Drivers	**Researchers**
Continuous demand for improved product/service quality	(Poon & Joseph, 2001)
Existing customers and their pressures power to adopt new technologies	(Andreu, et al., 2010; Beckinsale & Levy, 2004; Bigne, Aldas, & Anderu, 2008; Daniel, Wilson, & Myers, 2002; Dyerson & Harindranath, 2007)
External pressures from new customers	(Simpson & Docherty, 2004)
Social pressure of customers	(Vrana & Zafiropoulos, 2006)
Market demand	(Buhalis & Deimezi, 2004; Hung, et al., 2011)
Competitive pressures	(Beekhuyzen, et al., 2005; Hung, et al., 2011; Poon & Joseph, 2001)
The fear of being behind and level of competition among the industry firms	(Bigne, et al., 2008; Patricia, 2008; Teo, Lin, & Lai, 2009; Wesrthner & Klein, 1999)
Responding to competitors	(Simpson & Docherty, 2004)
Threats of competitive forces	(Jin, 2007)
Emergence of new virtual intermediaries	(Barnett & Standing, 2001)

Cont. Table 1

Fear of innovative competitors	(Buhalis & Deimezi, 2004)
pressures of business partners	(Nour, 2002; Raymond, 2001; Simpson & Docherty, 2004; Teo, et al., 2009)
Suppliers' development programmes	(Quayle, 2002)
Suppliers upgrading to newer technologies	(Vrana & Zafiropoulos, 2006)
Demands from business partners	(Rao, Metts, & Monge, 2003)
Rapidly changing business environment	(Grandon & Pearson, 2004; Kuan & Chau, 2001; Saffu & walker, 2008; Simpson & Docherty, 2004)
Environmental uncertainty	(Raymond, 2001)
Adapting to changes in technology	(Karagozoglu & Lindell, 2004; Law, leung, & Wong, 2004)
Globalization consequences	(Nour, 2002; Poon & Joseph, 2001)
Modernisation	(Buhalis & Deimezi, 2004)
Governmental rules	(Grandon & Pearson, 2004; Kuan & Chau, 2001; Saffu & walker, 2008; Teo, et al., 2009)
Business strategy	(Quayle, 2002)
Employees' pressure	(Beckinsale & Levy, 2004)
Owner/manager push	(Beckinsale & Levy, 2004)
Re-intermediation and future survival of travel agents	(Bennett & Lai, 2005; Heung, 2003; Stansfield & Grant, 2003; Warden & Tunzelana, 2004)

3 Research Gap and Objectives

Travel and tourism industry is increasingly moving online, therefore, travel and tourism businesses need to react accordingly, especially in developing countries (Migiro & Ocholla, 2005). Travel agents are one category of tourism businesses facing the threat of disintermediation in the global travel market. To survive in an increasingly competitive global environment, small businesses need to advent the Internet and achieve economies of scale (Kim, 2005). E-commerce creates opportunities for travel agents' re-intermediation in the market and supports their survival in the new competitive environment (Bennett & Lai, 2005; Hamed, 2003; Patricia, 2008). Drivers of e-commerce emerge not only because of supporting future survival, but as a result of external pressures from customers, suppliers, competitors, and even government regulations as well. This research aims to determine the significant drivers pushing travel agents to adopt e-commerce to support their future existence in Egypt as a developing country.

4 Research Methodology

Mixed method approach is used in this study to identify the drivers of e-commerce adoption by travel agents. The sequential explanatory design strategy grounded on the pragmatism paradigm is employed, starting with the quantitative stage and then moving to the qualitative stage that helps in interpreting the quantitative findings (Creswell, 2009). A questionnaire form was designed for quantitative data collection purposes. Eighteen drivers have been included in the initial form; each driver takes

the value of (5) if strongly agreed, ranging to the value of (1) if strongly disagreed by managers of travel agents. The form has been piloted by 50 managers of travel agents in order to determine the significant drivers of adoption by travel agents. Corrected item-total correlation statistics were used to retain factors with loadings among 0.35 and 0.80 in the final form according to Netemeyer, Bearden, and Sharma (2003). In doing so, a number of 6 out of 18 drivers have been considered in the final questionnaire (table 2). Mean statistics were used to count the agreed drivers. Subsequently, reliability statistics were calculated.

Table 2. Measurement scale

Items (*=Retained Items)	**Corrected Item-Total Correlation**
Consumer demand for enhanced service quality	.271
Increased pressures from new and/or existing customers	.308
Higher level of buyer-seller interaction	.313
Responding to competitor pressures*	.530
Fear of being left behind	.076
Emergence of new virtual intermediaries	.242
Supplier's development programmes*	.506
Business partner influence*	.410
Growing markets changes and demands	.188
Business environmental uncertainty	.227
Adapting to technology changes*	.577
Rapidly industry changes	.285
Globalization consequences*	.497
Governmental rules and regulations	.252
Business strategy (planning to expand and development)	.324
Pressure from employees	.336
Owner/manager and/or IT manager push	.155
Future survival of travel agency*	.590

Logistic regression is used in this study to identify the significant drivers where the outcome (i.e. dependent) variable is binary: adopter travel agents of e-commerce (coded as 1) versus non-adopters (0). The five-point Likert scale of drivers has been recoded to include three categories instead of five. Scales of 'Strongly Agree' and 'Agree' have been recoded to take the value of 1 instead of 5 and 4 respectively. 'Neither Agree nor Disagree' is recoded to 2 instead of 3. Finally, 'Strongly Disagree and Disagree' scales have been recoded to 3 instead of 1 and 2 respectively. The latter category has been used as the baseline in the logistic regression model in this study.

Where the qualitative stage is used to assist in interpreting the results revealed from the quantitative stage, NVivio (ver. 9) was used ot code qualitative data developed

based on quantitative findings. Face to face semi-structured interviews were used to collect qualitative data. In 22 interviews with general managers of travel agents in Egypt, interviewees were provided with a copy of the interview schedule before the event, and clear open questions to reduce bias and to promote validity and reliability issues (Saunders, Thornhill, & Lewis, 2007)

5 Sample of the Study

Among the three categories of travel agents in Egypt, category (A) travel agents is selected as the sample frame in this study. Category (A) includes agencies organising inclusive packages and all other related tourism services for groups or/and individuals, inside and/or outside Egypt, and finally executing the planned packages of other tour operators (Ministry of Tourism, 2008).

There are a total of 1,023 agents in category (A) (Egyptian Travel Agents Association, 2008). With the high percentage of agents based in Greater Cairo (826 agents), it is selected to be the geographic sample frame in this study. Where the study questions have been addressed to adopters and non-adopters of travel agents category (A), therefore, the questionnaire was delivered and collected from two groups of agents' managers. Using the sample frame, it is found that 387 agents are adopters of e-commerce and 439 agents are non-adopters. Adopters are agents who have a static website (initial level of e-commerce) and/or an interactive website that supports an online booking system (mature level of e-commerce adoption). By contrast, non-adopters are agents who do not have websites.

As a result of non-equal representation of the two groups in the sample frame, a stratified sample is used to reduce heterogeneity of the population and to increase the estimates' efficiency. To calculate strata sample size, the following formula was used: stratum size = (n_i/N)*total sample size, where n_i is the stratum size and N is the total number of sample frame. For adopters = (387/826)*278=135 subjects. For non-adopters = (439/826)*287=152 subjects. Technique of simple random sampling with replacement was used within each of these strata to select sample subjects. Finally, 208 and 210 valid forms were collected from adopters and non-adopters, respectively.

6 Validity and Reliability

To ensure validity procedures, the questionnaire was designed and piloted to be ready for data collection purposes. The primary questionnaire form was designed in English language, then validated by an expert panel and 15 PhD students to check its face-validity. Where the respondents' mother language is Arabic and using the mixed techniques in translation, for validity (Saunders, Thornhill, & Lewis, 2009), the questionnaire form was translated into Arabic to check questions' comprehensiveness. Subsequently, it was translated back into English to compare the two English forms and to avoid misleading translation issues. The advantage of mixed techniques in translation is getting the best match between source and target questionnaires to create the final questionnaire form.

More precisely, to achieve the mixed techniques of translation, firstly the source questionnaire was translated from English into Arabic language by a certified

translation centre in Egypt. Sequentially, the Arabic version was translated back from Arabic into English by another independent translator working for a certified translation centre in the UK. The next step was comparing the two English questionnaires by a British lecturer specialized in linguistics. The final questionnaire shows 6 drivers with a Cronbach's alpha value of 0.79 for the 6 items included in the final form for travel agents which reflects its reliability.

7 Research Results

7.1 Drivers of E-Commerce Adoption in Travel Agents

From descriptive analyses and using mean scores of managers' opinions in travel agents (table 3) it becomes clear that agents' responses ranged between 3 and 4 except for the driver of 'adapting to technology changes'. Generally, responses lie between neutral and managers' agreement.

Table 3. Statistics of drivers of e-commerce adoption in travel agents

Drivers	Mean	Std Deviation.
Responding to competitors pressures	3.26	1.19
Supplier's development programmes	3.29	1.22
Business partner influence	3.09	1.20
Adapting to technology changes	2.79	1.16
Globalization consequences	3.15	1.26
Future survival of travel agency	3.04	1.22

7.2 Significant Drivers of Adoption According to Logistic Regression Model

A number of 418 cases were included in the analysis. Logistic regression and a -2 Log likelihood and Chi-square test model revealed the following results for overall goodness of fit: it is found that including predictors in the model reduced the -2likelihood by 340.606. Rsquare value for Nagelkerke is 58% and the value for Cox & Snell is 43.5%, suggesting the model offers a reasonable explanation of variance. Furthermore, Chi-square (χ^2) value for Hosmer and Lemeshow test is 14.051 (df.=7) and $p>0.05$. Therefore, the null hypothesis of no difference between observed and predicted values of dependent variable is accepted and the model appears to be a good fit to the data. To test the null hypothesis that all coefficients=0, it is revealed that χ^2 value of Omnibus tests of model coefficients is 238.855 (df.=10) and $p<0.01$, therefore the null hypothesis is rejected and the independent variables as a whole are significantly affecting the dependent variable.

Additionally to assess the effects of how many explanatory variables (the independent variables) have in the model over the constant only, it is useful to compare the classification table in step 1 (i.e. independent variables included in the model) with the classification table in step 0 (i.e. no independent variables included in the model). In step 0, the % value was 50.2% and in step 1, it is 81.8%. This means that the

inclusion of independent variables significantly improves the explanation of the model regarding the drivers pushing the adoption of e-commerce by travel agents.

Table 4 indicates the significance of coefficients. Using the values of Wald statistics for drivers with p values less than 0.05, the null hypothesis can be rejected that B=0 and, thus, those drivers appeare as significant variables affecting the adoption of e-commerce in travel agents. Generally, the significant influences of the six drivers on the adoption of e-commerce according to managers who agree to these drivers in comparison to others, who have neutral opinions, are shown in table 4.

Table 4. Logistic regression output (variables in the equation)

	B	**S.E.**	**Wald**	**df**	**Sig.**	**Exp(B)**
Competitor_pressure(1)	.729	.275	7.041	1	.008	2.073
Suppliers_programs(1)	1.065	.275	15.006	1	.000	2.901
Suppliers_programs(2)	-21.215	40192.970	.000	1	1.000	.000
Business_partner(1)	.596	.281	4.507	1	.034	1.814
Business_partner(2)	.971	1.140	.725	1	.394	2.641
Technology_changes(1)	1.932	.355	29.542	1	.000	6.901
Technology_changes(2)	-20.508	40192.970	.000	1	1.000	.000
Globalisation_consequences(1)	1.395	.299	21.814	1	.000	4.035
Future_survival(1)	.766	.305	6.325	1	.012	2.151
Future_survival(2)	-.870	2.341	.138	1	.710	.419
Constant	-2.686	.296	82.290	1	.000	.068

To describe the variables in the model using the Logit of e-commerce adoption form:

$Logit_{ecomadoption}$= -2.686+ 0.729 * managers agree to competitors' pressures + 1.065 * managers agree to suppliers' development programs + 0.596 * managers agree to business partner's pressures + 1.932 * managers agree to adapting to technology changes + 1.395 * managers agree to globalisation consequences + 0.766 * managers agree to future survival of travel agents.

To express the same model in odds ratios (Exp (B) instead of B values):

$Odds_{ecomadoption}$= 0.068 x 2.073* managers agree to competitors' pressures x 2.901* managers agree to suppliers' development programs x 1.814* managers agree to business partner's pressures x 6.901* managers agree to adapting to technology changes x 4.035* managers agree to globalisation consequences x 2.151* managers agree to future survival of travel agents.

To interpret the odds ratios, it can be as follows:

- Agents' managers who agree on competitors' pressure are 2.073 times more likely to adopt e-commerce than others who disagree on these pressures;
- Agents' managers who agree on suppliers' development programs are 2.901 times more likely to adopt e-commerce than others who disagree on these pressures;

- Agents' managers who agree on business partner's pressures are 1.814 times more likely to adopt e-commerce than others who disagree on these pressures;
- Agents' managers who agree on adapt to technology changes are 6.901 times more likely to adopt e-commerce than others who disagree;
- Agents' managers who agree on globalisation consequences are 4.035 times more likely to adopt e-commerce than others who disagree on these pressures; and
- Agents' managers who agree on future survival of travel agents are 2.151 times more likely to adopt e-commerce than others who disagree on these pressures.

To sum up, for managers who agree on those drivers of e-commerce adoption in travel agents, it clearly emerged that adapting to technology changes is the strongest driver affecting their decision to adopt e-commerce, followed by globalisation consequences, the supplier's development programs, then the future survival of travel agents in the global tourism market, the competitors' pressures, and finally, the business partners' pressures.

7.3 Drivers of Adoption: Qualitative Analysis

Figure 1 illustrates the responses of interviews with general managers in travel agents. Analysis is described according to the number of times items were referred to in the 22 interviews. Responding to competitors was the most commonly driver of technology adoption by the managers. 'Competition', 'continue competing', 'competitors' pressure', and 'international competition' were the words used by the managers to define the strongest driver behind adopting e-commerce. Comparison with competitors, competitive position, and not wishing to be left behind competitors are some of the explanations managers used to justify their adoption of e-commerce. Future survival was also identified. Managers perceive that the future survival of their agencies is empowered by e-commerce and internet technology adoption. Additionally, 'continue', 'staying', 'survive' and 'exist' [in the travel market] are some of the phrases used in defining the drivers that push these agents to adopt e-commerce.

Adapting to technology changes came next. The interviewed managers used words like 'revolution', 'trend', and 'fashion' to describe the age of internet technologies. 'There is an emergent need to adopt internet technologies' said one manager, describing their need to adopt e-commerce. Some of the managers perceived e-commerce as the fashion saying 'this is the trend in the world now'. Other managers expressed the view that the new generations as well as some current customers are more interested in using technology and the Internet. Customer pressure is, thus, perceived by the managers as one of the drivers of e-commerce adoption. 'Potential' and 'online' customer requirements are behind the emergent need to adopt e-commerce in their activities.

The benefits to be gained from adoption are another driver behind the adoption of e-commerce in the view point of some managers. They claim that increasing the awareness of benefits gained from adoption pushes agents to adopt these technologies. Some managers request the support of government bodies and organisations to raise the awareness of e-commerce and their benefits for SMEs.

Further drivers include governmental regulations imputed by the ministry of tourism and other related bodies, which force companies to use internet technologies for registrations and collecting information when dealing with the ministry. Another key factor is business strategy, which encompasses adopting technology, expanding into global markets, and developing and increasing their business size.

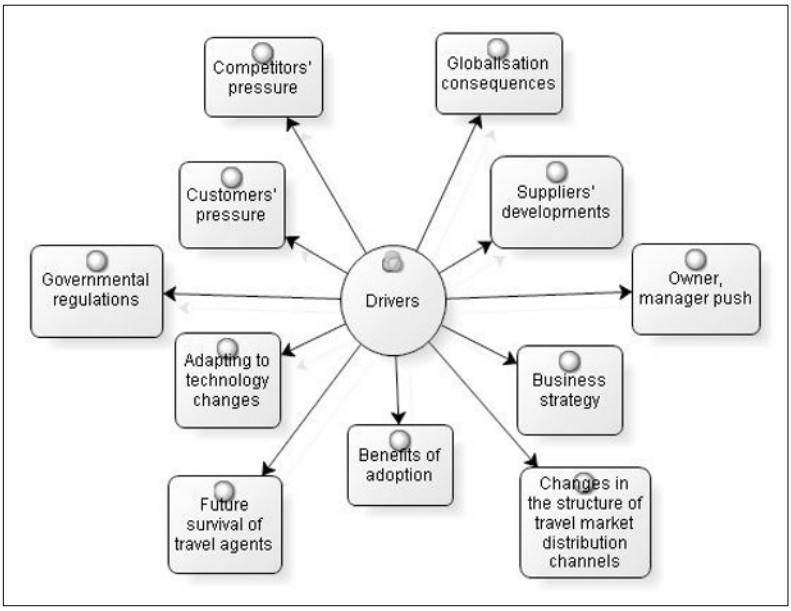


Fig. 1. NVivo output of the drivers of e-commerce adoption

Supplier's development programmes and upgrades to new technologies are another driver of adoption. This form of pressure from suppliers and partners stems from the need to manage problems of compatibility, and enhance business performance. For example, the travel supplier may choose a technology, such as Sabre or Amadeus, and then pushes the travel agent to adopt it as well. Some managers recognize the manager/owner's commitment to adopt technology as another driver of adoption. The consequences of globalisation, such as the possibility of international travel agents to open branches in various destinations, create another form of competition for travel agents, and are, thus, another driver of technology adoption. Finally, the changes in the travel market distribution structure and the appearance of virtual intermediaries are further pressures on travel agents to adopt e-commerce technologies.

8 Conclusion

This paper has highlighted the significant drivers of e-commerce adoption in the Egyptian travel agent branch. A total of 418 surveys were conducted in order to test a logistic regression model identifying the drivers pushing travel agents to adopt e-commerce. In addition, 22 interviews with managers in travel agents were taken. This study has met its aim identifying the significant drivers of e-commerce adoption in travel agents. It is found that agents' adaptation to technology changes is the strongest

driver of e-commerce adoption. The perception of globalisation consequences and its impact on the necessity to adopt technology is ranked as the second important driver in the agents' priorities. The development of suppliers' programs push agents to use the same programs in order to keep their relationships with suppliers. Future survival in the fierce competitive global markets occupies the fourth driver of technology adoption by travel agents. Finally, the competitors' pressure is the last driver of adoption. Qualitative results confirmed quantitative findings, however, adding new drivers, namely benefits of adoption, governmental regulations, and manager/owner's commitment to adopt technology.

These drivers could explain why agents are caring about adaptation to technology changes and developed programs of suppliers as a gateway to enhance their competitive positions and, in turn, supporting their future survival that leads to effectively re-intermediating themselves in the global travel and tourism market. These drivers also claim that agents are in need to adopt technology to fulfil their business defects. However, this leads to a question of why the majority of travel agents in Egypt are still non-adopters (more than 50% of agents do not even have websites) although they believe that adopting technology is enhancing their business operations. This suggests that drivers are not the only factors affecting the decision of technology adoption in travel agents. There might be inhibitors and barriers of adoption represented, for instance, by resource limitation, customers' readiness, suitability of technology to the nature of services or even the belief that technology adoption will not sufficiently enhance their competitive positions as expected.

9 Limitations and Suggestions for Further Research

When reviewing the results of this study, one limitation is the need to include the benefits and inhibitors of adoption to the research model to fully investigating factors affecting the adoption decision. Although managers found these drivers pushing agents to adopt e-commerce technologies, further research should be addressed to interpret why the majority of travel agents do not even have websites. Thus, future studies might focus on what are the barriers to adopt and how they affect travel agents' decisions to adopt e-commerce technologies. This will help both, agents and governments to take coordinated actions to maximize the benefits of adoption and to overcome existing adoption barriers.

References

Andreu, L., Aldas, J., Bigne, J. & Mattila, A. (2010). An analysis of e-business adoption and its impact on relational quality in travel agency–supplier relationships. *Tourism Management* 31: 777-787.

Barnett, M. & Standing, C. (2001). Repositioning travel agencies on the internet. *Journal of Vacation Marketing* 7(2): 143-152.

Beckinsale, M. & Levy, M. (2004). SM*Es and internet adoption strategy: who do SMEs listen to?* Retrieved 14 February, 2009, from http://is2.lse.ac.uk/asp/aspecis/20040016.pdf

Beekhuyzen, J., Hellens, L. & Siedle, M. (2005). *Cultural barriers in the adoption of emerging technologies.* Retrieved 26 February, 2009, from http://www.ucd.smartinternet.com.au/Documents/Cultural_Barriers.pdf

Bennett, M. & Lai, C. (2005). The impact of the internet on travel agencies in Taiwan. *Journal of Tourism and Hospitality Research* 6(1): 1-17.

Bigne, J., Aldas, J. & Anderu, L. (2008). B2B services: IT adoption in travel agency supply chains. *Journal of Services Marketing 22*(6): 454- 464.

Buhalis, D. & Deimezi, O. (2004). E-tourism developments in Greece: information communication technologies adoption for strategic management of the Greek tourism industry. *Tourism and Hospitality Research, 5*(2): 1467-3584.

Creswell, J. (2009). *Research design: qualitative, quantitative, and mixed methods approaches* (3rd ed.). London: SAGE Publication.

Daniel, E., Wilson, H. & Myers, A. (2002). Adoption of e-commerce by SMEs in the UK: towards a stage model. *International Small Business Journal 20*(3): 253-270.

Dyerson, R. & Harindranath, G. (2007, 5-9 August). *ICT adoption and use by SMEs in the UK: a survey of South East.* Paper presented at the Portland International Center for Management of Engineering and Technology (PICMET), Portland, Oregon - USA.

Gammack, J., Molinar, C., Chu, K. & Chanpayom, B. (2004). *Development needs of small to medium size tourism businesses: a report prepared for APEC workshop group*. APEC international Centre for Sustainable Tourism.

Grandon, E. & Pearson, J. (2004). Electronic commerce adoption: an empirical study of small and medium US businesses. *Information & Management 42*: 197-216.

Hamed, H. (2003). *E-commerce in tourism Sector.* Cairo: Faculty of Tourism and Hotels, Helwan University.

Heung, V. C. S. (2003). Barriers to implementing e-commerce in the travel industry: a practical perspective. *International Journal of Hospitality Management 22*(1): 111-118.

Hung, Y., Yang, Y., Yang, H. & Chuang, Y. (2011). Factors affecting the adoption of ecommerce for the tourism industry in Taiwan. *Asia Pacific Journal of Tourism Research 16*(1): 105-119.

Jin, N. (2007). *A study of information technology adoption for small and medium sized enterprises: strategic competitiveness*. Paper presented at the 2007 International Conference on Wireless Communications, Networking and Mobile Computing.

Karagozoglu, N. & Lindell, M. (2004). Electronic commerce strategy, operations, and performance in small and medium sized enterprises. *Journal of Small Business and Enterprise Development 11*(3): 290-301.

Karanasios, S. (2008). *An E-commerce Framework for Small Tourism Enterprises in Developing Countries.* Unpublished Doctorate thesis, School of Information Systems, Faculty of Business and Law, Victoria University.

Kim, C.-W. (2005). Enhancing the role of tourism SMEs in global value chain: a case analysis on travel agencies and tour operators in Korea. In *Global tourism growth: a challenge for SMEs* 6-7 September, Gwangju, KoreaKwangju, Korea. Retrieved July 20, 2008 from http://www.oecd.org/dataoecd/26/48/36886129.pdf

Kuan, K. & Chau, P. (2001). A perception-based model for EDI adoption in small businesses using technology-organization-environment framework. *Information & Management* 38: 507-521.

Lacovou, C., Benbasat, I. & Dexter, A. (1995). Electronic data interchange and small organizations: adoption and impact of technology. *MIS Ouarterly* 19(4): 465-485.

Law, R., leung, K. & Wong, J. (2004). The impact of the internet on travel agencies. *International Journal of Contemporary Hospitality Management* 16(2): 100-107.

Liu, C. & Arnett, K. (2000). Exploring the factors associated with Web site success in the context of electronic commerce *Information & Management* 38: 23-33.

MacGregor, R. & Kartiwi, M. (2010). Perception of barriers to e-commerce adoption in SMEs in a developed and developing Country: a comparison between Australia and Indonesia. *Journal of Electronic Commerce in Organizations* 8(1): 61-82.

Mehrtens, J., Cragg, P. & Mills, A. (2001). A model of internet adoption by SMEs. *Information and Management 39*: 165-176.

Migiro, S. & Ocholla, D. (2005). Information and Communication Technologies in Small and Medium Scale Tourism Enterprises in Durban, South Africa. *Information Development* 21(4): 283-294.

Netemeyer, R., Bearden, W. & Sharma, S. (2003). *Scaling procedures: issues and applications*. London: Sage Publications.

Nour, S. (2002). *ICT opportunities and challenges for development in the Arab World, Discussion paper*. Paper presented at the Paper prepared for the UNU/WIDER Conference on the New Economy in Development 10-11 May 2002.

Patricia, D. (2008). *The Internet, threat or tool for travel agencies?* Retrieved 5 February, 2009, from http://steconomice.uoradea.ro/anale/volume/2008/v2-economy-and-business-administration/017.pdf

Poon, S. & Joseph, M. (2001). A preliminary study of product nature and electronic commerce. *Marketing Intelligence & Planning 19*(7): 493-499.

Premkumar, G. & Roberts, M. (1999). Adoption of new information technologies in rural small businesses. *Omega, International Journal of Management Science 27*: 467-484.

Quayle, M. (2002). E-commerce: The challenge for UK SMEs in the twenty-first century. *International Journal of Operations & Production Management 22*(10): 1148-1161.

Rao, S., Metts, G. & Monge, C. (2003). Electronic commerce development in small and medium sized enterprises: a stage model and its implications. *Business Process management Journal 9*(1): 11-32.

Raymond, L. (2001). Determinants of web site implementation in small business. *Internet Research: Electronic Networking Applications and Policy 11*(5): 411-422.

Saffu, K. & Walker, J. (2008). Strategic value and electronic commerce adoption among small and medium-sized enterprises in a transactional economy. *Journal of Business & Industrial marketing 23*(6): 395-404.

Saunders, M., Thornhill, A. & Lewis, P. (2007). *Research methods for business students* (4th ed.). London: Financial Times Prentice Hall.

Saunders, M., Thornhill, A. & Lewis, P. (2009). *Research methods for business students* (5th ed.). London: Financial Times Prentice Hall.

Simpson, M. & Docherty, A. (2004). E-commerce adoption support and advice for UK SMEs. *Journal of Small Business and Enterprise Development 11*(13): 315-328.

Standing, C., Borbely, S. & Vasudavan, T. (1999). A study of Web diffusion in travel agencies. In *Thirty-second Annual Hawaii International Conference on System Sciences* (Vol. 5, pp. 1-9). Maui, Hawaii: HICSS.

Stansfield, M. & Grant, K. (2003). An investigation into issues influence the use of the internet and electronic commerce among small-medium sized enterprises. *Journal of Electronic Commerce Research 4*(1): 15-33.

Teo, T., Lin, S. & Lai, K. (2009). Adopters and non-adopters of e-procurement in Singapore: An empirical study. *Omega 37*(5): 972-987.

Thomas, R., Shaw, G. & Page, S. (2011). Understanding small firms in tourism: A perspective on research trends and challenges. *Tourism Management 32*: 963-976.

Thong, J. (1999). An integrated model of information systems adoption in small businesses. *Journal of Management Information Systems 15*(4): 187-214.

Voges, K. & Pulakanam, V. (2011). Enabling factors Influencing internet adoption by New Zealand small and medium size retail enterprises. *International Review of Business Research Papers 7*(1): 106-117.

Vrana, V. & Zafiropoulos, C. (2006). Tourism agents' attitudes on internet adoption: an analysis from Greece. *International Journal of Contemporary Hospitality Management 18*(7): 601-608.

Warden, S. & Tunzelana, S. (2004). E-Commerce: a critical review of SMME organisational barriers in tourism. In *Proceedings of 6th WWW Applications Conference 2004*, Rand Afrikaans University, Johannesburg, South Africa.

Wesrthner, H. & Klein, S. (1999). ICT and the changing landscape of global tourism distribution. *Electronic markets 9*(4): 256-262.

Wiertz, C. (2001). *The Internet adoption decision in small and medium enterprises (SMEs)*. Retrieved 12 February, 2009, from http://www.bonissen.de/ib/papers/Small%20Biz.pdf

A Follow Up of Internet Adopters' Use, Perceptions and Channel Preferences of Electronic Travel Services

Niklas Eriksson

Department of Business, Information Technology and Media
Arcada University of Applied Sciences, Finland
niklas.eriksson@arcada.fi

Abstract

For this study a consumer scenario was developed to get insights into Internet adopters' use, perceptions and channel preferences of electronic travel services. The findings are drawn upon empirical data collected in Finland with a self-administered Web-questionnaire in fall 2004 and a follow up in spring 2011. The results confirm a steady growth of electronic travel service use, but imply further developments of online services that support the last face of the investigated consumer scenario; reflection and evaluation. An overwhelming margin of the respondents prefers the electronic channel (computer-based or mobile device-based) over conventional interaction (e.g. calling and face to face interaction) for the studied consumer scenario. The results also confirm that the mobile channel should primarily be seen as a closely integrated supplement channel rather than a standalone channel only focusing on certain consumer activities.

Keywords: Internet; Travel; Mobile services; Consumer survey

1 Introduction

The use of the Internet for doing commerce or interacting with customers has been growing rapidly in the worldwide travel and tourism industry and dramatically changed the market conditions for organizations operating in the industry (Buhalis & Law, 2008). Mobile commerce, or e-commerce over mobile devices, has on the other hand had many conflicting predictions on its future popularity. For example, in the beginning of the m-commerce era there were several estimates that were overly optimistic (e.g. Durlacher Research, 2000). Later studies in Finland (e.g. Durlacher Research, 2005, Hyvönen & Repo, 2005, Ministry of Transport and Communications, 2005) showed that the mobile service market was growing steadily, but it seemed to be a sleeping giant (Waldén et al., 2007).

At that time the European market was, on the other hand, lagging behind when compared to countries like Japan and Korea. Mobile phones are now predicted to overtake PCs as the most common Web access device worldwide by 2013 (Gammage, 2010) and consumers' smartphone adoption will drive more activities usually associated with the PC (Husson & Ask, 2011). In this paper we will present results from an online consumer survey conducted in fall 2004 and a follow up conducted in spring 2011. The setup of the questionnaire has been the same for both studies. This lets directly compare results thereby showing changes in Internet adopters' use, perceptions and channel preferences of electronic travel services. The surveys have focused on the Finnish market, which has generally been seen as a technologically advanced market (at least outside Finland) when it comes to the populations' willingness and readiness to adopt new electronic services (Waldén et. al. 2007).

Internet adopters, which are the focus here, are an interesting group of consumers as they are likely to be the ones to adopt mobile services or the mobile Internet (Anckar & Dincau, 2002), at least in countries where the stationary Internet has a strong foothold (like Finland).

We have taken a multi-channel approach in order to investigate the use, perception and channel preference of electronic travel services among Internet adopters. Channel features have been subject for major research and managerial questions, especially the combination or the separation of the electronic channel and the conventional channel, also called the physical channel or off-line channel in the early Internet age (e.g. Simons et al. 2002; Schoenbachler & Gordon, 2002; Kaufman-Scarborough & Lindquist, 2002). Also within travel the existence or non-existence of conventional vendors such as the travel agent has been investigated (e.g. Anckar, 2003; Law et. al., 2004).

Due to the development of technology such as mobile communications the electronic channel has emerged into separate channels, all with their unique characteristics, and - presumably - a varying suitability for performing certain tasks. Online activities should in fact be studied in a multi-channel environment as consumers are likely to trade among various media (Kaufman-Scarborough & Lindquist, 2002). Therefore, we need to understand what is distinctive about the electronic channels in their own right, and look at where they clearly win over the conventional channel but also where they win over each other. Moreover, there is a need for longitudinal studies to understand the impact of Internet on travel services (Law et. al, 2004).

The paper is structured as follows: in section 2 we go through the study in 2004 and our main research objectives; in section 3 and 4 we present a consumer scenario and channels of interaction, which will be used to investigate our research questions; in section 5 we go through our data collection procedures; in section 6 - 8 we present our results and analysis; in section 9 we will draw some conclusions.

2 The Study in 2004 and Research Objectives

In 2004 empirical data was drawn upon the following main research question for a consumer scenario:

[RQ1] Use: to what extent are electronic travel services used through computers and/or mobile devices by Internet adopters?

[RQ2] Perceived benefit: as how beneficial are the electronic travel services perceived by Internet adopters?

[RQ3] Channel preferences: which channel of interaction with travel service providers is / would be mainly used by Internet adopters; 1. conventional, 2. computer-based, 3. mobile device-based?

We then concluded that the use of electronic travel services among online consumers was very high and that the services were perceived as very beneficial. It also seemed the electronic channel had reached a level where it passed the conventional channel in preference, at least for typical pre-site activities. The low preference of mobile-device based travel services also indicated that online consumers saw the mobile

device channel as a supplement channel rather than a channel that would compete with e-commerce or conventional interaction. Nevertheless, we also emphasized that wireless and mobile technology was in a very early stage of development and that consumers' ability to understand its true value potential was limited at that time. Therefore, it is essential to follow up on the research questions presented above and compare them to the data collected in 2011.

All three questions will be asked for the same predefined consumer scenario as in 2004. The scenario will be discussed and presented next.

3 Consumer Value Process

Value is what customers experience in their internal processes and when interacting with the service provider as consumers (Grönroos, 2002). On electronic markets, firms can create value for consumers in a manner that is different from that which has been achieved in conventional business (Han & Han, 2001). Many pointed out early on that the demand side of m-commerce is very much a search for value (e.g. Keen & Mackintosh, 2001, Anckar & Dincau, 2002). One way of identifying true consumer value in e-commerce and m-commerce is by using consumer scenarios. As noted by Seybold (2001), thinking in terms of consumer scenarios has always been useful, but with the Internet the technique is more powerful than ever. A consumer scenario or a consumer process usually is made up of several activities and therefore the researcher needs to identify the ones a consumer does to accomplish a certain task, such as purchase a service.

Generally speaking consumers' value process is divided into three phases; initial contact, purchasing and consumption. In the first phase, initial contact, consumers mainly create an interest towards a company and its products or services. In the second phase, purchasing, a consumer searches for information, evaluates the product or service and makes the actual order. In the third phase, consumption, the consumer uses the purchased product or service and relates to the whole experience (Grönroos, 2002). Similar consumer processes have been presented by Blackwell et al. (2001), and in travel and tourism by Crotts (1999). The most simplified traveller process is usually divided into the following steps namely 1. Pre-trip activities 2. On-site activities and 3. Past-trip activities (Werthner & Ricci, 2004).

For this study, which was delimited to the purchasing and the consumption of travel services, a consumer scenario with six activities was identified; *1. Search for information 2. Reservation 3. Payment 4. Cancel or Change reservation 5. Check in 6. Reflection and evaluation*. The first four activities can be characterized as pre-trip activities as they are often performed before the consumer is on-site, whereas the fifth and sixth activities can be regarded as onsite activities or the latter as past-trip activity. The six activities and the example features that explain them in detail are presented in table 1. The scenario may, however, differ for every individual travel consumer and it may as well differ depending on the nature of the travel services used and the situation of use. Therefore the scenario used for this study should only be seen as a tool to find deeper understanding of consumer adoption of electronic travel services.

The scenario was developed bearing in mind the above presented theories of consumer and traveller processes and by talking to different practitioners especially within the hospitality and airline sector. The example features used for each activity in the process are features that to some extent were already provided electronically both as fixed Internet services and mobile services in the airline industry by air carriers like Scandinavian Airlines[1] and Finnair[2] and in the lodging sector by hotels, like Omena Hotels[3] in 2004 when the first survey data for this study was collected.

Table 1. Consumer scenario in interaction with a travel service provider

Scenario activities	Example features
1. Search for information	The consumer e.g. compares prices and checks availability
2. Reservation	The consumer makes the actual reservation of a travel service e.g. hotel rooms, flight tickets, travel packages and train tickets
3. Payment	The consumer pays by e.g. credit card or Internet bank transaction for the travel service
4. Cancellation / Change of reservation	The consumer makes changes to or cancels the reservation
5. Check in	The consumer checks in / activates the purchased travel service.
6. Reflection and evaluation	The consumer gives feedback and checks such things as bonus points during or past a trip

4 Channels of Interaction

Channel management is a challenge from a company perspective. According to Neslin et.al. (2006), there are five major challenges; 1. Understanding customer behavior 2. Data integration 3. Channel evaluation 4. Allocating resources across channels and 5. Coordinating channel strategies. Here we will focus on understanding customer behavior from a consumer process perspective for three different channels of interaction as described in our research questions. Therefore three different channels of interaction will be investigated for the created consumer scenario, namely 1. conventional interaction, 2. computer-based interaction and 3. mobile device-based interaction. For example, Payne & Frow (2005) included in the conventional channel such interaction as a face-to-face encounter with a representative of a company and telephony (traditional phone and call center contact). The computer-based channel includes mainly Internet services through different computers but also interaction through e.g. self service machines. Also wireless computers (laptops and tablets) are included in the computer-based channel as they are not the same as mobile devices

[1] Scandinavian Airlines (SAS Group) is the Nordic region's largest listed airline and travel group, in terms of passengers and operating revenue. www.sas.fi

[2] Finnair is one of the world's oldest operating airlines and flies within Finland and to about 60 international destinations. www.finnair.fi

[3] OmenaHotels is a hotel chain primarily operating in Finland. Omena operates without neither reception personnel nor a reception desk, and thus no check-in or check-out procedures. All possible work tasks have been completely automated through IT. www.omena.com

even though there is considerable overlap; wireless computers do not necessarily need to support true mobility due to size (harder to bring along in every situation compared to a pocket size device). The mobile device-based channel is made up of Internet or mobile services through a mobile device. We refer to a mobile device as a mobile handheld device or mobile pocket device such as a mobile phone or smart phone. The miniature size as 'fits into a pocket and can be carried everywhere' has also been emphasized by others when defining a mobile device (e.g. Tsalgatidou et al., 2000).

5 Data Collection and Samples

In order to accomplish the research objectives a questionnaire was developed based on the consumer scenario presented and the channels features discussed. The primary data needed for the survey was collected through a self-administered Web-questionnaire, which was linked from the Web site of one cooperating company within the lodging sector in Finland. For the first data collection in fall 2004 the questionnaire received a total number of 766 applicable answers. The second data collection in spring 2011 the questionnaire received a total number of 927 applicable answers.

Of the respondents in 2004, 33% were males and 67% were females. The average age of the respondents was 34 years; the youngest was 18 years old and the oldest 67 years old. The respondents were all Internet adopters since the questionnaire was filled out on-line and nearly all (93.9%) reported that they visit the Internet every day or several times per week. 38.2% had, at least at some point, used the Internet through a mobile device and 36.8% were frequent travellers, who travel at least once per month either business or leisure for a minimum duration of one day.

Of the respondents in 2011, 24.1% were males and 75.9% were females. The average age of the respondents was 38 years; the youngest was 18 years old and the oldest 71 years old. The respondents were all Internet adopters since the questionnaire was filled out on-line and nearly all (98.5%) reported that they visit the Internet every day or several times per week. 67.7% had, at least at some point, used the Internet through a mobile device and 29% were frequent travellers, travel at least once per month either business or leisure for a minimum duration of one day.

We checked the samples for equality as they were not randomly selected from a predefined population (See table 2.). We can conclude that the two samples are similar in character as both have a clear majority of female respondents, representation of every age group, a clear majority of non-frequent travellers and respondents with very high Internet use. However, it has to be pointed out that the 2011 sample is significantly more skewed towards females, older consumers and non-frequent travellers. The way Internet is used in Finland is, however, fairly similar regardless of gender, and Internet use is nowadays very widespread among different age groups (Statistics Finland, 2010). The 2011 sample includes, not surprisingly, significantly more frequent Internet users and a lot more mobile Internet users. In general 86% of the Finnish population (age 16 – 74) is Internet users, with an almost equal use among men (87%) and women (85%). Men, however, use mobile devices more frequently than women for connecting to the Internet (Statistics Finland, 2010). The female dominance in our samples may therefore influence the perceptions about

the use of mobile devices for electronic travel services. The findings, which will be presented next according to research question, should therefore not be overgeneralized for all Finnish Internet adopters.

Table 2. Analysis of samples

Variable	Sample 2004 N = 766	Sample 2011 N = 927	χ^2/df	Sig. (2–sided)
Gender:				
Males	33.0%	24.1%	17.819 / 1	0.000
Females	67.0%	75.9%		
Age:				
16 – 24	21.5%	14.5%	59.157 / 6	0.000
25 – 34	29.0%	23.1%		
35 – 44	29.0%	29.2%		
45 – 54	15.5%	21.5%		
55 – 64	2.9%	8.5%		
65 – 74	0.3%	2.1%		
Internet adoption - use at least several times per week)	93.7%	98.5%	24.704 / 1	0.000
Mobile Internet adoption - has at least at some point tried	38.2%	67.7%	252.598 / 1	0.000
Travel frequency - travel at least once per month business and/or leisure	36.8%	29.0%	9.548 / 1	0.002

6 The Use of Electronic Travel Services (RQ1)

The use of electronic travel services in 2004 through a computer accounted for more than 50% of all the respondents for the first four activities in the pre-defined consumer scenario; Search for information, Reservation, Payment and Cancellation / Change of reservation (See table 3.). In 2011 the use of a computer for not only the first four activities but also the sixth activity, reflection and evaluation, accounted for more than 50% of all respondents. Also the electronic channel (use of a computer and/or a mobile device) for the fifth activity, check in, accounted for more than 50% of all respondents in 2011. A slight increase of only mobile device use is shown for every activity (Total use 0.4% to 0.9%) but the major increase since 2004 is the use of both a computer and a mobile device for all six activities (Total use 4.0% to 13.5%) . The Pearson Chi-square test shows that there is a significant difference in the distribution of use between 2004 and 2011 for all six activities.

Table 3. The use of electronic travel services (RQ1)

	Use 2004					Use 2011					
		NO	CO	MD	CO / MD		NO	CO	MD	CO / MD	Diff.
Scenario activities	**N**	**%**	**%**	**%**	**%**	**N**	**%**	**%**	**%**	**%**	**χ2 / 3**
1. Search	761	4.6	**86.7**	0.1	8.5	918	2.9	**72.3**	0.4	23.9	73.3*
2. Reserve	756	15.6	**79.1**	0.4	4.9	918	3.3	**79.4**	1.0	15.7	120.1*
3. Pay	754	27.2	**69.9**	0.5	2.4	920	6.3	**86.6**	0.7	6.2	144.9*
4. Cancel / Change	744	38.3	**54.6**	0.5	3.8	911	14.2	**68.2**	1.5	14.8	164.3*
5. Check in	735	**81.0**	15.0	0.7	3.4	898	47.7	**36.0**	**1.4**	**14.8**	412.8*
6. Reflect / evaluate	736	**53.5**	44.8	0.4	1.2	902	39.9	**51.2**	0.5	6.1	44.3*
Total	748	36.7	**58.4**	0.4	4.0	911	19.1	**65.6**	0.9	13,5	

Bolded numbers show a total percentage > 50%
* Significant at 0.001 level

NO = No use
CO = Only a computer (Internet services through a computer)
MD = Only a mobile device (Internet services or mobile services through a mobile device)
CO / MD = Both a computer and a mobile device

7 Perceived Benefits of Electronic Travel Services (RQ2)

The mean values, presented in table 4, show that there is a general increase in perceived benefit of electronic travel services from 2004 to 2011 among respondents who have actual experience of using such services. However, based on t-tests the increase is not significant for any of the activities but for 'Reservation'. 'Reflection and evaluation' still seems to add the least value to the online travel consumers as it accounts for the lowest mean of all the activities both in 2004 (3.20) and 2011 (3.23).

Table 4. Perceived benefit of electronic travel services by respondents using them. (RQ2)

	Benefit 2004			Benefit 2011			
Scenario activities	**N**	**Mean[1]**	**Std. dev.**	**N**	**Mean[1]**	**Std. dev.**	**Diff. t-Sig.**
1. Search for information	729	3.65	0.499	887	3.69	0.487	0.107
2. Reservation	640	3.72	0.475	879	3.78	0.420	0.004*
3. Payment	558	3.66	0.551	855	3.68	0.527	0.548
4. Cancellation / Change of reservation	462	3.67	0.536	787	3.71	0.476	0.169
5. Check in	195	3.44	0.681	529	3.53	0.603	0.097
6. Reflection and evaluation	374	3.20	0.700	577	3.23	0.693	0.446

[1]Scale: [4] very beneficial; [3] rather beneficial; [2] not very beneficial; [1] definitely not beneficial. "Can't say" answers are excluded.
*Significant at 0.01 level

8 The Preferred Channel of Interaction (RQ3)

As shown in table 5, the electronic channel (computer-based and mobile device-based) clearly wins both 2004 and 2011over the conventional channel for every activity in the scenario except for the 'check in' activity. However, in 2011 electronic check in is preferred by a majority of the respondents (>50% prefer computer or mobile device). The computer-based channel clearly wins over the mobile device-based channel for every activity both 2004 and 2011. In fact, the computer based channel has clearly grown in preference for every activity since 2004 (Total 64.0% to 80.4%), while, interestingly, the preference of the mobile device channel has decreased since 2004 for every activity (Total 6.5% to 3.4%). The Pearson Chi-square test shows that there is a significant difference in channel preference 2004 and 2011 for each activity. Check-in stands out as the most preferred activity on a mobile device both 2004 (9.2%) and 2011 (8.2%).

Table 5. Preferred channel of interaction with travel service providers. (RQ3)

	Preferences 2004				Preferences 2011				
Scenario Activities	**N**	**CON %**	**CO %**	**MD %**	**N**	**CON %**	**CO %**	**MD %**	**Diff. $\chi 2$ / 2**
1. Search for information	761	9.3	**84.2**	6.5	825	2.7	**93.7**	3.6	39.5*
2. Reservation	756	31.3	**63.5**	5.2	824	9.0	**88.8**	2.2	140.2*
3. Payment	754	18.9	**75.4**	5.7	806	5.1	**92.9**	2.0	89.1*
4. Cancellation / Change of reservation	744	42.4	**52.3**	5.3	805	19.3	**76.2**	4.6	162.2*
5. Check in	735	**51.1**	39.7	9.2	753	45.4	**46.4**	**8.2**	71.8*
6. Reflection and evaluation	736	23.9	**69.1**	7.0	754	10.6	**84.5**	4.9	50.5*
Total	748	29.5	**64.0**	6.5	795	15.4	**80.4**	3.4	

Bolded numbers show a total percentage > 50%
* Significant at 0.001 level

CON = Conventional interaction e.g. face to face or calling
CO = Computer-based interaction / Internet services through a computer or a self service machine
MD = Mobile device-based interaction / Internet services or mobile services through a mobile device

9 Implications, Limitations and Further Considerations

This study confirms a steady growth of electronic travel service use among Internet adopters in Finland for the researched consumer process; starting with search for information and ending with reflection and evaluation. The use of electronic services in 2011, either computer based and/or mobile device based, accounts for the majority of all respondents (>50%) for each step in the pre-defined consumer process. Especially the latter part of the process (check in and reflection / evaluation) has seen the biggest growth in use over the past seven years. Certainly much due to that these kinds of services are now more available to consumers.

The industry has also succeeded in providing valuable electronic services and, more importantly, succeeded in improving since 2004. In particular the perceived benefit of 'reservation' stands out as significantly more beneficial online in 2011 than in 2004, which indicates a clear improvement in online reservation transactions. Interesting is that the perceived benefit of the last step 'Reflection and evaluation' is still found clearly the least beneficial of the steps in the process.

Moreover, 'Reflection and evaluation' hasn't significantly improved since 2004, despite the enormous growth of online social media platforms where many travel

service providers are present. This result certainly indicates that there are still needs for industry improvements in the latter part of the consumer online process.

The reported channel preference also shows that the electronic channel is the winner for all steps in the studied consumer process. In 2011 the electronic channel (computer and/or mobile device) has passed the conventional channel in use and preference among Internet adopters in Finland, not only for typical pre-site activities as in 2004, but for the whole consumer process investigated. Early studies (e.g. Law et. al. 2004) of disintermediation of conventional travel services such as conventional travel agents forecasted that online and offline services would remain equally important for consumers (also Internet adopters). According to this research that is not the case as the consumer preference by an overwhelming margin is the electronic channel (either computer or mobile device) for every step in the consumer process investigated. Interesting, however, is the decrease of the mobile device channel as preference for every step in the process. It seems like online consumers back in 2004 were more optimistic regarding their preference of using a mobile device than today. Their actual online use of a mobile device together with a computer has, not surprisingly, increased since 2004, which means that more subjects base their choice of preference on real mobile use than in 2004. Despite huge technological advances this result certainly confirms that the mobile channel is primarily perceived as a supplement channel rather than a standalone channel that would directly compete with or dismiss the off-line channel or the computer-based channel for the investigated consumer process. That doesn't mean that the mobile channel isn't important or doesn't create value to the consumer.

On the contrary, as this study shows, a growing number of online consumers use both a computer and a mobile device (not either or). Therefore, the mobile channel cannot be seen as a separate channel only provided by certain specialized vendors but as a closely integrated channel, which the consumer is likely to switch to more and more for any of the activities in the investigated process. It also has to be emphasized that this goes for any of the six activities in the process and not for only particular ones. Therefore, when designing mobile services, not only certain parts of the process but the whole consumer process should be supported. For example, it may not be good enough to let the consumer search for information on his or her mobile device but then leave the reservation, payment and so on for him or her to do in another channel. Moreover, travel providers cannot rely on full size web sites in the mobile channel as they significantly reduce the self-arrangement experience on a mobile device (Eriksson, 2011).

Moroever, mobility (on the move aspect) plays a significant role whether a consumer will use a mobile device or not (Anckar & Dincau, 2002, Anckar & Eriksson, 2003). In this study we didn't differentiate between fixed and mobile consumer settings, as we looked primarily at the consumer process from a situation independent perspective to gain an overall view of the three channels under investigation. However, as discussed in the setup of the consumer scenario, especially the activity 'check in' is more on trip or on site in nature than the other activities.

Moreover, we didn't differentiate between the complexity of the described activities e.g. check in procedures may be found simpler than for example search and booking

procedures and thereby more suitable on a mobile device due e.g. to limited screen size. In this study 'Check in' stands out as the most preferred activity on a mobile device both 2004 and 2011, which supports that mobility and/or activity complexity influence consumers perceptions regarding the mobile channel. Therefore, an interesting question is if the mobile device channel is the preferred channel when consumers are on the move and when typical on site activities are investigated. Moreover, despite the tremendous growth of the adoption of smartphones and mobile services there may still be barriers out there, which influence consumers' perceptions regarding the use of mobile services in travel. We will look into these questions in our upcoming research.

This study has its limitations as, due to the online data collection method, non-Internet adopters are excluded. On the other hand as many as 86% of the population in Finland uses the Internet (Statistics Finland, 2010). The two samples compared were similar in character and represent fairly well Finnish Internet adopters in general. However, both samples were skewed towards females, especially the sample from 2011. Therefore, the findings shouldn't be overgeneralized for all Finnish Internet adopters. Moreover, it has to be pointed out that the results are valid primarily for the Finnish society, which is characterized by high acceptance of information technology.

It should also be emphasized that we included tablet devices into the computer-based channel and not in the mobile device channel due to size. There is, however, a considerable overlap between tablets and mobile pocket devices like smartphones regarding the user experience and mobility features (e.g. GPS). Companies should, therefore, be aware that the growing tablet market may require a different consumer approach regarding development of applications and user experience compared to e.g. lap top computers.

References

Anckar, B. (2003). Consumer Intentions in Terms of Electronic Travel Distribution: Implications for Future Market Structures. *e-Service Journal* 2(2): 68-86.

Anckar, B. & D'Incau, D. (2002). Value creation in mobile commerce: Findings from a consumer survey. *Journal of Information Technology Theory and Application* 4(1): 43-65.

Anckar, B. & Eriksson, N. (2003). Mobility: The Basis for Value Creation in Mobile Commerce*?*. *In Proceedings of the International Conference SSGRR'03*. L'Aquila, Italy: Telecom Italia learning Services.

Blackwell, R. D., Miniard, P. W. & Engel, J. F. (2001). *Consumer behavior (9th ed*.). Mason, OH: Thomson South-Western.

Buhalis, D. & Law, R. (2008). Progress in information technology and tourism management: 20 years on and 10 years after the Internet – The state of the eTourism research. *Tourism Management* 29(4): 609-623.

Crotts, J. C. (1999). Consumer decision making and pre-purchase information search. In Mansfield and Pizam (Eds.), *Consumer behaviour in travel and tourism*. Binghamton, NY: Haworth Press.

Durlacher Research (2000). *UMTS Report, An Investment perspective*. Retrieved from www.durlacher.com.

Durlacher Research (2005). *UMTS Report, An Investment perspective*. Retrieved from www.durlacher.com.

Eriksson, N. (2011). User Experience of Trip Arrangements – a Comparison of Mobile Device and Computer Users. Conditionally accepted to *Journal of E-Services and Mobile Applications*.

Gammage, B. (2010). *Coping with the New Balance of Power. Gartner top predictions 2010, Gartner Webinar 14th January 2010*. Retrieved March 5, 2011, from http://www.gartner.com/it/content/1260200/1260221/january_14_top_predictions_2010 bgammage.pdf.

Grönroos, C. (2002). *Service Management and Marketing – A Customer Relationship Management Approach. (2nd edition*). Chichester: Wiley.

Han, J. & Han, D. (2001). A Framework for Analyzing Customer Value of Internet Business. *Journal of Information Technology Theory and Application* 3(5): 25-38.

Husson, T. & Ask, J. (2011). *2011 Mobile Trends*. Forrester Research, Inc.

Hyvönen, K. & Repo, P. (2005). Mobiilipalvelut suomalaisten arjessa. *Kuluttajatutkimus-keskuksen vuosikirja*. Kerava: Savion Kirjapaino Oy.

Kaufman-Scarborough, C. & Lindqvist, J. D. (2002). E-shopping in a multiple channel environment. *Journal of Consumer Marketing* 19(4): 333-350.

Keen, P. & Macintosh, R. (2001). *The Freedom Economy – Gaining the M-commerce Edge in the Era of the Wireless Internet*. Berkeley: McGraw-Hill.

Law, R., Leung, K. & Wong, R. (2004). The impact of the Internet on travel agencies. *International Journal of Contemporary Hospitality Management* 16(2): 100-107.

Ministry of Transport and Communication (2005). Mobile services market in Finland 2004. In *Publications of the Ministry of Transport and Communications*, No. 34, Helsinki.

Neslin, S. A., Dhruv, G., Leghorn, R.,Venkatesh, S., Teerling, M. L.. Thomas, J. & Verhoef, P. C. (2005). Challenges and Opportunities in Multichannel Customer Management. *Journal of Service Research* 9(2): 95-112.

Payne, A. & Frow, P. (2005). A Strategic Framework for Customer Relationship Management. *The Journal of Marketing* 69(4): 167-176.

Schoenbachler, D. D. & Gordon, G. L. (2002). Multi-channel shopping: understanding what drives channel choice. *Journal of Consumer Marketing* 19(1): 42-53.

Seybold, P. (2001). Get Inside the Lives of Your Customers. *Harvard Business review* 79(5): 80-89.

Simons, L. P.A., Steinfield, C. & Bouwman, H. (2002). Strategic positioning of the Web in a multi-channel market approach. *Internet Research: Electronic Networking Applications and Policy* 12(4): 339-347.

Statistics Finland (2010). *Every second Finn uses the Internet several times a day*. Retrieved September 6, 2011, from http://tilastokeskus.fi/til/sutivi/2010/sutivi_2010_2010-10-26_tie_001_en.html

Tsalgatidou, A., Veijalainen, J. & Pitoura, E (2000). Challenges in Mobile Electronic Commerce. In *proceedings of IeC 2000. 3rd Int. Conf. on Innovation through E-Commerce.* Manchester, UK.

Walden P., Han S., Carlsson, C. & Majlender, P. (2007). The Sleeping Giant – A Longitudinal Study Surveying the Mobile Service Market in Finland. In H. Österle, J. Schelp, and R.Winter (Eds.), *Proceedings of the 15th European Conference on Information Systems.* St Gallen, Switzerland: University of St Gallen.

Werthner, H. & Ricci, F. (2004). E-commerce and tourism. *The communication of ACM* 47(12): 101-105.

Modelling and Comparing Malaysian Hotel Website Diffusion

Noor Hazarina Hashim[a],
Miriam Scaglione[b], and
Jamie Murphy[c]

[a]Department of Management
Universiti Teknologi Malaysia
m-hazarina@utm.my

[b]Institute of Tourism,
Applied University of Western Switzerland
miriam.scaglione@hevs.ch

[c]Murdoch Business School
Murdoch University
jamie.murphy@murdoch.edu.au

Abstract

This study draws upon diffusion modelling, a coupled Bass Rogers model, to examine hospitality website adoption within Malaysia and across countries. The results suggest 38% of Malaysian hotels may never adopt a traditional website. Rather, these hotels may start with new online channels, such as social networking and communities, leapfrogging over those that adopted websites. Across countries, the results highlight the role of innovation and imitation in adopting websites. Malaysian and international chains hotels showed similar adoption related to innovation and imitation. For Swiss hotels, the influence of imitation was much stronger than for Malaysian and international chain hotels.

Keywords: Malaysian hotels, Bass-Rogers Model, Adopter categories, Website diffusion

1 Introduction

Few technologies have the scope and impact on businesses as digital technologies and the Internet do. Network values in Metcalfe's Law, cost and space reductions with Moore's Law, and converging technological devices help fuel explosive Internet growth (Hanson & Kalyanam, 2007). For over a decade scholars have investigated the Internet's widespread impact on business and society. Yet many business website studies have at least two limitations.

Firstly, the population often omits businesses late to adopt websites; the population is right-censored. Similar to non-respondents in survey research, omitting businesses yet to adopt websites gives an incomplete view of technology diffusion. Early hotel Internet studies investigated a hotel subset, the early to adopt, and may have examined just innovator and early adopter hotels (see for examples: Murphy, Forrest, Wotring, & Brymer, 1996; Siguaw, Enz, & Namiasivayam, 2000). Likewise, research on technology and innovation in developing economies should not always be grounded on theories derived from leading economies (Da Silveira, 2001). As argued by Zhu and Kraemer (2005, p. 62), theories developed in "mature markets and industrialised

economies need to be re-examined in the context of developing countries, because these countries may have very different economic and regulatory environments".

Businesses also face evolving digital communication methods and metrics (Hoffman & Fodor, 2010). Today's travellers rely on user-generated content for travel decision-making more so than published travel company information (Pan & Fesenmaier, 2006; Xiang, Gretzel, & Fesenmaier, 2009). Such rapid generational technology changes, as in eTourism, may lead to leapfrog effects (Kauffman & Techatassanasoontorn, 2009; Stremersch, Muller, & Peres, 2010). That is, the late adopters bypass early adopters in effective technology use (Davison, Vogel, Harris, & Jones, 2000). More concretely, some hotels may start a strong Facebook or Twitter presence rather than a traditional website (Anderson & Wolff, 2010).

The second limitation, poor estimates of total market size and adopter categories, compounds the incomplete view. Rogers (2003) models technology diffusion as a normal distribution across five adopter categories: *innovators*, the first 2.5% to adopt an innovation; *early adopters*, the next 13%; *early majority*, the next 34%; *late majority*, the next 34% and finally *laggards*, the last 16% to adopt. Yet, each adopter category differs in characteristics and innovation use (Hsu, Lu, & Hsu, 2007; Rogers, 2003). There is, however, little justification that Rogers' category sizes fit all products (Mahajan, Muller, & Bass, 1990). Rather, products may have abnormal adopter category distributions (Peterson, 1973).

A diffusion model that addresses these two limitations, the coupled Bass Rogers (BR), classifies adopter categories and estimates both, product growth and innovation diffusion (Peres, Muller, & Mahajan, 2010; Scaglione, Schegg, & Murphy, 2009; Stremersch, et al., 2010). Addressing limitations of some website adoption studies, this paper draws on a coupled BR model and Malaysian hospitality data to calculate total website adoption and adopter categories.

This paper begins with technology diffusion literature followed by the methodology to predict hotel website diffusion in Malaysia. The paper compares the Malaysian findings with Swiss hotels and international chain hotels to compare hospitality website adoption across countries. The study closes with conclusions, limitations and recommendation for future research. This supports Da Silveira's (2001) argument that research on technology and innovation in developing economies should not always be grounded on theories derived from leading economies. As argued by Zhu and Kraemer (2005, p. 62), theories developed in "mature markets and industrialised economies need to be re-examined in the context of developing countries, because these countries may have very different economic and regulatory environments".

2 Literature Review

2.1 Organisational Diffusion

An early diffusion academic, French sociologist Gabriel de Tarde at the turn of the 19th century discussed society's acceptance and rejection of innovations (Kinnunen, 1996). Diffusion research continued during the 1940s and 50s as independent studies in anthropology, education, public health and medical sociology, marketing, management, communication, and general sociology (Rogers, 2003). Everett M.

Rogers' 1962 book, *Diffusion of Innovations* (DOI), the most cited work in innovation research (Jeyaraj, Rottman, & Lacity, 2006; Zhu & Kraemer, 2005), discusses adoption of innovations at the individual and organisational level. Rogers (2003, p. 12) defines an innovation as "an idea, practice, or object that is perceived as new by an individual or other units of adoption."

Organisational diffusion research, the focus of this study, examines characteristics related to the adoption and subsequent implementation of innovations. As Figure 1 shows, organisational diffusion studies comprise adoption and implementation research streams (Fichman, 2000; Rogers, 1962, 2003). Rather than adopting an innovation, implementation studies investigate organisational use of innovations. Adoption studies, far more popular than implementation studies, tend to focus on and model innovation diffusion rates and factors related to adoption (Fichman, 2004).

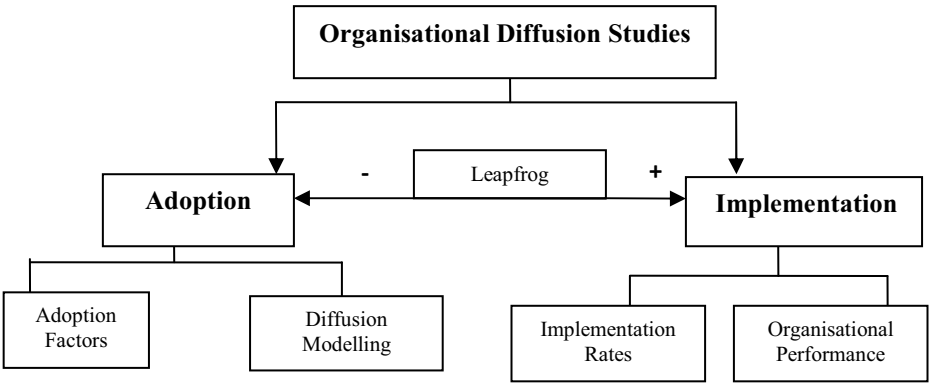


Fig.1. Organisational innovation diffusion studies

A meta-analysis of adoption studies identified three factors − organisational, innovation and environmental – common to the organisational adoption of an innovation (Jeyaraj, et al., 2006). Thanks in part to half a century of research, adoption factor studies seem to be reaching maturity. Most studies conclude that individuals and organisations with "greater innovation-related needs and abilities or the Right Stuff are more likely to adopt an innovation than those with less needs and abilities (Fichman, 2004, p. 315)".

2.2 Diffusion Modelling

Diffusion modelling studies examine market diffusion, adoption rates and adopter categories (Bass, 1969; Mahajan, et al., 1990). Internet diffusion follows an abnormal *distribution* due to complex and diverse economic, political, cultural, infrastructural and geographical factors (Andrés, Cuberes, Diouf, & Serebrisky, 2010; Kim, 2011), and adopter categories differ in technology implementation. At the individual level, American Online users categorised as innovators preferred content-based online information while laggards preferred interpersonal information (Stafford, 2003). Similarly, a study on multimedia message services found significant differences in perceived use between potential adopters and users (Hsu, et al., 2007).

At the organisational level, a study of medium-large US and Singapore firms found significant website design differences between early and late adopters (Teo & Pian, 2004). The early adopters, in addition to providing information, provided transactions, customer services and personalisation. The late adopter websites mainly provided information. As well, a study of top 1000 US and Taiwan companies found most Taiwanese companies online used their sites for presenting product information. The US companies gave information and also used their sites for customer relationship (Liao, To, & Shih, 2006).

Despite differences across adopter categories, there is a lack of theoretical support and few reliable and valid measures for adopter category constructs (Mahajan, et al., 1990). Nor is there justification for uniform adopter category sizes for all products. Many products have abnormal adopter category distributions (Hsu, et al., 2007; Peterson, 1973) and Roger's (2003) normal distribution underestimates many diffusion patterns (Mahajan, et al., 1990). Furthermore, recent IT studies show abnormal diffusion patterns across product generations, such as today's smart phones (Kauffman & Techatassanasoontorn, 2009; Stremersch, et al., 2010). These rapidly improving technologies enjoy a shorter takeoff period but similar overall growth relative to other technologies (Stremersch, et al., 2010).

2.3 Modelling Innovation Diffusion

Given limitations with Roger's (2003) model, studies develop data and innovation-specific adopter categories (Mahajan, et al., 1990). Diffusion models such as Peterson (1973), Bass (1969) and Stremersch et al., (2010) also fit abnormal distributions. Rather than adopter categories based on given percentages, their categories are geographically and innovation specific. The coupled Bass and Rogers diffusion model, "has the dual advantages of allocating adopter categories based on actual data and reflecting two coefficients—innovation and imitation—that influence adoption (Scaglione, et al., 2009, p. 626)."

Bass (1969) improved Rogers (1962) model by quantifying two factors that drive individual and organisational adoption, *innovation* driven by external channels such as mass communication and *imitation* driven by internal communication such as word of mouth. Equation 1, below, shows the Bass function (1969) for the growth of a new product.

$$\frac{dN(t)}{dt} = \left[p + q\frac{N(t)}{m} \right](m - N(t)) \qquad (1)$$

N(t) is the cumulative number of adopters at time t, m is the total market adoption, parameters p and q, the coefficients of innovation and imitation, respectively (Bass, 1969).

2.4 The Bass Rogers Model

Mahajan et al. (1990) used Rogers' bell-shaped adopter categories and Bass' parameters to develop the Bass Rogers (BR) model. The three inflection points in Figure2 below–T1,T and T2–set the adopter categories in the BR model (Mahajan et al., 1990; Mahajan, Muller, & Wind, 2000). Up to T1, the adoption rate increases rapidly and delineates Early Adopter and Early Majority categories. From T1 to the

peak T, the highest adoption rate, the adoption rate continues to rise, but at a slower rate. T separates the Early Majority and Late Majority. From the peak, T, the adoption rate decreases slowly until T2, when the rate slows to a crawl. T2 separates the Late Majority and Laggards. Figure 2 shows these inflection points and adopter categories for Malaysian hotels.

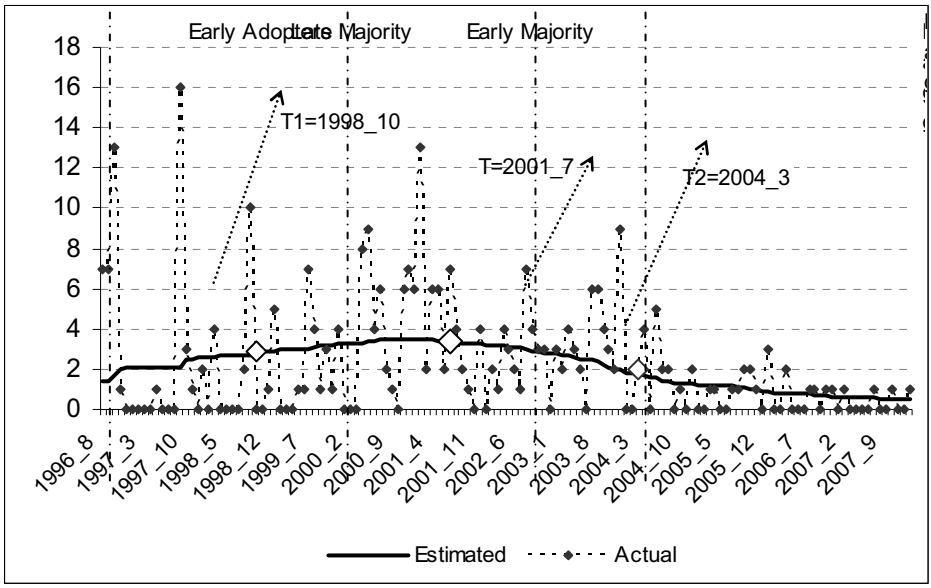


Fig.2. Malaysian BR model estimates, monthly adoption

2.5 Innovation and Imitation

The innovation coefficient *p* suggests the propensity of an initial purchase in the first time interval, up to T1 (Schmittlein & Mahajan, 1982). The imitation coefficient *q*, reflects the propensity of an initial purchase in the second time period, from T1 to T. The second time interval is driven by $[p+qN(t)/m]$. This latter probability is an addition of the probability of adopting by innovation (p) plus the probability of adopting by imitation. As p is constant, the probability of adopting by imitation is proportional to the total number of adopters before ($N(t)/m$). With pure innovation, p=1, Equation 1 is an exponential function. With pure imitation, q=1, Equation 1 is a logistic function (Meade & Islam, 2006).

Apart from these extreme cases, the sum and ratio of q and p suggest the shape of the diffusion curve. The sum controls the dispersion. The ratio $q/p > 1$ indicates a bell shaped non-cumulative adoption and S-shaped cumulative adoption. The bigger the ratio, the more pronounced the S-shape (Meade & Islam, 2006). Shapes of the S curves and q/p ratios help compare adoption across countries.

Van den Bulte and Stremersch's (2004) analysis of q/p ratios across 28 countries and individual adoption of 52 consumer durables found, among other aspects, a positive relationship between the q/p ratio and income inequality. The bigger a country's income inequality, the more that innovation rather than imitation drove adoption.

Likewise, the bigger a country's income equality, the more that imitation rather than innovation drove adoption

At the organisational level, a study of German banks showed that relative to non-interactive innovations, the diffusion of interactive innovations (i.e. Electronic Funds Transfers and home banking for private customers) was slower until reaching a critical mass of adopters. The S-Shape for the non-interactive innovations was more pronounced than for interactive innovations. As shown in Figure 3 shows, critical mass makes the innovation valuable, and once reaching critical mass, the adoption rate accelerates.

Critical mass is less an objective number of adopters and more the perceived number of adopters. If individuals have similar resistance/attraction, namely the *threshold*, to the innovation based on the objective number of critical mass' adopters, the innovation will never launch, but individuals and organisations have different threshold of resistance to innovation adoption (Mahler & Rogers, 1999).

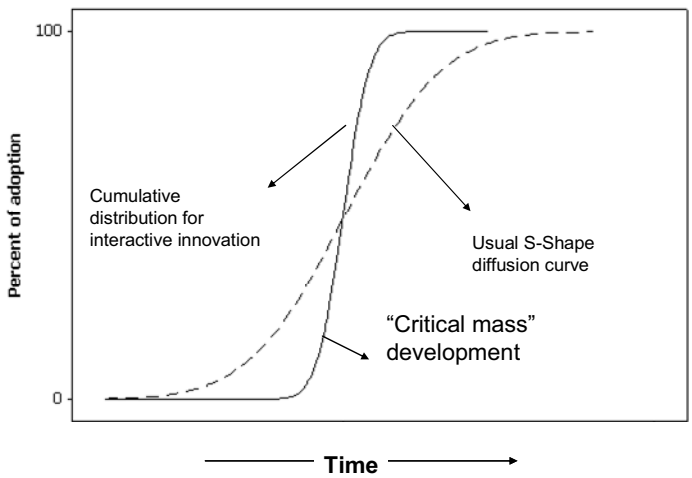


Fig.3. Cumulative distribution for *interactive* (solid line) and *non-interactive* (dotted line) innovations based on (Mahler & Rogers, 1999)

Nonlinear least square (NLS), a popular method to estimate parameters *p* and q, can have problems with a correctly specified model (Van den Bulte & Lilien, 1997). Even if the NLS estimates converge in probability to the estimated parameter value, they may be biased. There is no way to exclude that the estimation differs systematically from the parameter. A second problem with estimating Bass parameters is available observations. Censored data has only partial information available. As censored data decreases, the estimated coefficient of imitation (*q*) decreases whereas *p* and *m* increase (Van den Bulte & Lilien, 1997, p.242).

The estimated final market size, *m*, merits cautious interpretation. The calculated eventual penetration a new technology may seriously underestimate the potential market. Furthermore, the technology will evolve and may die.

2.6 Leapfrogging Across Technology Generations

A 2010 article in the magazine *Wired* argues that the web is dead. "Over the past few years, one of the most important shifts in the digital world has been the move from the wide-open Web to semi closed platforms that use the Internet for transport but not the browser for display (Anderson & Wolff, 2010)." Future Internet use, business and individual, may shift to applications for social networking or peer-to-peer sharing rather than the traditional website. As the eventual use of an innovation takes about 30 years to mature (Fidler, 1997), tourism and hospitality website studies may be examining an endangered species.

Furthermore, those late to adopt a technology may leapfrog early adopters in effective technology use (Davison, et al., 2000). For some individuals, organisations (Amir, 2004; Gallagher, 2006; Hobday, 1995; Rosenkranz, 1997) and countries (Gray & Sanzogni, 2004; Rosenkranz, 1997; Steinmueller, 2001), being late to adopt has advantages. This leapfrog effect accelerates technology implementation in industries such as manufacturing, aviation, energy and health (Amir, 2004; Gallagher, 2006; Hobday, 1995; Rosenkranz, 1997). For developing nations - Thailand, Egypt, and Malaysia - technology leapfrogging helps narrow productivity and output gaps, and helps reduce research and development costs (Gray & Sanzogni, 2004; Steinmueller, 2001). The following methodology models website diffusion in a developing country, Malaysia, as a preliminary step towards investigating leapfrogging.

3 Methodology

The data stemmed from an Internet adoption study of all 540 hotels registered with the Ministry of Tourism Malaysia (Hashim, Murphy, Purchase, & O'Connor, 2009). As the Ministry database provided no website or email addresses, keying the 540 hotel names into Google and Yahoo returned 315 hotel urls and 10 closed hotels. The 530 operating hotels ranged from 10 to 1,234 rooms, 39% were chain affiliated and 42% had one or two stars, followed by 28% with three stars and 30% with over three stars.

As the BR model requires time-series data, this study used website age from the Wayback Machine (www.archive.org). This online tool showed a significant positive correlation between website age and domain name age ($r=.933$, $p<.001$) for Malaysian hotels (Murphy, Hashim, & O'Connor, 2007). Gathering the website age from the Wayback Machine yielded 305 hotel website ages from the 315 hotel websites. Password-protected, dynamic sites and owner exclusion requests help explain the 10 hotels without a website age. The remaining 305 hotels had an average age of five years and age range in 2007 from seven months to over nine years.

BR parameters p, q and their ratio q/p help compare an innovation's diffusion across countries (Van den Bulte & Stremersch, 2004). For cross country evaluations, this research takes Scaglione, Steiner, Schegg, & Murphy's (2005) analysis of Swiss hotels' adoption of domain names and international chain hotels retrieved from hotelsmag.com in July 2006 (Scaglione, Ismail, Trabichet, & Murphy, 2010).

4 Results and Discussion

4.1 Malaysian Hospitality Website Adoption and Adopter Categories

This study classified adopters using the coupled BR model (Scaglione, et al., 2009) and the 305 Malaysian hotels with website age. The authors estimated Bass model parameters using two nonlinear regressions, which both replace continuous time by discrete periods (Lilien & Rangaswamy, 2002). These two estimations use SAS Institute V9.2 routines Proc nlin and Proc Model (SAS Institute Inc., 2011). Calculations for the total market, namely all hotels that will adopt websites at the end of the diffusion process were 316.9 (std=13.6), 316.9 (std=12.8), respectively for Proc nlin and Proc Model. The results converge on 317, but omit the 10 hotels with an unknown website age, for a total market size of 327 websites or 62% of the 530 hotels. Table 1 shows Rogers' (2003) proposed adopter category percentages, BR estimates and the observed adoption.

Table 1. Market Size and Adopter Categories for Malaysian Hotels

	Rogers	BR limits	BR estimate	Observed	Average website age (days)	Hotels actual	BR hotels eventual
Innovators	2.5%	0.2-2.8%	2.2%	2.2%	4248	7	7
Early Adopters	13.50 %	9.5-20%	15.5%	15.5%	3962	49	49
Early Majority	34%	29.1-32.1%	32.6%	32.6%	2942	103	103
Late Majority	34%	29.1-32.1%	33.5%	33.5%	2035	106	106
Laggards	16%	21.4-23.5%	16.1%	12.7%	1004	40	51
Total	100%	100 %	100%	96.2%		305	317
Adopters without Wayback age							10
Final Total Adopters							327
Total							530

Table 1 shows that the percentage of B-R categories follows the theoretical values except for Late Majority and Laggards. For Late Majority, the estimate (33.5%) is 1.4% higher than the upper value of the BR theoretical limit of (32.1%), which is acceptable in terms of model estimation errors. The case of Laggards is in the opposite direction, 40 actual adoptions and 51 eventual adoptions. This difference shows that the diffusion has not stopped.

4.2 Comparing Hospitality Website Diffusion

Table 2 below shows a comparison of the BR estimates for Malaysian, Swiss and international chain hotels, listed by the date when adoption peaked, T. The first sector to peak was international chain hotels, in September 2000. The Malaysian hotels

followed almost a year later, in July 2001, and the Swiss hotels hit their maximum adoption rate another seven months later in February 2002. The inflection points T1 and T2 tend towards symmetry around the peak T for international and Malaysian hotels; Swiss hotels show an asymmetric distribution around T. International chain hotels lagged about 28 months from T1 to T and 26 months from T to T2. For Malaysian hotels the first lag is 34 and the second 33 months. Swiss hotels differ, with 62 and 22 months respectively. Thus, Swiss hotels shows a slower diffusion process before the maximum peak of adoption will be reached than the others and at least comparable or higher speed after it. Moreover, Swiss hotels are the one that reach the maximum peak the latest in time, 8 months after Malaysian hotels and 17 months after international hotels chains.

Table 2. Comparative BR model coefficients and adoption dates

Sector	Time range	T1	T	T2	*p*	*q*	*q/p*	Source
International hotel chains	Dec-96 Feb-07	Jun 98	Sep 00	Nov 02	.0043	.0455	10.7	hotelsmag. com (July 2006)
Malaysian hotels	Nov-96 Jun-08	Oct 98	Jul 01	Apr 04	.0037	.0364	9.85	this paper
Swiss hotels	Oct-96 Feb-06	Jan 97	Feb 02	Nov 01	.0038	.1381	36.8	(Scaglione, Steiner, Schegg, & Murphy, 2005)

The comparison shows $q/p > 1$, suggesting an S-shaped cumulative distribution as in Fig 3, across all three sectors. Malaysian and International chains have comparable *q/p* ratios, 9.85 and 10.7 respectively, whereas the Swiss is three times greater. All three sectors had similar coefficients of innovation (*p*), but the Swiss hotels had over triple the coefficient of imitation (*q*) relative to Malaysian and international chain hotels. The Swiss hotel's asymmetric adoption distribution curve and their highest *q/p* ratio suggest concerns about competitive pressure, imitation, driving adoption relative to Malaysian and international chain hotels. The Swiss hotels lagged in the formation of a critical mass of adopters relative to the Malaysian and International chain hotels. As a result, the Swiss early majority hotels took three times longer to adopt than did the Swiss late majority category. Critical mass and imitation may have played a more important role for Swiss hotels relative to the Malaysian and international chain hotels.

5 Conclusion and Recommendations

5.1 Academic Implications

Comparing Malaysian hotels Internet diffusion with Swiss and international chain hotels provides a holistic view of Internet diffusion process. In addition, replicating studies into other geographical areas provides insights and helps increase the

generalisability of previous findings and theory on Internet diffusion. This study adds to diffusion literature by identifying total market size and adopter categories using BR model. All three sectors adopted websites at about the same time, yet critical mass and imitation seemed to hinder the early adoption rate and accelerated the late adoption of Swiss hotel websites. The results suggest that 38% of Malaysian hotels may never adopt traditional websites.

5.2 Business Implications

As the Internet becomes essential for operational and strategic purposes (Bai, Law, & Wen, 2008; Nasution & Mavondo, 2008) hotels without the Internet face competitive disadvantages. This study shows that 38%, over one in three Malaysian hotels will never have a website similar to today's traditional websites. However, to ensure the hotels benefit from the technology, the decision to adopt necessitates long and short term planning. However, web usage continues to evolve. These never-to-adopt Malaysian hotels may leap over their colleagues and start their Internet presence with new applications such as a Wordpress blog, Facebook fan page or Googlepage rather than a traditional website. Swiss hospitality sector shows a higher level of parsimony in this adoption those Malaysian hotels.

6 Limitations and Future Research

One limitation of this study is no complete hotel list, rather a census of the 540 hotels registered with the Ministry of Tourism Malaysia. Extending and comparing the population to other locations, particularly developing countries, could improve the generalisability of the results. Future research could also examine other factors relating to Internet leapfrogging such as system legacy (LaRose, 2009) and cloud computing (Knorr & Gruman, 2008). In addition, the analyses rely on the website age provided by the Wayback Machine site. As some sites do not provide access, future research could ask hotels for the website age. Future research could also examine the role of other country-level variables such as Internet users, the evolution of overnights, and Gross Domestic Product as exogenous variables in the *Bass model* (Bass, Jain, & Khishnan 2000; Scaglione, Schegg, Steiner, & Murphy, 2005). This latter research will shed light on the different dimension in the perception of critical mass and allow comparisons across countries.

References

Amir, S. (2004). The Regime and the Airplane: High Technology and Nationlism in Indonesia. *Bulletin of Science Technology Society* 24(2): 107-114.

Anderson, C. & Wolff, M. (2010). *The Web is Dead. Long Live the Internet*. Retrieved September 13, 2011, from http://www.wired.com/magazine/2010/08/ff_webrip/all/1

Andrés, L., Cuberes, D., Diouf, M. & Serebrisky, T. (2010). The Diffusion of the Internet: A Cross-Country Analysis *Telecommunications Policy* 34(5-6): 323-340.

Bai, B., Law, R. & Wen, I. (2008). The Impact of Website Quality on Customer Satisfaction and Purchase Intentions: Evidence from Chinese Online Visitors. *International Journal of Hospitality Management* 27(3): 391-402.

Bass, F. M. (1969). A New Product Growth Model for Consumer Durables. *Management Science* 15(5): 215-227.

Bass, F., Jain, D. & Khishnan, T. (2000), Modeling the Marketing-Mix influence in New-Product Diffusion. In V. Mahajan, E. Muller, and Wind Y. (Eds.), *New Product Diffusion Models*. Boston: Kluwer Academic Publishers.

Da Silveira, G. (2001). Innovation Diffusion: Research Agenda for Developing Economies. *Technovation* 21(12): 767-773.

Davison, R., Vogel, D., Harris, R. & Jones, N. (2000). Technology Leapfrogging in Developing Countries - An Inevitable Luxury? *EJISDC* 1(5): 1-10.

Fichman, R. G. (2000). The Diffusion and Assimilation of Information Technology Innovations. Framing the Domains of IT Management. In R. W. Zmud (Ed.), *Projecting the Future through the Past*. Cincinnati, Ohio: Pinnaflex Publishing.

Fichman, R. G. (2004). Going Beyond the Dominant Paradigm for Information Technology Innovation Research: Emerging Concepts and Methods. *Journal of the AIS* 5(8): 314-355.

Fidler, R. F. (Ed.). (1997). *Mediamorphosis - Understanding New Media*. Thousand Oaks, California: Pine Forge Press.

Gallagher, K. S. (2006). Limits to Leapfrogging in Energy Technologies? Evidence from the Chinese Automobile Industry. *Energy Policy* 34(4): 383-394.

Gray, H. & Sanzogni, L. (2004). Technology Leapfrogging in Thailand: Issues for the Support of E-commerce Infrastructure. *The Electronic Journal on Information Systems in Developing Countries* 16(3): 1-26.

Hanson, W. A. & Kalyanam, K. (2007). *Internet Marketing and e-Commerce*. Mason, Ohio: Thompson South Western.

Hashim, N. H., Murphy, J., Purchase, S. & O'Connor, P. (2009). Website and Email Adoption by Malaysian Hotels. *International Journal of Hospitality Management 29*(1): 1-4.

Hobday, M. (1995). East Asian Latecomer Firms: Learning the Technology of Electronics. *World Development* 23(7): 1171-1193.

Hoffman, D. & Fodor (2010). Can you Measure the ROI of Your Social Media Marketing, *Sloan Management Review* 52(1): 39-49.

Hsu, C.-L., Lu, H.-P. & Hsu, H.-H. (2007). Adoption of the mobile Internet: An empirical study of multimedia message service (MMS). *Omega* 35(6): 715-726.

Jeyaraj, A., Rottman, J. W. & Lacity, M. C. (2006). A Review of Predictors, Linkages, and Biases in IT Innovation Adoption. *Journal of Information Technology* 21: 1-23.

Kauffman, R. J. & Techatassanasoontorn, A. A. (2009). Understanding Early Diffusion od Digital Wireless Phones. *Telecommunications Policy* 33(3): 432-450.

Kim, S. (2011). The Diffusion of the Internet: Trend and Causes. *Social Science Research 40*(2): 602-613.

Kinnunen, J. (1996). Gabriel Tarde as a Founding Father of Innovation Diffusion Research. *Acta Sociologica* 39(44): 431-442.

Knorr, E. & Gruman, G. (2008). What Cloud Computing Really Means. *InfoWorld*. Retrieved from http://www.infoworld.com/d/cloud-computing/what-cloud-computing-really-means-031

LaRose, M. (2009). *Definition of Legacy System*. Retrieved from http://www.ehow.com/about_5175378_definition-legacy-system.html

Liao, C., To, P.-L. & Shih, M.-L. (2006). Website Practices: A comparison between the Top 1000 Companies in the US and Taiwan. *International Joournal of Information Management 26*(3): 196-211.

Lilien, G. L. & Rangaswamy, A. (2002). *Marketing Engineering*. New Jersey: Pearson Education.

Mahajan, V., Muller, E. & Bass, F. M. (1990). New Product Diffusion Models in Marketing: A Review and Directions for Research. *Journal of Marketing* 54(1): 1-26.

Mahler, A. & Rogers, E. M. (1999). The diffusion of interactive communication innovations and the critical mass: the adoption of telecommunications services by German banks *Telecommunications Policy* 23(10-11): 719-740.

Meade, N. & Islam, T. (2006). Modelling and Forecasting the Diffusion of Innovation: A 25-Years Review. *International Journal of Forecasting* 22(3): 519-545.

Murphy, J., Forrest, E. J., Wotring, C. E. & Brymer, R. A. (1996). Hotel Management and Marketing on the Internet. *Cornell Hotel and Restaurant Administration Quarterly* 37(June): 70-82.

Murphy, J., Hashim, N. H. & O'Connor, P. (2007). Take Me Back: Validating the Wayback Machine. *Journal of Computer-Mediated Communication* 13(1). Retrieved from http://jcmc.indiana.edu/vol13/issue11/murphy.html.

Nasution, H. N. & Mavondo, F. T. (2008). Customer Value in the Hotel Industry: What Managers Believe They Deliver and What Customer Experience. *International Journal of Hospitality Management* 27(2): 204-213.

Pan, B. & Fesenmaier, D. R. (2006). Online Information Search: Vacation Planning Process. *Annals of Tourism 3k3*(3): 809-832.

Peres, R., Muller, E. & Mahajan, V. (2010). Innovation Diffusion and New Product Growth Models: A Critical Review and Research Directions. *International Journal of Research in Marketing 27*: 91-106.

Peterson, R. A. (1973). A Note on Optimal Adopter Category Determination. *Journal of Marketing Research* 10(3): 325-329.

Rogers, E. M. (1962). *Diffusion of Innovations* (1st ed.). New York: The Free Press.

Rogers, E. M. (2003). *Diffusion of Innovations* (5th ed.). New York: The Free Press.

Rosenkranz, S. (1997). Quality Improvements and the Incentive to Leapfrog. *International Journal of Industrial Organization* 15(2): 243-261.

SAS Institute Inc. (2011). *SAS/STAT® 9.22 User's Guide*. Cary, North Carolina: SAS Institute Inc.

Scaglione, M., Ismail, A.F., Trabichet, J-P. & Murphy, J. (2010) An investigation of Leapfrogging and Web 2.0 Implementation. In U. Gretzel, R. Law and M. Fuchs (Eds.) *Information and Communication Technologies in Tourism 2010.* Vienna: Springer.

Scaglione, M., Schegg, R. & Murphy, J. (2009). Website Adoption and Sales performance in Valais' Hospitality Industry. *Technovation* 29(9): 625-631.

Schmittlein, D. & Mahajan, V. (1982). Maximum Likelihood Estimation for An Innovation Diffusion Model of New Product Acceptance. *Marketing Science 1*(1): 57-78.

Siguaw, J. A., Enz, C. A. & Namiasivayam, K. (2000). Adoption of Information Technology in U.S. Hotels: Strategically Driven Objectives. *Journal of Travel Research 39*(2): 192-201.

Stafford, T. F. (2003). Differentiating between Adopter Categories in the Uses and Gratifications for Internet Services. *IEEE Transactions on Engineering Management 50*(4): 427-435.

Steinmueller, W. E. (2001). ICTs and the Possibilities for Leapfrogging by Developing Countries. *International Labour Review 140*(2): 193-2001.

Stremersch, S., Muller, E. & Peres, R. (2010). Does New Product Growth Accelerate Across Technology Generations? *Marketing Letters 21*(1): 103-120.

Teo, T. S. H. & Pian, Y. (2004). A Model for Web Adoption. *Information & Management 41*(4): 457-468.

Van den Bulte, C. & Lilien, G. L. (1997). Bias and Systematic Change in the Parameter Estimates of Macro-level Diffusion Models. *Marketing Science 16*(4): 338-353.

Van den Bulte, C. & Stremersch, S. (2004). Social Contagion and Income Heterogeneity in New Product Diffusion: A Meta-Analytic Test. *Marketing Science 23*(4): 530-544.

Xiang, Z., Gretzel, U. & Fesenmaier, D. R. (2009). Semantic Representation of Tourism on the Internet. *Journal of Travel Research* 47(4): 440-453.

Zhu, K. & Kraemer, K. L. (2005). Post Adoption Variations in Usage and Value of E-Business by Organizations: Cross-Country Evidence from the Retail Industry. *Information Systems Research* 16(1): 61-84.

Interacting with a Social Web of Smart Objects for Enhancing Tourist Experiences

The PIEMONTE Team[1]

[a] Dipartimento di Informatica
Università di Torino, Italy
cena@di.unito.it

[b] Telecom Italia R&T, Italy

[c] Università degli Studi di Scienze Gastronomiche, Italy

[d] Slowfood

Abstract

In this paper we introduce the idea of interaction with networks of socially intelligent objects as a way of supporting tourists and introducing them to the culture of a territory and as a way of maintaing the cultural heritage of a territory alive. We illustrate this idea with an application we designed in the field of gastronomy. Socially intelligent objects are able to maintain and aggregate knowledge about themselves and their world and are able to establish social relations with other objects and people. In this way they become hubs which allow tourists to get in touch with the world of the objects which is made of a territory, its culture and traditions, people and other objects. We support natural forms of interaction without requiring any electronic infrastructring of the objects. In particular, we designed an interaction paradigm supporting a playful enhancing experience when interacting with objects. We also support a continuum of experience in the real and virtual world.

Keywords: Social web of things, Social web, Semantic web, Social intelligence, Cultural heritage, Eno-gastronomy

1 Introduction

The vision of the project is to support users in sharing information and experiences in the places where such experiences are lived. The basic idea is to transform objects and places into intelligent social entities that can communicate among them and with people, tell users stories and exchange information with them, aggregate information and knowledge, establish social relations (friendship) with other objects and people and allow people to get in touch with their friends.

On the one hand, this makes it possible to support immersive experience where users can interact with objects, and through this interaction access a variety of information and services in an easy and natural way. On the other hand, in this way objects become a sort of intelligent memory where the cultural heritage of a territory can be preserved and kept alive.

---

[1] The PIEMONTE Team is: F. Cena[a] (contact author: cena@di.unito.it), F. Antonelli[b], G. Biamino[a], F. Carmagnola[a], E. Chiabrando[a,b], L. Console[a], V. Cuciti[b], M. Demichelis[b], F. Fassio[c,d], F. Franceschi[b], R. Furnari[a], C. Gena[a], M. Geymonat[b], P. Grimaldi[c], P. Grillo[a], E. Guercio[b], S. Likavec[a], I. Lombardi[a], D. Mana[b], A. Marcengo[b], M. Mioli[a], M. Mirabelli[b], M. Perrero[a,b], C. Picardi[a], F. Protti[a,b], A. Rapp[b], R. Sandon[c], R. Simeoni[b], D. Theseider Dupre'[a], I. Torre[a], A. Toso[a], F. Torta[d], F. Vernero[a]

The project relies on a number of ingredients:

- Everyday objects. We assume that objects are *not* infrastructured, that is, no sensor or electronics or tag is embedded into objects.
- Techniques for making objects intelligent (*smart*), i.e. making them able to:
  - maintain information and knowledge, aggregate it into coherent "stories" and offer services to users.
  - exchange information with users and integrate information told by users with their knowledge base, and exchange information with other objects.
  - maintain social relations with people and objects; the relation may depend on objects properties, on the information provided by users, on users' behavior, on context (i.e. time or location).
- Social networking and Web 2.0. We allow users to share information but we introduce a notion of *Social Network of Intelligent Objects and People* where people and objects belong to a unique network and they are both active entities. This is different from traditional approaches which consider networks of people in which objects are only the topics of discussion.
- Natural interaction. We developed natural forms of *playful interaction* between people and objects.
- Continuum of experiences. In the interaction we consider multiple devices and support interaction with objects using smartphones (real world) and on the web (virtual world) in a seamless way. In this way interaction may be started on the web, continue in reality and then back again on the web and so forth. As we shall see, this is a key point for tourist applications.

These ideas have been implemented and tested in our application in the gastronomy domain, as a way of getting in touch with the traditions and cultural heritage of a territory. Objects to get in touch with include food products, market stalls, restaurants, shops, recipes, but also geographic places and actors such as cooks, producers, shop owners, etc. Interacting with a food product (e.g., a bottle of wine) is a way of getting in touch with the cultural heritage behind that wine, made of stories and traditions, and with its social network, made of grapes, producers, shops, but also people who talked about it, people who tasted and liked it, recipes that are perfect to taste the wine with, other products that are "friends" of the wine (e.g., a cheese traditionally served with the wine). This is in line with the philosophy of SlowFood: "A new model of sustainable gastronomy must be created, where food is *good, clean and fair* and where people are aware of the territory they live in or visit, of its resources, history, traditions and of the actors and the processes in the chain; a model where biodiversity is preserved and where networks of actors can share experiences." (Petrini, 2006)

This paper describes the applications that we developed for experimenting with these ideas. After introducing a scenario to clarify our goals, we describe the suite of applications we developed: the application supporting mobile interaction between people and objects using smartphones, the application supporting web interaction and the continuum of experiences and finally the backshop application for supporting

stakeholders, providing both registration tools (for entering objects in the system) and analysis tools (for getting feedback about the behavior of the registered objects). In this way, we can close the loop for supporting cultural heritage and for creating direct and sustainable chains. We conclude by discussing the evaluation of the applications.

2 A Scenario

Before the visit. *Anna decides to make a trip to Susa Valley and she uses our application to plan the visit. Anna is interested in gastronomic products, local food and goat cheese in particular, as well as in Romanesque art. She particularly loves exploring the cultural heritage and local traditions, related to art and food. At home, she visits our web application in order to discover the main attractions of the Valley. The application stores Anna's user model, hence it shows her a local farmers' market in Coazze, a small village, having also a very nice handcraft museum and many restaurants where it is possible to taste the local cheeses. Such information has been inserted by the local tourist agency, which used our backshop application to insert the information about events in the Valley, about its cultural heritage and its attractions (museums, monuments, etc), local companies (shops, restaurants, etc.) and their services. Among other things, it has inserted the information about the local farmers' market and the handcraft museum. At the same time, the local producers also used our backshop application to insert the information about their products. For example, Mario Rossi registered himself, and the cheeses he produces in the system.*

During the visit. *When Anna arrives to the Valley, in Coazze she goes to the local farmers' market suggested by the application, and our application on her portable device, acting as a radar, alerts her of the presence of the stall of a renowned producer of Cevrin, a traditional local goat cheese. Anna takes a picture of the cheese label with her smartphone. The cheese is recognized as an object in the system network and Anna is put in contact with its virtual avatar. She discovers that a friend of hers had previously visited it in another location and left some suggestions on what to buy. The cheese also tells her stories on how the farmers used to produce it in the past; the stories were also complemented by other visitors with family anecdotes or historical facts. Anna reads a mash-up of such stories and, upon reading about a traditional dish made with Cevrin, she bookmarks it, remembering that her grandma used to cook it. In addition, she is presented with some interesting objects and people (e.g., people that have a profile close to hers or said something about Cevrin).*

Anna is then presented with the social network of Cevrin, including other cheeses with similar properties (e.g. type of milk), its producers, the information about the places where it is produced, a museum of ancient handcraft related to cheese production, some restaurants where Cevrin can be tasted, a recipe of a salad made with the cheese, and a local wine (Erbaluce di Caluso) that her friends and many other people found perfect to drink with it. Anna follows the link to Erbaluce di Caluso and in the network of the wine she discovers a shop nearby where she can buy it. She decides to visit the wine maker facilities where she buys a bottle. She also discovers that a local restaurant prepares the traditional dish she bookmarked earlier. Anna has no time to eat in the restaurant, but memorizes the address and prefers to visit the handcraft museum. Finally, our application shows her some comments of her friends that suggest to visit a lovely town nearby, where it is possible

to see ancient Romanesque church, a style she particularly likes. Thus, after leaving a positive comment about the museum, she decides to visit this place.

After the visit. *Once home, Anna tastes the wine she bought and finds out she likes it a lot, she agrees with the comments that suggested it together with the goat cheese. She connects to our web application, finds the virtual transposition of the market stall and adds a positive evaluation of the cheese-wine connection. She also adds positive comments about the handcraft museum and the little town, suggesting to other people to visit its lovely church. Moreover, Mario Rossi, using our backshop application, can read the comments about his products and discover new relationships among his products, other products and other producers. At the same time, the local tourist agency can use our backshop application to discover the most visited attractions and what people liked the most. In this way, it is possible to create tourists' profiles in order to better organize the tourist attractions and offer personalized tours.*

This story puts in action the key ideas of the project. The first key idea is the one of an *enhanced intelligent thing*. Things are enhanced in their capability to interact with people *and* other objects. They display intelligent behavior: they have knowledge about themselves and about the world around them and they can tell people about this world. They can tailor the interaction to the specific person interacting with them. For example, the cheese knows who Anna is and what to tell her about itself, e.g., it knows that Anna likes the salad. The enhanced intelligent things collect facts and user comments about themselves, aggregate these comments and use them to build their own social relations and their own public images.

Hence, objects have *social behavior*. This introduces the second key issue, the *mixed social networks*. Rather than having a social network of users that discuss and express opinions on objects, we propose a social network of people *and* objects, where the two categories intermingle. The social networks of a cheese include objects such as similar products of the same typology (other cheeses with similar properties), related products (such as a wine that some other users found perfect to drink with it), the cheese's producers, touristic information about the place where it is produced, related tourist attractions (the museum of ancient handcraft related to cheese production), restaurants where the cheese is served, some recipes made with the cheese, etc.

The networks are created and managed in different ways, exploiting knowledge about the domain and taking into account the behavior of users, as we will discuss later. Things have an active social role in the management of these networks and can establish social relationships among themselves and with people: the cheese becomes a friend of local wine after many people said that they marry well together. It is important to notice at this point that in this application we are interested in generic objects (such as Cevrin cheese or the Erbaluce di Caluso wine) and not in individual instances (a specific bottle of wine or a specific wheel of cheese). We are interested in the concept of the cheese, its relations with other products, what people say about it and what it can narrate to the people. In this sense all the interactions with cheese wheels are the interactions with the same thing, the cheese itself.

The third important issue is *natural playful personalized interaction* between users and enhanced objects to create a wholesome experience. Users should not perceive the use of the system as something that distracts them from their real-life experience,

rather as its enhancement. Examples of this are the possibility to access the system by taking a picture, as Anna does with the cheese label, as well as recording comments, rather than typing them. Secondly, interaction must be tailored to each user and the interaction with each user is personalized, taking into account user's individual interests and preferences. Moreover, the system can create links between users based on similarities of their profiles. This is an important issue for objects to display an intelligent behavior. A further ingredient we offer to the user is a *continuum of experience*, from a virtual experience on the web to a real experience interacting with objects, returning to the web again and so forth. This enhances the user experience.

The territory and stakeholders can get a lot of feedback from the users' experience and this is valuable effect of the use of our framework which can contribute positively to the tourist growth of the territory and to sustainability in the relationships between tourists (customers) and the territory and its stakeholders and to maintaining the cultural heritage of the territory alive.

3 Our Framework

Our framework is a suite of applications we designed for achieving the goals discussed above. In the following we briefly discuss the three applications.

3.1 Mobile Application

Our Apple iPhone mobile application introduces a novel and peculiar paradigm for supporting the user interaction with social networks of smart objects. This interaction consists of two main phases: (i) getting in touch and (ii) interacting.

Getting in touch. A basic assumption of our project is that infrastructuring of the environment must be minimized. We aim at supporting interaction with everyday objects, with no embedded electronics or tags. Thus, we developed the following ways of establishing the contact between a user and a thing (fig. 1(a)):

- *Taking a picture*. The user frames the label of a product with the camera; the label is recognized and a contact between the user and the thing is created (fig. 1(b)).
- *Geo-positioning*. The user can start the navigation by getting in touch with a place (e.g., Coazze village in the scenario above) and thus with the objects related to the place, i.e., the objects around him.
- *Bookmarking*. The user can create a list of "preferred" objects and start the navigation from one of them. In other words, the user can have a list of friends including users and objects and can get in touch with one of them at any time.
- *Searching for objects.*

We also consider the possibility that a thing calls the user. If the application is turned on and a user gets close to a thing that the recommender considers relevant for him/her, the system can suggest him/her to get in touch with this thing.

Interacting with the thing and its world: the wheel. Once a contact with a thing has been established, the user can interact with it and access its social network. Since we aim at using objects as gateways for accessing the cultural heritage of a territory, we

designed an interaction model which allows users to explore the world starting from a contacted thing. We developed a *"wheel" model* (fig. 1(c)), where the wheel can be seen as the main square of a village, i.e., the traditional place for meeting. In this place the user can interact with the thing and its friends, exchanging information and knowledge, being introduced to and exploring their social networks.

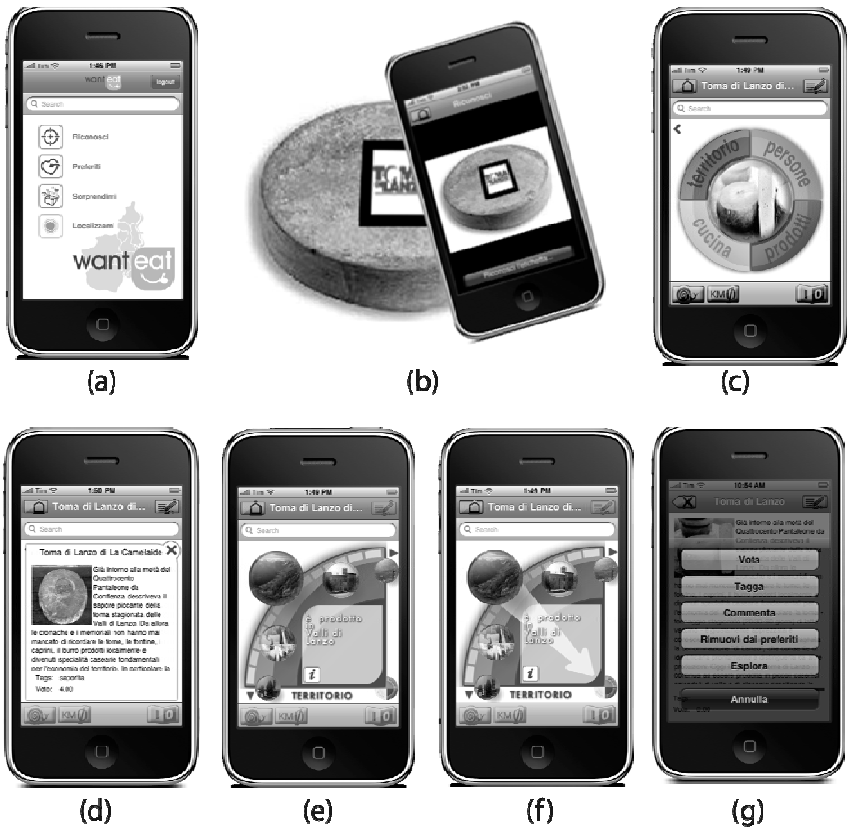


Fig. 1. Example of the wheel on the iPhone

The thing the user is interacting with is in the centre of the wheel. By touching the thing in the center of the wheel, the user can get in touch with it. The thing tells the user about itself, providing both the general knowledge and the information synthesized from the interaction with other people (including tags, comments, votes), see fig. 1(d). The user can communicate with the thing, as well as add tags, comments and votes or bookmark the thing, adding it to his/her list of friends (fig. 1(g)).

The thing in focus is surrounded by the social network of its friends (fig. 1(c)), which contains both people (including users) and other objects. Each friend belongs to one of the four sectors (which depend on the thing of the center). In the example in fig. 1(c), the thing in focus is a food product; the first sector of the wheel ("Territorio"

– "Territory") contains the friends related to the territory, the production and supply chain of the thing in focus (e.g., producers, shops, production places, etc.). The sector "Persone" ("People") contains people that are friends of the thing in focus (e.g., who bookmarked it); the sector "Prodotti" ("Products") contains other food products that are friends of the thing in focus (e.g., a wine that goes well with a cheese); the sector "Cucina" ("Kitchen") contains entities related to the cuisine, such as cooks, restaurants, recipes. Each sector can be expanded by touching it as shown in fig. 1(e). The items can be explored by rotating the wheel, as an old style telephone dialing selector. One item at a time is highlighted and its relation to the thing at the center is explained. The user can continue exploration by changing the thing in the focus. This can be done by dragging one thing towards the small wheel in one of the corners (fig. 1(f)). At this point the whole wheel is recomputed and displayed to the user.

4 Behind the Scene: the Framework Architecture

In this section we sketch the architecture of our framework, focusing on the modules that support social intelligence. The framework is partially distributed including: (i) the *"server side"* in charge of the creation and management of the network of objects and people and containing the logic for creating intelligent adaptive behavior; (ii) the *"client side"*, running on the user's smartphone, in charge of managing the interaction discussed above; (iii) the *infrastructure* for interacting with objects.

Server Side

Five main modules constitute the server side of our framework: (i) knowledge base manager; (ii) social network manager; (iii) user profile manager; (iv) recommender and (v) interaction manager.

The *Knowledge base manager* exploits the *ontology*, defining all the concepts in the domain and the relations between them. For example, the ontology includes a taxonomy of gastronomic products (e.g., different types of cheeses based on different properties, like the type of milk and the production techniques), linking them to actors (such as producers or market stall owners) and to places in the geographic ontology. Inferences on the ontology allow the system (i) to associate inferred properties to objects; (ii) to create links between objects. The former are part of the description provided to users; the latter contribute to the creation of social networks of objects. Things that are related in the ontology (e.g., originating from the same place or being produced in similar ways) can in fact be linked. Other links may be created by rules added to the ontology. Examples of these rules are those associating wines and foods, based on their properties (e.g., aged red wines with meat dishes).

Dynamic information is associated to a thing by users in the form of ratings, tags, comments. The system aggregates these pieces of information in different ways. Clouds are generated from the tags, while comments are aggregated based on the profile of the users who provided them. For example, the system will present to a user only the comments/tags/ratings provided by her/his friends or by the users similar to him/her. We re also exploring "digital storytelling" techniques for generating stories starting from the pieces of information provided by the users and their actions. The *Social network manager* is in charge of managing networks of objects and people.

The links between items are computed in two ways: (i) links derived from inferences in the ontology (e.g., linking products and places or linking products with similar properties) and (ii) links deriving from the behavior of users. Links may connect users to other users or objects or objects to other users or objects, thus creating mixed networks of users and objects. Links may depend on the properties of the entities, linking entities with many common properties. For example, this may lead to linking two cheeses that are similar or a cheese and a wine or a museum that belong to the same territory. Links may also depend on the users behavior. These links are created by the system periodically after an analysis of the users' behavior and interaction with objects. An item X is linked to another item Y in case, for example, X has been mentioned as a tag or in a comment on Y, or in case X and Y appear frequently together in the bookmarks of several users, or in case X and Y have been visited frequently on the same occasion by several users. These links are created by tracking and analyzing the actions performed by the users and looking for significant correlations between objects. The same approach can be used to create further links between users and objects whenever the name of a user is frequently associated by other users to a thing (or vice-versa).

The system we designed is *adaptive* and provides information and services tailored to the specific user and context of usage (Brusilovsky et al., 2007). The *User modeling* module maintains a model of each user which reports the user's interests with respect to the concepts in the domain ontology. The model is built and updated considering the actions performed by the user and inducing his/her interests from his/her behavior (see Carmagnola et al. (2008)). The user model is also the basis for computing similarities between users as the distance between their models and thus to create links between users with similar models.

Our system exploits a *Recommender* which in turn exploits the user model and the context of interaction (e.g., the location or time of the interaction) and suggests potentially interesting objects to a user. Moreover, whenever many items have to be presented to the user (e.g., in a wheel sector or as a result of a search), the recommender contributes to ranking them. However, the policy we used is a complex one, taking into account not only the ranking of the recommender but also the type of objects and the location, as well as some randomness factors in order to support a serendipitous exploration.

Finally, the *Interaction manager/dialog manager* module manages the interaction with the user client; it processes the user requests and dispatches them to the other modules; it then combines the answers generated by the modules above, depending on the user's request, and sends them to the client.

Client Side

The client side of our mobile application has been implemented on Apple iPhone and is in charge of (i) managing the interaction between a user and objects and (ii) creating the interaction environment discussed in previous sections (the "wheel"), sending requests to the server (interaction manager) and processing the answers to generate the appropriate way of presenting them. Taking a picture is the main method provided to users for getting in touch with objects. The clients for our web application and our backshop application are implemented as web applications.

Infrastructure and Implementation.

Minimizing all infrastructures is an important goal of the project. The recognition of the label of a thing is thus the main way to create a contact with the user and this does not require intervention on the objects. Regarding the location, we assume some objects such as shops or restaurants are localized while the user's position is known due to the GPS included in Apple iPhone. Regarding the implementation of the framework, we decided to centralize the core of the system (agents of the server side); this choice has been driven by the fact that we aim at interacting with concepts, via their instances (see the discussion above) and we are not interested in tracking or interacting with individuals.

It is worth noting that the amount of information to be exchanged by the server and the mobile device is limited and can be easily managed using standard web service approaches and 3G mobile telecommunication.

5 Evaluation with Users

We adopted a user-centered approach in the design of the interaction model. Thus we performed user tests starting at the early stages of the project. The first prototype of our mobile application has been tested with 12 users who have been asked to perform a number of tasks, ranging from thing recognition to navigation with the wheel. All the users liked the service; no major usability problems were pointed out; the users evidenced some minor issues that have been fixed in the final prototype. The users also provided interesting comments about the advantages of using the application while visiting an area or while visiting a market for discovering products.

In September 2010 we populated our system with information about "Provincia di Torino" (Turin region) and its gastronomy. A test of the system has been performed afterwards with two main goals: checking the stability and efficiency of the server and validating the consistency of data entry. In October 2010 the system was presented and tested by a wide public at Salone del Gusto. During 5 days of Salone del Gusto the application has been used by several hundreds of casual users, who were very enthusiastic about the services and opportunities offered by our application. We collected 684 structured interviews from users who evaluated the system.

The results of the evaluation were very positive about both the experience of interacting with objects (or better their social network) to access knowledge about a territory and the novel interaction model we proposed. The evaluation with users covered many aspects. More than 80% of the users found the idea of using our mobile application for accessing cultural heritage information useful; 78% of the user would use the application, if available when visiting a place and 89% of them would recommend the application. 90% of the users appreciated the idea of interacting with objects and the interface we developed was considered easy to use, comprehensible, pleasant and playful.

In particular, on a scale from 1 to 4 the average was 3,42 (s.d. 0,68 for *ease of use*, 3,46 (s.d. 0,67) for *comprehensibility*; 3,43 (s.d. 0,67) for *pleasantness*; 3,12 (s.d. 0,84) for *usefulness*. Most users were very positive about all the aspects.

The administration tools proved to be very interesting for stakeholders. They support the generation of synthesized information about how a thing is perceived and more generally how a territory is perceived. Producers, as well as managers of the trial territory, appreciated the richness of information and of the knowledge that can be obtained by the business intelligence tools. Almost all the producers involved in the experimentation gave a positive evaluation of the tools; many others that met us at the Salone del Gusto asked to participate and upload the information about their products.

6 Related Work

Our project carries on innovative research aimed at providing users with an immersive experience where they can interact with objects as living entities that make them access a variety of information about the cultural heritage in a natural way. For this purpose, the project merges recent research from different promising areas like embodied and natural human-computer interaction, mixed reality, ambient intelligence, adaptive systems.

We exploit ideas from *mixed and augmented reality*, which propose to overlay the real-world environment with digital objects. For example, *Magic Book* (Billinghurst et al., 2001) visualizes virtual 3D models on the page of a physical book, which acts as a handle: moving the book, models can be moved and observed from different points of view.

Other interesting applications of augmented reality can guide visitors inside a building (Reitmayr & Schmalstieg, 2003) or guide tourists in a city (Feiner et al., 1997). The latter is an early example of mobile augmented reality that, due to the recent advances in capabilities of handheld platforms, is gaining significant interest. In this area, a relevant application is *Wikireality* (Gray et al., 2009), a system exploiting GPS data and live camera to query Wikipedia through photos of the surrounding landscape.

The project is related to *adaptive systems exploiting user models* (Brusilovsky et al., 2007), providing a novel area of application. It also elaborates on the notions of social networking, extending it in different ways with respect to the common social applications. Our approach is innovative in the fact that we create mixed networks involving people and objects. The approach we adopted for managing the user model and for social recommendation is derived from Carmagnola et al. (2008).

Finally, our project exploits social networks of intelligent objects to introduce tourists to the culture of a territory. Considering the application domains we have chosen, we have to mention some further related work. In the cultural heritage domain, most of the related projects concern mobile guides that put the user and her needs at the focus of attention and provide location based services.

Among them we mention Peach (Stock et al., 2007), a large, interdisciplinary development project aimed at experimenting with new advanced technologies to enhance cultural heritage appreciation; Carletto (Damiano et al., 2006), a virtual and dramatized guide for the Savoy apartments in Turin; Cicero (Mantyjarvi et al., 2006), a location-aware application developed for the Marble Museum located in Carrara.

These guides can be easily consulted in a context-dependent way through different kinds of natural interaction modalities, such as multimedia presentation (Stock et al., 2007), the presence of animated and conversational agents (Damiano et al., 2006) and scan and tilt modalities (Mantyjarvi et al., 2006). However, these works do not offer users a comfortable interaction experience, due to the limited dimensions of the screens, the use of special input devices or unusual gestures that make the interaction complex and unnatural, problems that our project wants to overcome. None of them uses the concept of social network of intelligent objects to discover cultural heritage.

Related to the gastronomic domain, the most relevant projects to mention concern systems that allow the traceability and identification of food goods. Most of them use the RFID technology: (Gandino et al., 2007), for example, is a system that traces food in the agri-food chain. These systems are related to our project as providers of enabling technologies, which can be used to allow our Enhanced Objects to automatically acquire information related to the history of food, such as the ingredients of a product, and integrate it with information provided by users and intermediaries.

Other works concern the identification of food goods in shops and markets with different uses (Buser, 2007). Buser exploits the RFID technology to identify products and provide recommendations to customers in conventional grocery, with a personal shopping assistant attached to the shopping basket. The strategies used in virtual online groceries are analyzed and borrowed in the physical brick and mortar grocery store. In our project we go further, creating synergies and exchanges between virtual and physical spaces; moreover the user can interact with the Enhanced Objects in several ways, providing comments, votes, etc. and not just receiving recommendations. Aggregated user trails can then be used to allow social navigation support, as in another domain related project Kalas, a social navigation system for food recipes (Svensson et al., 2005) which provides evidence that users are more influenced by other users' comments than by those of the recipes authors.

7 Conclusions

We introduced the idea of interacting with a social web of smart objects to get in touch with the cultural heritage of a place. We discussed an interaction paradigm for supporting this idea, and we implemented and tested a prototype in the field of gastronomy. This novel way of aggregating and structuring the knowledge and experiences and the possibility to share it with users is a powerful paradigm which can be employed in many different domains, reusing the approach to knowledge representation and organization. The new interaction metaphors can bring playful and innovative aspects to other applications. The concept of a social network of smart things and people is a promissing environment for exploring further industrial and academic areas.

References

Billinghurst, M., Kato, H. & Poupyrev, I. (2001). The MagicBook: a Transitional AR Interface. *Computers and Graphics* 25(5): 745-753.

Brusilovsky, P., Kobsa, A. & Nejdl, W. (Eds.) (2007). *The Adaptive Web, Methods and*

Strategies of Web Personalization, vol. 4321 of LNCS. Springer.

Buser, D. C. (2007). Context-based recommender systems in conventional grocery–an economic analysis. In *40th Hawaii Int. Conf. on System Sciences, IEEE*.

Carmagnola, F., Cena, F., Console, L., Cortassa, O., Gena, C., Goy, A., Torre, I., Toso, A. & Vernero, F. (2008). Tag-Based User Modeling for Social Multi-Device Adaptive Guides. *User Modeling and User-Adapted Interaction* 18: 497-538.

Damiano, R., Lombardo, V., Pizzo, A. & Nunnari, F. (2006). Dramatization meets narrative presentations. In *Proceeding of the 2006 conference on ECAI 2006: 17th European Conference on Artificial Intelligence* August 29 – September 1, 2006, Riva del Garda, Italy. Amsterdam, The Netherlands: IOS Press.

Feiner, S., MacIntyre, B., Höllerer, T. & Webster, A. (1997). A Touring Machine: Prototyping 3D Mobile Augmented Reality Systems for Exploring the Urban Environment. In *Proc. of 1st IEEE Intern. Symp. on Wearable Computers*.

Gandino, F., Montrucchio, B., Rebaudengo, M. & R, S. E. (2007). Analysis of an rfid-based information system for tracking and tracing in an agri-food chain. In *Proc. First Annual RFID Eurasia Conf.*, *IEEE*.

Gray, D., Kozintsev, I., Wu, Y. & Haussecker, H. (2009). Wikireality: Augmenting Reality with Community Driven Websites. In *Proc. of 2009 IEEE Int. Conf. on Multimedia and Expo*.

Mantyjarvi, J., Paterno, F., Salvador, Z. & Santoro, C. (2006). Scan and tilt: towards natural interaction for mobile museum guides. In *Proceedings of the 8th Conference on Human-Computer interaction with Mobile Devices and Services, MobileHCI '06*.

Petrini, C. (2006). Slow Food Nation: why our food should be good, clean and fair. Rizzoli.

Reitmayr, G. & Schmalstieg, D. (2003). Location Based Applications for Mobile Augmented Reality. In *Proc. of 4th Australasian Conf. on User Interfaces*.

Stock, O., Zancanaro, M., Busetta, P., Callaway, C., Krüger, A., Kruppa, M., Kuflik, T., Not, E. & Rocchi, C. (2007). Adaptive, intelligent presentation of information for the museum visitor in peach. *User Modeling and User-Adapted Interaction* 17: 257–304.

Svensson, M., Hook, K. & Coster, R. (2005). Designing and evaluating kalas: A social navigation system for food recipes. *ACM Trans. Comput.-Hum. Interact* 3(12): 374–400.

Flashpacking: A Discussion of Independent Travel in a Digital World

Cody Morris Paris

Middlesex University Dubai, United Arab Emirates
c.paris@mdx.ac

Abstract

The purpose of this paper is to develop an understanding of the impact of the convergence of ICT and independent travel from the perspective of a small group of tech-savvy flashpackers. Key informant e-interviews were conducted with a small group of eight flashpackers with an established virtual presence. This paper employs the mobilities paradigm to discuss the convergence of technology and backpacking tourism. The study of the flashpacking phenomenon through a 'mobilities lens' provides a vehicle for developing an understanding of the complexities of the intersection of physical travel, technologies, and daily life. Several themes emerged from the interviews including the conceptualization of connection and disconnection, impact of technologies on traditional physical spaces of backpacking, social interaction, and increasing fluidity of 'tourist' and 'non-tourist' experiences.

Keywords: Social Media, Backpacking, Technology in Society, Hybrid Spaces.

1 Introduction

The backpacker phenomenon provides a context through which critical questions about the relationship between contemporary society, tourism, and technology can be discussed. Recently the emergence of the flashpacker has been embraced as an increasingly important sub-segment of backpacking both in the academic literature and the tourism industry (Jarvis & Peel, 2010). The flashpacker has been viewed as a 'key constituent of contemporary society' emerging from the economic, demographic, technological, and social changes in the world (Hannam & Diekmann, 2010). Flashpackers' travel experiences are mediated by communications technology and technological devices (Paris, 2010; Hannam & Diekmann, 2010).

Flashpackers seemingly embody both the backpacker culture and that of the 'digital nomad' (Makimoto & Manners, 1997), often living a location independent lifestyle. Flashpackers could also be considered as part of a new global elite (Bauman, 2000; 2007). These elite are hypermobile mentally, corporeally, and virtually. They have the means to move fluidly across the globe through the various travelscapes using the 'nomadic institutional structure' and have the ability to connect instantly with multiple networks from virtually anywhere through an array of mobile technologies (O'Regan, 2008).

Hannam and Diekmann (2010) note that the emergence of the flashpacker "exemplifies the changing demographics in western societies where older age at marriage, older age having children, increased affluence and new technological developments, alongside increased holiday and leisure time have all come together" (p. 1-2). In Bauman's (2000; 2007) discussion of 'liquid modernity' he suggest that nomadism is a universal trait of the liquid modern person. Research concerning the relationship between technology and society needs to adapt in order to address the

complexities of liquid modernity. Technological innovations have transformed the daily lives of individuals and communities worldwide. Several recent studies indicate that backpackers are very active users of communication technologies, particularly social media (Paris, 2010), and the emergence of the 'flashpacker' sub-segment is representative of the increasing fluidity of the travel experience and daily life caused by the proliferation of information and communication technologies (ICT).

Increasingly, the tourism experience is mediated by ICT (Xiang and Gretzel, 2009; Tussyadiah & Fesenmaier, 2009). This mediation occurs before, during, and after an individual's trip (Paris, 2010). Tourism products are booked and information is collected via social media and e-word-of-mouth (Litvin, Goldsmith, & Pan, 2007) prior to the trip (Xiang & Gretzel, 2009). Individuals maintain connections with home, collect information, book travel, upload photos, blogs, twitter, Facebook, and download travel guides while travelling. After the trip, individuals also use social media and other technologies to portray, reconstruct and relive their trips (Xiang & Gretzel, 2009; Pudliner, 2007). Some have argued that technology can detract from tourist experiences, which are emphasized by the contrast to everyday life (Uriely, 2005). The distinction between tourist experiences and home experiences has blurred, and now experiences are more liquid as tourists experiences can flow through virtual networks and thus are accessible even during everyday life without the necessity of physical movement (Uriely, 2005).

New virtual moorings (Hannam, Sheller, & Urry, 2006) have developed that allow backpackers to be fully integrated in their multiple networks and maintain a sustained state of co-presence between the backpacker culture and their home culture (Mascheroni, 2007). Backpackers manage their multiple networks while travelling and at home through social networking sites, email, and other technologies, which have blurred the boundaries between home and away. The close virtual proximity that backpackers maintain allows them to be able to be instantly in contact with friends, family, work, school, and fellow travellers. Further, the backpacker ideals of independence, freedom, and physical travel are all enhanced and impaired by the virtual mobility of backpacker information, identities, and culture.

The objective of this study is to examine the convergence of information technology and backpacking and develop an understanding of some of the socio-technical implications of social media and other technologies for the backpacking experience and backpacking culture. Understanding how technology has affected flashpackers' backpacking experience could provide a snapshot into one how technological innovations are impacting one aspect of life: travel and tourism.

2 Theoretical Background

The New Mobilities Paradigm (Hannam, Sheller, & Urry, 2006) seeks to establish a movement-driven social science. The mobilities paradigm examines the increasingly complex world through the exploration of the varying levels of movement of people, ideas, objects, and information. The mobilities paradigm provides a strong basis for understanding how the advancement of Internet and related technologies have been incorporated into the daily lives of people around the world.

This paper employs the mobilities paradigm to discuss the convergence of technology and backpacking tourism. The study of the flashpacking phenomenon through a 'mobilities lens' provides a vehicle for developing an understanding of the complexities of the intersection of physical travel, technologies, and daily life.

The recent innovations of the Internet and communications technologies have created a more networked patterning of social life, home life, and work life (Hannam, Sheller, & Urry, 2006). No longer are the simple discussions of the opposition of 'real' vs. 'virtual' or online vs. offline relevant, instead it is more useful to examine the hybridity between physical spaces and cyberspace. These technologies have allowed many people to maintain intermittent co-presence with these networks. Co-presence is further enhanced by 'virtual travel' as many social interactions need to take place over long distances, where corporeal travel is not as easy. This virtual proximity is proliferated by advances in cyberspace, including email, social networking sites, blogs, and other virtual extensions of personal identity. The virtual proximity of an individual's multiple networks allow them to shift easily between or simultaneously interact with more than one network. In the increasingly complex world, where people need to maintain close networks over large geographical distances, virtual mobility allows for the strengthening of interactions (Urry, 2002). The virtual mobility of personal networks allows people to connect to their networks anywhere and at anytime, especially with advances in personal wireless technologies (Hannam, Sheller, & Urry, 2006). The spatial division between 'home and away' is now less important, which allows people greater flexibility with concerns to their movements through time and space. The profusion of information available and adoption of e-commerce by travel service providers have made independent travel much easier.

A number of studies have examined aspects of the virtual mobility phenomenon with specific reference to backpackers. For example, Mascheroni (2007) explored the convergence of new communication media, the Internet and mobile phones, and travel by backpackers. The author concluded that "global nomads produce and maintain mobile spaces of sociality, founded on a complex interaction of face-to-face interaction and mediated communication, co-presence and virtual proximity, corporeal travel and virtual mobilities" (p.527). Backpackers have increasingly used email and social networking sites to stay in contact with fellow backpackers met during their trip (Paris, 2010), and their social network thus becomes accessible anywhere at any time (Mascheroni, 2007). An increasing number of backpackers are carrying laptops, smart phones and wireless access in hostels and guesthouses, as well as cafés, restaurants, and bars in backpacker enclaves provide an almost constant connectivity of the Internet while travelling.

The interaction with fellow travellers in online communities has provided backpackers with a useful source of travel information. Many backpackers use narrative emails and/or travelogues via personal websites, blogs, or social networking sites to share their experiences with friends and family back home, as well as fellow travellers. Mascharoni (2007) points out that travelogues allow backpackers to maintain mobile spaces of sociality that follow individuals during their trip, are constantly updated, and with an email address that represents the only permanent address of the traveller. The mobile sociality allows backpackers to maintain connections with fluid social networks made up of friends, family, travellers met

while on the road, local people, and unknown travellers (Paris, 2011; Mascheroni, 2007). Against this brief backdrop of literature, the purpose of this paper is to develop an understanding of the impact of the convergence of ICT and independent travel from the perspective of a small group of tech-savvy flashpackers.

3 Methods

Key informant e-interviews were conducted with a small group of eight flashpackers with an established virtual presence. Potential individuals were contacted using a snowball sampling procedure. This group was selected as part of a larger research project that also sought to examine the online social interactions among the small group of individuals, thus pre-existing connections were assured through the sampling method. An initial contact, who is an active backpacker and maintains a travel blog, YouTube account, Twitter account, Facebook account, was contacted. The initial key informant was asked to recommend other backpackers who were tech savvy and actively contribute to the production of online content. Five additional individuals were contacted and asked to participate and recommend other potential participants, who then recommended a total of 10 other individuals.

All fifteen individuals were screened, with only individuals who actively maintained a minimum of three of the following were invited to participate: a blog, Facebook profile, Twitter and YouTube account. Eleven individuals met this requirement and were then sent an email explaining the study. Eight of the eleven individuals agreed to be interviewed. The interviewees ranged in age from 23-45, and were mostly from North America or Europe, with the exception of one Brazilian and one Australian. Six of the respondents were male and two were female. Three of the respondents were married, and two were currently travelling with their spouse. Four of the respondents were currently travelling on long-term trips of more than a year, of which two earned their primary income through their online backpacker blogs and as free-lance travel writers. Two of the respondents were recent university graduates, and another two were employed in the IT industry. The semi-structured interviews were conducted via email and Skype. Questions focused on how impacts of information and communication technologies have been incorporated with and changed the backpacker experience. In order to preserve the anonymity, the names of each individual have been changed. Each of the interviews were transcribed and then subjected to a thematic analysis. The following section discusses the main findings.

4 Discussion of Results

Several themes emerged from the interviews including the conceptualization of connection and disconnection, impact of technologies on traditional physical spaces of backpacking, social interaction, and increasing fluidity of ‘tourist’ and ‘non-tourist’ experiences.

4.1 Hybridization of Backpacker Spaces

Backpacker enclaves, made up of hostels, restaurants, and bars, provide the spaces for meaningful interactions, communication, expression of shared values, and the backpacker travel identity (Sorensen, 2003; Murphy, 2001). Backpacking literature

often depicts backpacker enclaves as meta-spaces that provide a space for adjustment, reduced culture shock, respite from life on the road, perceived control, and often provide comforts of home. These spaces also include what Molz (2006) referred to as a 'system of surveillance.' The system of surveillance is the hybridizing of the physical and virtual spaces of backpacking. Backpackers 'surveil' themselves by documenting their experiences for others to see using connections such as internet café's, Wi-Fi, or mobile phones. Increasingly these connections are offered free of charge within a backpacker enclave. This hyper-connective ability has also created some experiential conflicts in the physical destination. Brandon (respondent 6) reflected on a recent hostel experience:

> I've sat in hostel common rooms where 10 backpackers were silently staring at screens gathering information about the city they are in on Twitter rather than talking to each other, meeting new friends, and sharing information through the 'traveller network' that is right in the room. I can't count the times that I have sat next to a stranger on a bus that I would have spoken to and interacted with had they not spent the entire journey playing with iPhone apps.

The hybridization of backpacker spaces facilitated by developments in information and communication technologies has created a mobile sociality that exists virtually and physically. Social media offer individuals a place for co-presence and interaction with various networks, but also are spaces of expression through digital media. Many backpackers maintain social connections with people they once met physically, as strangers, as well as contacts in future destinations. All of these contacts are virtually proximate to each other allowing for backpackers to interact instantly and simultaneously with people from all over the world. As Sara (respondent 4) stated,

> The rise in social media has fostered a community where it is incredibly simple to find people to meet up with and keep in touch with. I can find a couch to crash on in New Zealand, ask a friend for travel recommendations on Vietnam and email that Canadian kid I met in Bosnia last summer within the span of 5 minutes.

4.2 Social Interactions

Each of the individuals interviewed did agree that innovations in communication technologies have affected the backpacking culture, especially in terms of the social interactions. Alan (respondent 5) claimed that these innovations have 'revolutionized' backpacker culture in terms of social interactions while on the road. Alan (respondent 5) further described traditional backpackers' social interactions on the road as, "…transient in nature. You met, you socialized, you enjoyed each other's company and then you parted ways. If lucky you'd bump into each other later down the road in a different city or country, but for the most part good-bye was just that." This account of the fleeting interactions of backpackers is echoed in the backpacker literature (Murphy, 2001). Recent developments have also been 'revolutionary' by increasing the ability for individuals to 'meet-up'. Alan (respondent 5) recounted a personal experience of when he had planned to meet a friend at the McDonalds in the train station in Florence several years ago: "I didn't have a cell phone at the time, making timing of the utmost importance. Unfortunately, not only was there a train strike, but

the train station and immediate area had two separate McDonalds." He ended up spending the majority of the day and his patience trying to find his friend. Nowadays, social media developments, Wi-Fi, Internet access, and mobile phones have decreased this hassle for some.

While social media, and in particular social networking sites, are used to maintain social relationships, most are primarily focused at already established relationships. Backpacking, traditionally, has consisted of fleeting social interactions with individuals while travelling. With the rise of social networking sites, like Facebook, maintaining connection with individuals met on the road is easier, but yet sometimes the relationships do not develop. As Don (respondent 3) noted, "People are relying heavily on Facebook to stay connected, especially with other travellers met along the way, but after time passes, unless the connection was really strong in person, then usually those people fade into the background and you'll never write them or hear from them again." Virtual relationships through Facebook are intimate and must be reciprocally nurtured.

Other social networking sites, however, are focused on providing a space for individuals to meet strangers and build relationships. This reciprocity is also evident in social networking sites that are meant to facilitate off-line connections, such as Couchsufing.com, a site that provides a space for backpackers to meet individuals who are both strangers and locals that wish to share their own personal space (couch) with other strangers. This community is unique in that it requires individuals to also commit to allowing fellow couchsurfers crash on their own couches at some point. Even though there are social networking sites directed at facilitating both on-line and off-line interactions, the strength of these relationships are dependent upon the effort put into them by individuals. The level of effort in maintaining online or offline relationships could be related to each individual's value of the mode of the relationship.

4.3 (Dis)Connection

The innovations of social media have changed the way backpackers communicate with home. Another one of the interviewees traced these innovations through the travel experience of his parents and himself. When his parents over-landed from Scotland to New Zealand, his grandparents could only expect an occasional postcard or letter, where as now when he travels his parents, friends, and other networks can follow his blog, view photos uploaded onto Flikr.com, and interact on Twitter and Facebook. Additionally Chris (respondent 2) produces podcasts from the road, both as guides for other travellers, but also as an auditory story telling for his followers. For some backpackers, who travelled pre-Web 2.0, a dramatic change in both communication and experience can be traced. Don (respondent 3) reflected on a backpacking trip in 1998 around Europe:

> I kept a written journal of my experiences and after 2 months home, I taught myself HTML and put it all up on a website. It was a long and slow process, but even strangers would find it and read along. There was no Facebook—you lived in the moment, sent postcards to people at home. Now social media, blogs, and quality internet connections around the world make it incredibly easy to update people on your travels and experiences. Plenty of

> backpackers do not keep blogs, but simply post updates and photos to Facebook from time to time. It's the norm.

These increased innovations in communication have provided a 'safety-net'. Instantaneous communication with people anywhere in the world can reduce the perception of risk of independent travel. This can be particularly true for women backpackers travelling on their own. Sara (respondent 4) said that, "it's very important for me to keep in touch with my immediate family. I want them to know where I am and that I'm safe. I usually email them daily or whenever I have internet connection. My friends usually just follow my blog." Sara (respondent 4) uses email, a much more private form of communication, to maintain daily contact with her family. On the other hand, she uses her blog to share her experiences with her friends. Homesickness is often something travellers experience at some point during their trips, especially long-term or trips to destinations far different than an individual's home. Social media allows for instant contact and thus, as Alan (respondent 5) pointed out, "often helps to alleviate/reduce homesickness. It also dramatically reduces the catch up time when re-integrating at the end of a trip."

Social media also provide a means for friends and family to seek out news and help when crises do occur. Previously, the ability to get news updates, and mobilize searches was nearly impossible. With contact only through occasional letters and postcards the location of backpackers could never really be known by family and friends back home. Instant, global contact mediated by social media now provides friends and family a better starting point from which to begin when the worst occurs. Chris (respondent 2) provided a first-hand account of the Chilean earthquake and the social media response. "Following the Chilean earthquake last week, we found dozens of comments on Facebook and Twitter asking where we were, if we were OK, and if we needed help. People were asking the Twitter community if anyone had heard from us and by the time I answer emails and logged into Twitter to say we were safe, people I had emailed had posted on twitter and Facebook that we were safe. It was nearly instantaneous and this wide group of virtual friends we've never met were rallying around to look for us."

While this instant contact with home, friends and family, and the virtual community can be very beneficial when travelling, and during disasters, a theme emerged from several of the interviews that suggests that being too connected can take a way from the experience of travelling. Brandon (respondent 6) urges that "being TOO connected to home dissipates your focus….and really takes away from the place you are exploring. How can you really enjoy what's going on in a local village if your mind is thinking about gossip and updates from home that come straight to your phone?"

Gasser and Simun (2010) suggest in their discussion of the travel experiences of 'Digital Natives' that while these individuals are physically traveling, they are mentally and emotionally at home. The innovations in communication technologies make it increasingly easier for individuals to have undisturbed connections with home. As the barriers for maintaining continuous contact with various networks have decreased, individuals now have the *choice* of who they will stay in contact with, how they will stay in contact, and when they will make contact. Some self-described

flashpackers, such as Don (respondent 3), carry lots of technological devices meant to maintain connections with home, virtual community, and to document their experiences to share virtually. Despite this Don (respondent 3) said that he, "Purposefully did not bring a cell phone with me on my trip, as a way to stay somewhat disconnected." This marks an interesting shift in the perception of 'being disconnected' from previous backpackers, who were disconnected by default of their form of travel. Using a net-book or laptop to connect and interact with a virtual network through social media appears not to be defined as 'being connected.' As Alan (respondent 5) explained, "I typically check in (depending on how wired my destination country is) between 3-6 times a week via email, Facebook and Twitter. I do not, however, take a cell phone with me…I enjoy my lack of phone/limited connectivity on the road."

4.4 The Virtual Backpacker Experience

The recent innovations in information and communications technologies have given rise to virtual-cultural hybrid spaces (Paris, 2010). Previously the backpacking culture was immobile, only accessible to a select few while travelling to backpacker enclaves and over backpacker trails. This limited accessibility also facilitated the noticed gap between backpacking ideals and backpacking experience. Backpacking culture could only be experienced in the close physical proximity to other backpackers, thus decreasing the independence and local immersion. Virtual moorings of backpacker culture arguably allows individuals to have intimate contact with the backpacking community from anywhere at any time.

Alan (respondent 5) was aware of the importance of technological innovations for backpacking culture. He said that, "It [technological developments] further empowers backpackers, and encourages backpackers to develop an extensive social network of friends and contracts which offer constant insights into different cultures and peoples." These hybrid spaces resulting from the technological developments allow individuals to experience and interact with the backpacker culture freely without the physical limitations of the past. This arguably allows them the freedom to travel completely physically 'off-the-beaten-track' if they so desire, as they will be safe in knowing that they will be able to share their experiences with the virtual backpacker community.

The virtualization of the backpacking culture provides spaces for individuals to maintain their sense of belonging and connection with like-minded individuals. As Sara (respondent 4) puts it, "Before I discovered Twitter, I really didn't have a community of travellers to connect with while at home. It was very discouraging, but now I have encouragement, I have advice, and I have an outlet for expression." The online backpacker community has developed a social structure that in itself complements the physical backpacker spaces by providing a place of continuity of the culture, instead of sporadic interactions on the road with the 'road' culture. The continuity of the backpacker culture from 'on-the-road' to the virtual to the home, also means a continuity of individuals' identity. Current innovations have led to an increase in the virtual proximities of an individual's networks. Individuals' online identities are losing their anonymity and privacy. Social networking sites, like Facebook and Twitter, allow individuals to interact with multiple networks

simultaneously, but they also decrease an individual's ability to maintain anonymity. Similarly, travellers on the road have seen a decrease in their anonymity, as the ability to maintain connections with other travellers they have met has benefited from the communication developments. In terms of developing authentic relationships with other travellers as well as individuals at the location, this was perceived as a positive development by some of the interviewees. Alan (respondent 5) believed that for him, "the reduction of anonymity in both [travel and internet] is a push towards more real friendships and interactions."

Another interesting effect of the Web 2.0 developments for backpacking and tourism in general is that User Generated Content can influence or at least be perceived to influence more people to travel. Tara (respondent 7) attributes both the development of Web 2.0 and the inclusion of backpacking in mainstream media as a contributing factor, as she concluded that, "more young people are backpacking and exploring due to the ability to explore blogs, travel sites, reality shows, and hear other fellow backpacker stories." Brandon (respondent 6) noted that by posting photos and stories online, people at home are able to dream and escape by virtually taking part in the traveller's adventures. He also notes a downfall to this; that it can have a negative effect on the exploration, expectations, and overall experience. As Brandon (respondent 6) stated, "when you arrive and you now have a filter planted firmly in your head rather than an open mind in which to form you own personal opinions."

The abundance of information and the ease of accessing it through Social Media and mobile devices has arguably resulted in the increased independence of backpackers (Paris, 2010), which is one of the main ideals of the backpacking culture. Alan (respondent 5) summed up the 'pre-Facebook' and 'pre-iPhone' backpacking experience as, "Less connected. Harder to research on the go. Less spontaneous." It is important to note that there were conflicting views on whether social media and communications technologies increased the independence of backpackers in the interviews. Chris (respondent 2) argued that the new developments have just replaced past dependencies, "Where someone used to religiously carry around their Lonely Planet bible, they're now perhaps using a Google search, Wikitravel or other personal favourites to find the same information." In Sara's (respondent 4) response, a similar sentiment emerged suggesting a dependence on information via the Internet might exist. Sara (respondent 4) said, "it is so much easier to plan and book your own travel on the internet. I would say I do 75% of my research for trips on the internet." On the other hand individuals have more control and flexibility in their travel planning. The virtualization of word of mouth recommendations allow individuals to instantly access vast stores of peer reviews on destinations, accommodations, restaurants, etc, while being physically mobile. Alan (respondent 5) is a big proponent of this innovation, "Social functions (such as real time reviews) on booking sites have made a huge difference. One of my favourites is Hostelworld.com which provides a multi-point review system with user comments for each hostel. There is nothing better than being able to look through tens if not hundreds of reviews left by real travellers." Some individuals try to find a middle-ground between planning and allowing time for drifting. As Chris (respondent 2) summed up, "With so much access to information and online booking, many people are planning much more in advance than in the past. We try to find a middle point between understanding our destination, booking flexibly

and a few days in advance, and also finding space and time to go with the flow and take local advice from people online and offline." While the general feeling of respondents is that the technological innovations and increased amount of information available online has made backpacking easier, Jess (respondent 8) pointed out that, "There is also conflicting information that sometimes complicates things, especially when you go off the beaten path. So we would say that it's [recent technological developments] made it [backpacking] both easier and more complex." As information is available instantly, anywhere, and at anytime, independence and pre-planning using consumer-generated media have become a balancing act. The resulting time-compression of the developments in information and communication technologies has also resulted in different constructions of time and the backpacking experience.

5 Conclusion

The nearly constant connectivity of backpackers while travelling creates a situation in which experiences are shared across physical distances instantaneously with multiple networks. This mobile sociality is now blurring the boundaries between home and away are allowing individuals to portray, construct (and reconstruct) and relieve their trips interactively within their mobile sociality as they are experiencing them. This instant mediation needs to receive more attention in future studies, as it has in the reconstruction of experiences through social media after returning home (Xiang & Gretzel, 2009; Pudliner, 2007).

Future studies should focus on re-conceptualizing what it means to be connected and disconnected, as well as explore more deeply the fluidity of the tourist experience as the increasing convergence of information and communication technologies and tourists' experiences allows individuals to remain embedded in their social networks while maintaining a state of physical mobility and localness. The ease of connection and the multiple channels of connectivity mean that the constraints of connecting are being quickly marginalized. The actual experience of travel could be impacted negatively because of the easy and instantaneousness of virtual interactions with home social network, which supports the argument by some that technology can detract from the tourist experience (Uriely, 2005). Several examples were given of situations in which individuals were so tuned into their virtual networks that they were not mindful of their physical location, experiences and social interactions. Additionally, backpacking has often been seen as a time of 'finding oneself' or a rite-of-passage, achieved through a period of detachment, which arguably is threatened by the constant connectivity with friends, family and home.

Developments of social media positively affect the backpacker experience by allowing greater independence through the access to information and the ability to let friends and family at home know they are fine, safe and secure, enhancing their ability to have a more authentic experience by focusing on building relationships with local people via social media, creating a greater ease of return home and reducing reverse culture shock, and documenting a trip through more advanced means. Essentially, being continually connected facilitates positive benefits allowing individuals to embody the shared cultural understanding of backpacking to a greater extent. Continual connection also empowers individuals with the ability to choose which networks they maintain connection with. Additionally, the findings suggest that

individuals might avoid particular technologies in order to create an artificial constraint to connecting to certain social networks. For example, the results indicate that many individuals chose to not maintain connection with work by leaving their cell phone at home, not checking their work email, and not using 'professional' social networking sites like LinkedIn. Greater understanding of these preferences and uses of different types of social media and mobile devices can help the industry to adapt strategies with maximum effectiveness.

Participation and interaction with the online backpacker communities implicate individual backpackers in the collective and collaborative creation of cultural knowledge. Each of the 'online communities' represent hybridized spaces of virtual and corporeal mobilities, social interactions and knowledge generation. It is important to note that a social context is needed to give meaning to the various mobilities, knowledge, and technologies on which the online communities are based (Molz, 2010). The backpacking culture provides the social context for these online backpacker communities. Thus the virtualization of the backpacker culture is a process in which the backpacking culture is manifested through the various information and communication technologies, the knowledge creation process, the social interactions, and the continued hybridization of the physical and virtual spaces through the individual use of mobile technologies.

The virtual moorings of backpacking are based on a fluidity of networks and purposeful collaborations (Molz, 2010) instead of a more hierarchical structure, as conventional with the backpacker road culture. Future studies into the social and cultural interactions in these hybrid spaces and the impact on the 'real-world' travel and tourist experience will become increasingly important in the near future. The convergence of information and communication technologies has created a more decentralized and democratized space where knowledge is collaboratively and collectively created, shared, filtered, contested, and consumed. Virtual spaces and processes can dramatically increase the time it takes for a piece of knowledge to become integrated into the shared culture. In the past, the backpacking road culture has been noted as experiencing very little or no change over many decades as a result of limited time that backpackers would spend in the physical spaces of backpacking, and the lack of access to the 'road culture' from home (Sorensen, 2003).

The rate of change and adaption could be argued to be much greater now because of the virtual nature of the backpacking culture. The virality of backpacker knowledge creates an environment in which knowledge can be experienced, created, shared, contested, reconstructed, and authorized in an extremely short amount of time. This has large implications for the backpacker tourist industry. Previously, just being located in close geographical proximity to a backpacker enclave, or being listed in a Alternative travel guidebook was enough to ensure viability. Nowadays the market place is much more complex because of the proliferation of online (and often unfiltered consumer based) information, competition between business online, and geographically independent technologically connected travellers. Just as the 'backpacker' industry adapted and evolved during the past hundred years to follow the progression of travellers from tramps to drifters to backpackers, the industry needs to adapt to the current forces of change that have resulted in the emergence of the flashpacker.

References

Bauman, Z. (2000). *Globalization: The human consequences.* New York: Columbia University Press.

Bauman, Z. (2007). *Liquid times: Living in an age of uncertainty.* Cambridge: Polity.

Gasser, U. & Simun, M. (2010). Digital lifestyle and online travel: Looking at the case of digital natives. In R. Conrady, and M. Buck (Eds.) *Trends and Issues in Global Tourism 2010*. Berlin: Springer.

Hannam, K. & Diekmann, A. (2010). From backpacking to flashpacking: developments in backpacker tourism research. In K. Hannam, and A. Diekmann (Eds.), *Beyond backpacker tourism: Mobilities and experiences*. Bristol: Channel View Publications.

Hannam, K., Sheller, M., & Urry, J. (2006) Editorial: Mobilities, immobilities and moorings. *Mobilities* 1(1): 1-22.

Jarvis, J. & Peel, V. (2010). Flashpacking in Fiji: Reframing the 'global nomad' in a developing destination. In K. Hannam & A. Diekmann (Eds.), *Beyond backpacker tourism: Mobilities and experiences*. Bristol: Channel View Publications.

Litvin, S., Goldsmith, R. & Pan, B. (2008). Electronic word-of-mouth in hospitality and tourism management. *Tourism Management* 29(3): 458-468.

Makimoto, T. & Manners, D. (1997). *Digital Nomad*. New York: Wiley.

Mascheroni, G. (2007) Global Nomads' Network and Mobile Sociality: Exploring New Media Uses on the Move. *Information, Communication, and Society* 10(4): 527-546.

Molz, G. (2006). Watch us wander: Mobile surveillance and the surveillance of mobility. *Environment and Planning A* 38(2): 377-393.

Molz, G. (2010). Connectivity, collaboration, search: Interactive travel and the mobile production of knowledge. In M. Buscher, K. Witchger, & J. Urry (Eds.), *Mobile Methodologies* (in press). London: Routledge.

Murphy, L. (2001). Exploring Social Interactions of Backpackers. *Annals of Tourism Research 28*(1): 50-67.

O'Regan, M. (2008). Hypermobility in Backpacker lifestyles: The emergence of the internet café. In P. Burns and M. Novelli (Eds.) *Tourism and mobilities: Local-global connections*. Walingford: CABI.

Paris, C. (2010). The virtualization of backpacker culture: Virtual moorings, sustained interactions, and enhanced mobilities. In K. Hannam & A. Diekmann (Eds.), *Beyond backpacker tourism: Mobilities and experience*. Bristol, Channel View Publications.

Paris, C. (2011). Understanding the Statusphere and Blogosphere: An analysis of virtual backpacker spaces. In R. Law, M. Fuchs, and F. Ricci (Eds.), *Information and Communication Technologies in Tourism 2011: Proceedings of the International Conference in Innsbruck, Austria, January 26-28 2011*. Vienna: Springer.

Pudliner, B. (2007). Alternative literature and tourist experience: travel and tourist weblogs. *Journal of Tourism and Cultural Change* 5(1): 46-59.

Sorensen, A. (2003). Backpacker Ethnography. *Annals of Tourism Research 30*(4): 847-867.

Tussyadiah, I. & Fesenmaier, D. (2009). Mediating tourist experiences: Access to places via shared videos. *Annals of Tourism Research* 36(1): 24-40.

Uriely, N. (2005). The tourist experience: Conceptual developments. *Annals of Tourism Research* 32(1): 199-216.

Urry, J. (2002). Mobility and proximity. *Sociology* 36(2): 255-274.

Xiang, Z. & Gretzel, U. (2009). Role of social media in online travel information search. *Tourism Management* 31: 179-188.

A 'Green Economy Tourism System' (GETS): Architecture and Usage

Alexandra Law[a],
G. Michael McGrath[b], and
Terry DeLacy[a]

[a]School of International Business
[b]School of Management and Information Systems
Victoria University, Melbourne, Australia
alexandra.law@live.vu.edu.au
michael.mcgrath@vu.edu.au
terry.delacy@vu.edu.au

Abstract

The world's climate is changing and we are seeing the emergence of a new 'green economy' around the world. Tourism destinations are challenged to adapt to this new reality in order to stay competitive and sustainable. This paper reports on the development of a decision support system (DSS) to assist destinations with this transition. Based on system dynamics theory, GETS is being developed with the aim to allow for scenario modelling around greenhouse gas reductions, enhancing environments and ecosystems; global, regional, local economic and socio-cultural trends; as well as intersecting with changing market demand, supply chain dynamics, destination competitiveness as well as brand positioning and funding options. This paper introduces the system architecture of GETS and provides a usage example to demonstrate how the desired system capabilities may be achieved.

Keywords: green tourism; sustainability; decision support systems (DSS).

1 Introduction

This paper introduces the system architecture and usage examples of a decision support system (DSS) which is being developed to assist tourism destinations into the new green economy. Countries worldwide are attempting to meet ambitious emission reduction targets and adapt to their individual climate change risks. The UK, for instance, has announced plans to reduce carbon dioxide (CO_2) emissions by 60% by 2050 (compared to 1990 levels) and Germany has committed to a 40% reduction by 2020 (Kannan, 2009; Röttgen, 2010). Such ambitious targets require significant changes, not only in a country's energy system (Kannan, 2009), but also in all other greenhouse gas (GHG) producing sectors (such as tourism, which is a major user of fossil fuels, particularly through transport as well as heating, cooling and lighting (Becken & Hay, 2007).

Currently, tourism is estimated to contribute around 5% to the world's total anthropogenic GHG emissions (Scott et al., 2008; World Economic Forum, 2009). However, it is the tourism sector's growth potential that is giving reason for concern. Estimates show that tourism's GHG emissions could grow by 161% by 2035 in a business as usual scenario (Scott, et al., 2008). In the context of global decarbonisation trends, such estimates highlight the need for climate change mitigation and adaptation in the tourism sector and emphazise the need to adopt a

more sustainable path with significant changes in policy and strategy. However, green economy planning in tourism is a complex process, characterized by high levels of uncertainty. For instance, tourism is not included as a sector in traditional emission inventories, and as such, little information is avaiable on sources and the magnitude of the sector's GHG emissions (Becken & Hay, 2007).

Another example of uncertainty for the tourism sector is the emergence of 'green demand', which remains difficult to quantify. While it is anticipated that climate change and environmental perceptions will alter destination choice and consequently influence tourism demand (i.e. Gössling, 2011; Gössling et al, 2008; Scott, et al., 2008; Simpson et al, 2008), there is a lack of information available for destination policymakers and planners to understand the dynamics behind these changes. For targeted mitigation and adaptation strategies, the relationship and interdependencies between the green economy drivers must be understood. However, a planning framework (of this type) for a green economy transition in tourism destinations does not currently exist. In this context, GETS is being developed to allow destination policymakers and planners investigate dynamic "what if" scenarios around their destination and green eocnomy developments.

The paper is organized as follows: project background is presented in the following section and the research approach is discussed in Section 3. The systems architecture is then outlined and this is followed by a DSS usage example. Concluding remarks are presented in Section 6.

2 Background

The development of a green economy DSS for tourism destinations requires a detailed analysis of the dynamic relationships that influence the behaviour of the system. While some information can be found in the academic and industry literature (see e.g. Georgantzas, 2003; Walker et al., 1999), little is available in the way of detailed planning guides, roadmaps and (probably most importantly) data. Consequently, much of this needs to be obtained in collaboration with a destination that is developing and implementing a tourism green economy strategy.

In 2009, the authors were part of a consultancy team to the Egyptian Government charged with developing strategies for a green economy transformation of the destination of Sharm El Sheikh on the Red Sea. The data and findings from this study allowed a systemic view on the drivers of change for a green economy transformation in tourism as well as an insight into the different stakeholder perspectives. Based on these findings a first holistic model for GETS was developed. The Sharm El Sheikh data highlighted four key elements for a successful green economy transformation in tourism:

1. GHG emission reductions
2. Growing the destinations market demand
3. Enhancing the destinations environments and ecosystems and
4. Sustainability of economic and socio-cultural trends.

The model (see Law et al., 2011 for more detail) highlights the complexity of the approach: all elements and factors are characterized by an integrative nature, reflecting that targeted and effective strategies rely on a holistic and systemic view (see Figure 1).

Fig. 1. GETS system requirements

One of the key challenges of developing and implementing GETS is the source, format and availability of data. The aim of GETS is to produce a system that can work for any destination, for example an island destination such as Bali, a whole country such as Australia, a city or a tourism region. This paper presents a system architecture (and usage examples) designed to demonstrate how this aim may be achieved.

3 Research Approach

As a relatively new field, information systems (IS) research borrows heavily from older disciplines; in particular, engineering and the design sciences. As Simon (1981) has noted, "design sciences do not tell us how things are done but how they ought to be to attain some ends". Much the same applies to IS development and Gregor (2002, p. 12) has posed the question: what constitutes a contribution to knowledge when research is of this type (oftentimes with no hypotheses, no experimental design no data analysis per se)? Hasan (2003, p. 4) responds to this by claiming that IS development, in many cases, should be considered a legitimate research activity (and method) because, not only is knowledge created about the development process itself, but also because "a deeper understanding emerges about the organizational problem

that the system is designed to solve". Markus et al. (2002) put forward a similar case in arguing that IS development is a particular instance of an *emergent knowledge process (EKP)* and that this constitutes original research where requirements elicitation, design and implementation are original and generate new knowledge on *how to proactively manage data and information in complex situations.* Hasan (2003: 6) contends that this often involves a staged approach, where "systems evolve through a series of prototypes" with results of each stage informing requirements for the next and subsequent iterations.

Nunamaker et al. (1991) take an approach consistent with the above but draw on an alternative research tradition in case studies and, in particular, action research. Again, using 'replication' strategies, each new instance (case or action research activity) builds upon and refines knowledge gleaned from previous studies (Yin, 1994). Nunamaker et al. (op cit.), however, nominate two features of IS development that distinguish it from more general action research: first, the techniques of IS development, the properties of the system itself and the situation where the system is to be deployed may all generate important knowledge; and, second, IS research projects are both constrained by the limits that current IT place on the development of systems and are enabled by the uniqueness of the technology (which can, as a tool, mediate knowledge generation and the communication of same).

The latter feature has been studied extensively by scholars in 'activity theory' (Vygotsky, 1978). Notably, activity theorists emphasize the holistic nature of the IS development process and, in particular, the critical nature of the cultural and social context within which systems are developed (see, for example, Engestrom, 1987; Nardi, 1996). The socio-technical view of IS, where hardware, software, people and processes are integrated into a complex, purposeful whole, is one of the key features that make information and communication technologies "like no other in the history of mankind" (Hasan, 2003: 4). Thus, to summarize: the development of our GETS IS is a legitimate research activity in its own right, which draws on the more established, traditional research approaches of the design sciences and especially case study/action research. Each new application of GETS (e.g. to a new destination) produces a new version of our prototype and extends our knowledge of the green tourism economy research domain. This is akin to employing a multi-case (study) research strategy - with each new case refining and extending results of previous iterations - and finally, many research findings and outputs are actually inherent in the various conceptual models (and implementations of these) that constitute the GETS IS.

4 System Architecture

A high-level view of the GETS architecture is illustrated in Figure 2. A fundamental objective of the GETS project is to produce a system that is iterative, scalable and open. *Iterative* (in this instance) means that each application (e.g. to a new destination or aspect of a destination) produces a new prototype that increases or refines our knowledge of the green economy domain; *scalable* means that the system must be able to cope equally effectively with large and small destinations; and *open* means that GETS must be capable of handling any type of data, irrespective of source or format.

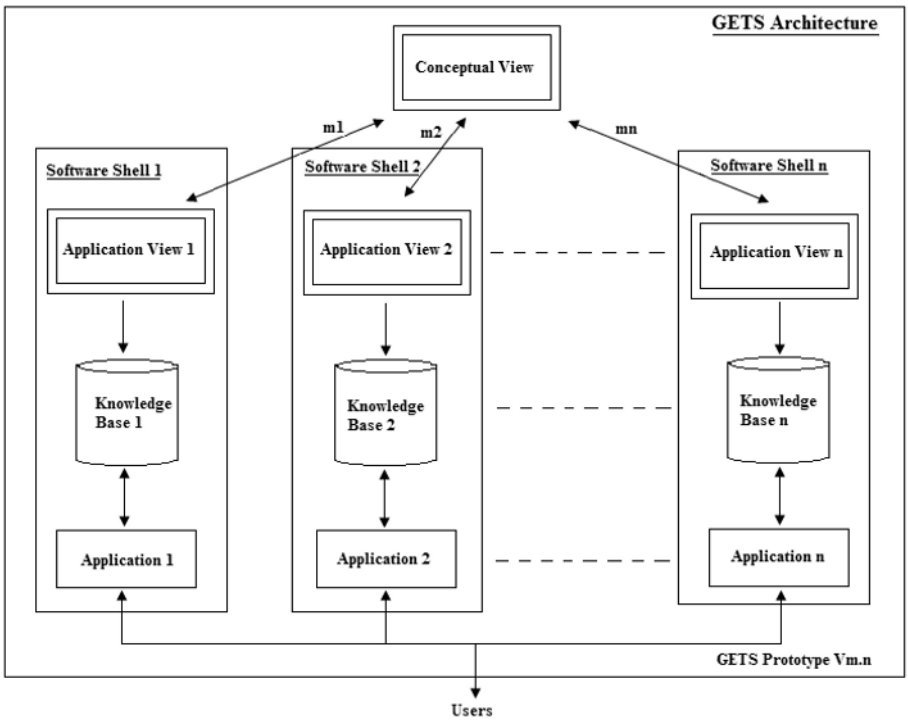


Fig. 2. GETS architecture – high-level view

One of the keys to realizing both an iterative and a scalable system is developing all code (and higher-order applications) around abstracted data models. Essentially, the aim is to allow new functionality to be added (e.g. as issues associated with a new destination introduce new system requirements) without having to revise existing applications. Similarly, development effort may be substantially reduced if large and small destinations (and sub-destinations) can be accommodated by the same applications without modification. Again, abstraction assists with this. We now illustrate how this is accomplished within GETS through a fairly simple (but realistic) example.

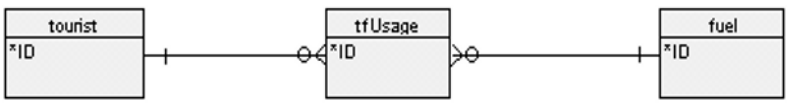
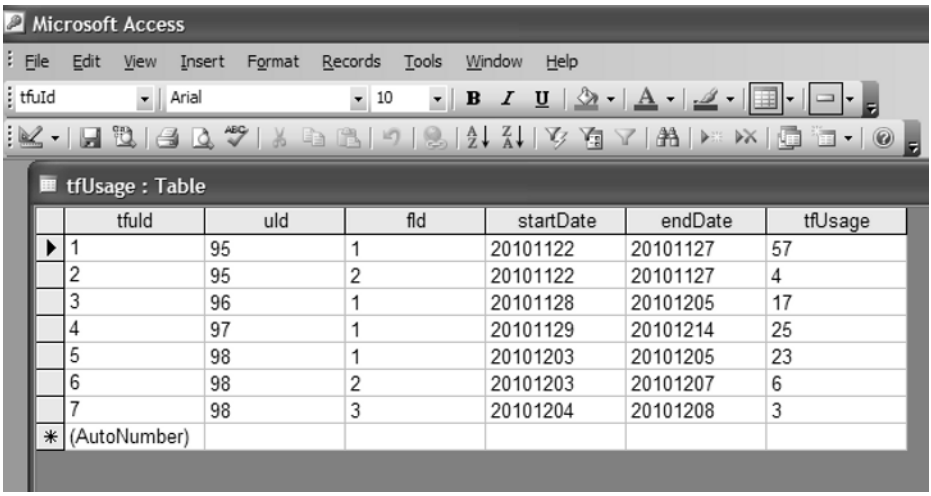


Fig. 3. E-R representation of tourist fuel usage

One indicator we are definitely interested in is tourist fuel usage, and a data model our application might employ is illustrated in entity-relationship (E-R) form (Chen, 1976)

in Figure 3. Implemented as a relational database application, the intersecting entity, *tfUsage,* would translate to something like the *Access* table presented in Figure 4 (which has been populated with hypothetical test data). Pseudo-code constructed to calculate tourist gasoline usage (assuming the *fId* for gasoline is 1) might then read something like the following:

```
begin
    set totalTGUsage to 0.
    for each tfuId
      if fId = 1
        then add tfUsage to totalTGUsage
        else no action.
    display totalTGUsage.
end.
```



tfuId	uId	fId	startDate	endDate	tfUsage
1	95	1	20101122	20101127	57
2	95	2	20101122	20101127	4
3	96	1	20101128	20101205	17
4	97	1	20101129	20101214	25
5	98	1	20101203	20101205	23
6	98	2	20101203	20101207	6
7	98	3	20101204	20101208	3
(AutoNumber)					

Fig. 4. Fuel usage database table (hypothetical)

Assume now that, sometime after implementing the above, the need arises to track energy usage for all classes of visitors, plus locals and local businesses. Moreover, usage needs to be aggregated at many different levels and for many different categories (e.g. hotels, types of hotels, business and leisure visitors etc.). This will certainly demand extensions to our E-R model, plus amendments to our initial relational tables and the code constructed around these: in short, before even tackling the enhanced requirements, substantial (and costly) maintenance must be undertaken in order to ensure existing functionality continues to work correctly.

Abstraction can alleviate these difficulties. One approach would be to replace the E-R schema presented in Figure 3 with that illustrated in Figure 5. In our revised schema, tourists, visitors, businesses etc. (plus their sub and super-types) all become *parties* of a particular *party type* and, where appropriate, these are all related to each other as *party-party involvement (ppi)* instances. *Resources* are treated in much the same way

and *party-resource involvements (pris)* are used to link parties to their resource usage (at whatever level of granularity is required).

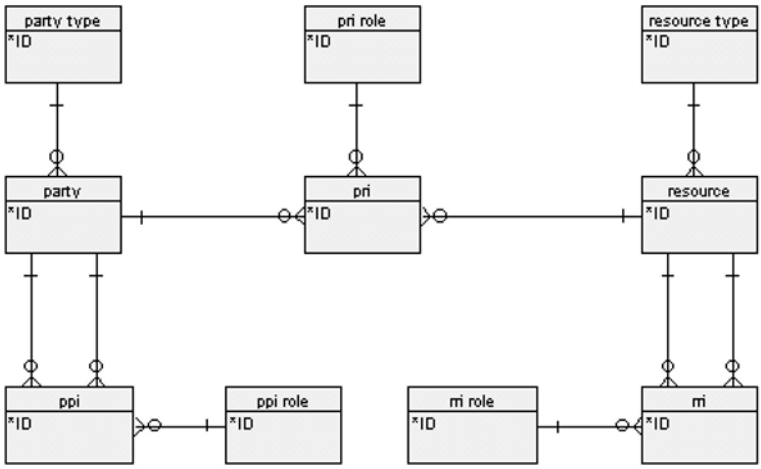


Fig. 5. Abstracted view of the original E-R model (see Fig. 3)

Two major benefits of this abstracted approach are: i) the same code can be used for multiple relationships (thus substantially reducing development costs and time); and ii) for the most part, application extensions do not require existing code and schema modifications – because new requirements can often be accommodated with simple revisions at the table level.

The history of software engineering has demonstrated that, in all but trivial systems, accurately capturing *all* requirements up-front is close to impossible (Somerville and Sawyer, 1997). In addition, there are other important benefits that result from abstraction and, for a detailed discussion of these, the reader is referred to the seminal work of Feldman and Miller (1986).

The open systems objective is realized by adopting a design for GETS consistent with ISO '3-Schema Architecture' principles (Griethuysen, 1982). Referring back again to Figure 2, the *Conceptual View* is a highly abstracted model of the total system, completely free of any implementation-level detail (hence, use of the term, 'conceptual'). *Application View 1, ----, Application View n* are information schemas developed for individual applications, implemented within specific software shells *(Software Shell 1, ----, Software Shell n).* Examples of these (used in applications implemented to date) are *Excel™, Access™,* a rule-based expert systems shell called *Flex™* and the system dynamics simulator, *PowerSim™.* There is a third ISO 3-Schema level, the *Internal View,* which deals primarily with technical aspects of each application (relating to efficiency etc.) and is beyond the scope of this paper. Basically, in recognition of the fact that no one modelling or development method is best for all applications, we have adopted a 'horses for courses' approach. At the same time, no one application is truly stand-alone and, consequently, data must be

shared between applications. Within GETS, this is accomplished through the *mappings, m1, ----, m_n,* which allow data within each application (irrespective of its form) to be mapped to and from the common, uniting conceptual view. This approach to application integration is far more efficient than the alternative of developing 1:1 interfaces between applications as required (McGrath, 1997).

5 System Usage: An Example

The following example is taken from a recent study into possibilities for introducing green tourism into the well-established coastal and water sports destination of Sharm el Sheikh in Egypt. This particular application is concerned with hotel energy usage and visitor goodwill, and was constructed using system dynamics (SD) models and principles. It was implemented using the software product, *PowerSim™*. SD has its origins in the work of Forrester (1961) and, more recently, has enjoyed something of a resurgence – largely due to Peter Senge's (1990) very influential work on 'the learning organization' and the development and release of easy-to-use, powerful, SD-based software modelling and simulation tools (such as *iThink™, Vensim™* and *Powersim™).* Recent examples of where SD has been used to good effect in tourism include the 'Tourism Futures Simulator' (Walker et al., 1999), the hotel value chain modelling work of Georgantzas (2003), and the tourism multipliers model of Loutif, Moscardini and Lawler (2000).

In this instance, the destination had decided to go green as part of its rejuvenation strategy and, consequently, was interested in possible impacts. Hotels[1] within this particular destination were divided into three categories: i) *non-green* – where almost no energy reduction initiatives had been implemented; ii) *moderate-green* - where most of the relatively easy (and cheap) initiatives had been implemented; and iii) *total-green* – where a significant number of capital-intensive initiatives had been implemented. The user is required to specify desired transformation rates (for example, 80% of hotels will be moderate-green within 5 years and 50% will be total-green within 8 years).

The user may test some social impacts of the energy reduction strategy; specifically, the impacts of visitor and local goodwill. Visitor goodwill, which manifests itself as the payment of premium prices for the accommodation, may be generated because there is some evidence that tourists (western tourists in particular) are tending to favour destinations that appear to be serious about environmental improvement (De Lacy and Lipman, 2010), while local goodwill, which manifests itself as community satisfaction may occur as a result of both environmental improvement and improved infrastructure and facilities that might result from more, higher-yielding visitors.

Baseline data gathered during the Sharm el Sheikh green tourism planning study was then used to instantiate SD models developed for scenario testing and to extend the basic SD model to include energy reduction impacts on visitor tariffs, revenue and yield. Referring back to Figure 2, this data represents an application view (derived

[1] It was established that hotels within the destination were responsible for over 70% of tourism-related energy usage and CO2 emissions, exclusive of energy used in travel to and from the destination – hence the accommodation focus.

from the conceptual view and associated knowledge bases) for evaluating various scenarios related to energy usage in the accommodation sector. Implemented in *Excel™,* this is a GETS application in its own right and deals with much of the basic data analysis that must be conducted in any green tourism planning study. Here though, we are primarily concerned with the SD energy reduction scenario generation exercise introduced above. This has its own application view but much of the data it uses is sourced from the *Excel™* application.

As previously noted, hotels move from a status of non-green (NG) to moderate-green (MG) and, finally, to total-green (TG) and the user may set and test target transfer rates. The SD 'stock-flow' model employed here is presented in Figure 6 and, without going into detail on how these models are used to underpin SD simulation (for a comprehensive treatment see, for example, Maani and Cavana (2000)), the flow of destination energy through the three phases is fairly apparent from the model.

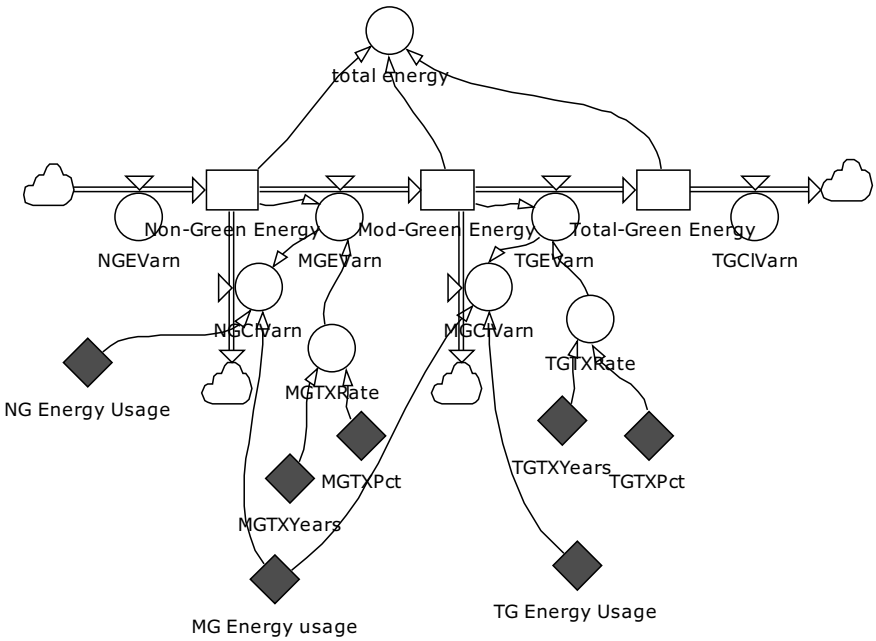


Fig. 6. Hotel energy usage – SD core process model

Each of the 'stocks', 'flows' and 'converters' (the latter represented by circles) represent simulation variables and the diamonds represent constants. All of these must be initialized prior to simulation commencing and data from Table 1 is used here. For example, with an initial NG-MG-TG split of 70-30-0, the initial values for the *Non-Green Energy, Mod-Green Energy* and *Total-Green Energy* stocks are 2,047.9 million, 761.8 million and 0 MJ/year respectively. Some key simulation constants are initialized from slider values set by the user and sample simulation run output is presented in Figure 7 below. In this instance it can be seen that the Sharm

destination's total annual accommodation sector energy usage is predicted to drop from 2.9 million to around 1.4 million MJ/year over a 10-year period.

Many other possible scenarios may be evaluated using the SD models developed as part of the Sharm strategic planning process. A key outcome of this exercise was the development of a marketing and branding strategy for the destination based around the 'Green Sharm' concept. The belief was that green tourism would attract fewer, higher-yielding visitors and that this would help to alleviate many of the serious environmental and social problems confronting the destination. While this may sound simple enough (in concept, if not in implementation), it is actually underpinned by many assumptions and a great deal of complexity. With the example presented, we have demonstrated how an interactive, sophisticated model such as GETS can be operationalized by using advanced decision support techniques such as SD. In this situation, SD can be employed: i) to expose and clarify much of this underlying complexity; ii) to model key aspects of the green tourism domain in a formal, precise and rigorous manner; iii) to use the models to generate and evaluate possible future scenarios; and iv) to, hopefully, prevent the unintended consequences of policy decisions that frequently bedevil complex decision-making settings of this type (Vennix, 1996).

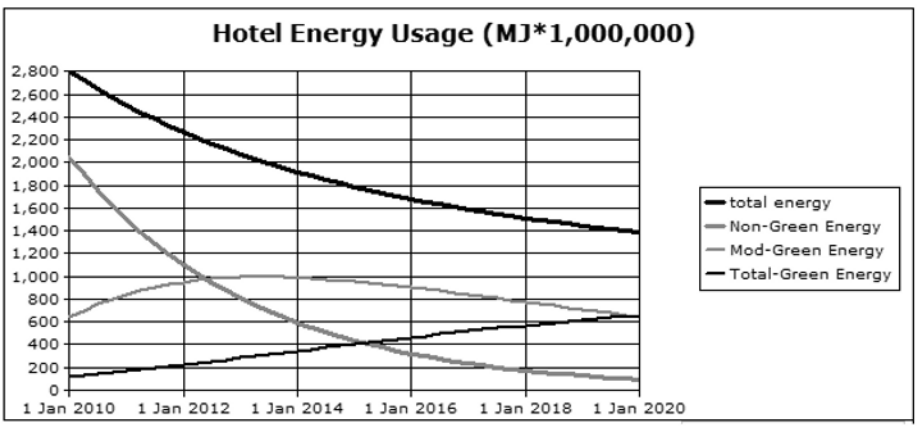


Fig. 7. Hotel energy usage – sample SD simulation output

6 Conclusion

This work adds to an increasing body of applied research concerned with green economy strategic planning. In this particular paper, we outlined a DSS implementation designed to support this activity and were primarily concerned with the architecture of this system. In particular, the paper focused on data abstraction and the benefits that flow from that – specifically: i) a reduction in DSS development and maintenance costs; ii) a reduction in control flow complexity through data-driven programming techniques; iii) an improved capability to cope with business environment volatility; and iv) the ability to more effectively integrate application schemas.

One area we have not covered in this paper is the issue of how our DSS might be employed to advise on green-based destination rejuvenation strategies. A prototype component of this type, based on Butler's (1980) 'Tourism Area Life Cycle' (TALC) has been developed (Pornphol & McGrath, 2011) and is due to be field-tested in Phuket, Thailand in late-2011. In this application, the knowledge base contains instances of TALC cases taken from the literature, with particular emphasis on instances of destinations in stagnation/decline where rejuvenation strategies (often green economy-based) have been applied. Users enter details (e.g. information on the current status of the destination environment, tourism enterprises, types of current visitors and significant problems as search parameters) and a case-based reasoning component retrieves those knowledge base cases that most closely match the new case as a 'Best Match Report'. The user may then review these cases to ascertain whether they provide pointers to actions, strategies or policies that might assist the destination with its current (or potential) problems.

An example might be the Waikiki environmental quality enhancement initiatives adopted by the City and County of Honolulu during the 1990s (Patoskie, 1992). If adoption of a particular initiative is under serious consideration, the user may then perform some "what if" analysis using the type of SD module outlined in this paper. It is our view that this type of functionality would constitute a very useful enhancement to the current GETS system.

References

Becken, S. & Hay, J. E. (2007). *Tourism and climate change: Risks and opportunities*. Clevedon, Buffalo: Channel View Publications.

Butler, R.W. (1980). The Concept of a Tourism Area Cycle of Evolution: Implications for Management Resources. *The Canadian Geographer* 24(1): 5-16.

Chen, P.P.S. (1976). The Entity-Relationship Model- Towards a Unified View of Data. *ACM Transactions on Database Systems* 1(1): 9-36.

De Lacy, T. & Lipman, G. (2010). GreenEarth.travel: Moving to Carbon Clean Destinations. In C. Schott (Ed.), *Tourism and the Implications of Climate Change: Issues and Aactions.* Bingley, UK: Emerald Group Publishing.

Engestrom, Y. (1987). *Learning by Expanding: An Activity-Theoretical Approach to Developmental Research.* Helsinki, Finland: Orienta-Konsultit.

Feldman, P. & Miller, D. (1986). Entity Model Clustering: Structuring a Data Model by Abstraction, *The Computer Journal* 29(4): 348-360.

Forrester, J.W. (1961). *Industrial Dynamics.* Cambridge, MA: MIT Press.

Georgantzas, N.C. (2003). Tourism Dynamics: Cyprus' Hotel Value Chain and Profitability. *System Dynamics Review* 19(3): 175-212.

Gössling, S. (2011). *Carbon management in tourism: Mitigating the impacts on climate change*. Oxon (UK) and New York (USA): Routledge.

Gössling, S., Peters, P. & Scott, D. (2008). Consequences of climate policy for international tourist arrivals in developing countries. *Third World Quarterly* 29(5): 873-901.

Gregor, S. (2002). A Theory of Theories in Information Systems. In S. Gregor and D. Hart (Eds.), *Information Systems Foundations: Building the Theoretical Basis.* Canberra: ANU.

Griethuysen, J.J. (1982). *Concepts and Terminology for the Conceptual Schema and the Information Base.* ISO Technical Report ISO/TC97/SC5/WG3.

Hasan, H. (2003). Information Systems Development as a Research Method. *Australasian Journal of Information Systems* 11(1): 4-13.

Kannan, R. (2009). Uncertainties in key low carbon power generation technologies: Implication for UK decarbonisation targets. *Applied Energy* 86(10): 1873-1886.

Law, A., DeLacy, T., McGrath, G. M. & Whitelaw, P. A. (2011). Tourism destinations in the emerging green economy: Towards blending in brilliantly. In *Proceedings of CAUTHE 2011 National Conference*, 8-11 February, Adelaide.

Loutif, M., Moscardini, A.O. & Lawler, K. (2000). Using System Dynamics to Analyse the Economic Impact of Tourism Multipliers. In P.I. Davidsen, D.N. Ford and A.N. Mashayeekhi (Eds.), *Proceedings of the 18th International Conference of the System Dynamics Society* (pp. 132-232). Albany, NY: System Dynamics Society.

Maani, K.E. & Cavana, R.Y. (2000). *Systems Thinking and Modelling: Understanding Change and Complexity.* Auckland, New Zealand: Prentice-Hall.

Markus, L., Majchrzak & Gasser, L. (2002). A Design Theory for Systems that Support Emergent Knowledge Processes. *MIS Quarterly* 26(3): 179-212.

McGrath, G.M. (1997). Behavioural Issues in Information Systems Design, Development and Implementation: A Process Modelling Framework. In *Proceedings of the 3rd Pacific-Asia Conference on Information Systems,* Brisbane, 1-5 April.

Nardi, B.A. (Ed.) (1996). *Context and Consciousness: Activity Theory and Human-Computer Interaction.*Cambridge, MA: MIT Press.

Nunamaker, J.F., Chen, M. & Purdin, T. (1991). Systems Development in Information Systems Research. *Journal of Management Information Systems* 7(3): 89-106.

Patoskie, J. (1992). Traditional Tourist Resorts: Problems and Solutions – Honolulu (Hawaii, USA). In *International Forum on Tourism to the Year 2000: Prospects and Challenges,* Acapulco, Mexico.

Pornphol, P. & McGrath, G.M. (2011). Information Systems as an Emergent Knowledge Research Process. To appear in *The International Journal of Interdisciplinary Social Sciences* 6.

Röttgen, N. (2010). *Climate protection as an opportunity: Plenary statement at the UN Climate Change Conference COP16, Cancun, 08/12/2010*. Retrieved from http://unfccc.int/files/meetings/cop_16/statements/application/pdf/101208_cop16_hls_germany.pdf

Scott, D., Amelung, B., Becken, S., Ceron, J. P., Dubois, G., Gössling, S., et al. (2008). *Climate change and tourism: Responding to global challenges*. Madrid, Paris: World Tourism Organization (UNWTO), United Nations Environment Programme (UNEP).

Senge, P.M. (1990). *The Fifth Discipline: The Art and Practice of the Learning Organization,* Milsons Point, NSW: Random House.

Simon, H. (1981). *The Sciences of the Artificial.* Cambridge, MA: MIT Press.

Simpson, M. C., Gössling, S., Scott, S., Hall, C. M. & Gladin, E. (2008). *Climate change adaptation and mitigation in the tourism sector: Frameworks, tools and practices.* Paris, France: UNEP, University of Oxford, UNWTO, WMO.

Sommerville, I. & Sawyer, P. (1997). *Requirements Engineering: A Good Practice Guide.* Chichester, UK: Wiley.

Vennix, J.A.M. (1996). *Group Model Building: Facilitating Team Learning Using System Dynamics.* Chichester, UK: Wiley.

Vygotsky, L.S. (1978). *Mind and Society.* Cambridge, MA: Harvard University Press.

Walker, P.A., Greiner, R., McDonald, D. & Lyne, V. (1999). The Tourism Futures Simulator: A Systems Thinking Approach. *Environmental Modelling and Software* 14: 59-67.

World Economic Forum (2009). *Towards a low carbon travel and tourism sector*. Geneva, Switzerland: World Economic Forum.

Yin, R.K. (1994). *Case Study Research: Design and Methods.* Thousand Oaks, CA: Sage.

eTcoMM – eTourism Communication Maturity Model. A Framework to evaluate the Maturity of a DMO when it comes to the Online Communication Management: The Case of Canton Ticino and Lombardy

Elena Marchiori,
Giulia Pavese, and
Lorenzo Cantoni

webatelier.net - Faculty of Communication Sciences
Università della Svizzera italiana (University of Lugano, Switzerland)
{elena.marchiori; giulia.pavese; lorenzo.cantoni}@usi.ch

Abstract

The online environment represents an important communication resource of a Destination Management Organization (DMO). Evaluating the level of preparation of a DMO on its processes dedicated to the online communication management is a challenge for organizational performances, especially if compared with the efforts spent by other tourism players on the web. This study presents the development of a maturity model, named eTcoMM, dedicated to assess the online communication management preparation of a DMO on five online related operational areas corresponding to: staff training, search engines management, web promotion through back-links, website usages analysis, and electronic Word-of-Mouth analysis. Results of its first application with seven Italian and Swiss local DMOs are presented and discussed.

Keywords: maturity model, online communication management, DMOs

1 Introduction

The understanding of the level of maturity reached by the Destination Management Organization (DMO) in charge to manage the online communication of a tourism destination is a challenge for the industry which is still under research. This study investigates the levels of maturity a DMO can reach in regard to its management of processes related to the online communication. In order to investigate those different levels of maturity, a model has been defined. Maturity is considered as a scale formed by evolutionary levels where to place an organization, according to its preparation when it comes to the processes implemented within a specific operational area. The model developed and presented in this study is called eTcoMM, which stands for eTourism Communication Maturity Model, and includes two main elements that constitute the essence of this research: the online communication in the online Tourism domain, which represents the study object, and a maturity model, which is the tool needed for a systematic investigation of the online communication processes/activities within a DMO.

The perspective chosen to analyze the maturity of a DMO regarding the online communication management concerns three aspects: (i) training activities of the people in charge to manage the online communication within the tourism organization; (ii) activities related to analyse the publics receiving this communication; and (iii) activities related to the communicative context: the online

environment. Five areas have been chosen to assess the maturity for five specific areas corresponding to: staff training, search engines management, web promotion through back-links, website usages analysis, and Electronic Word-of-Mouth (eWOM) analysis. Once the maturity model has been built, a validation of the model has been performed with a sample composed by seven DMOs from Canton Ticino, Switzerland, and Lombardy, Italy. Results of this first step are presented in this study. This research has been developed to achieve three goals: (i) develop a maturity model of the web related communication activities of DMOs; (ii) to test the model using a sample of local tourism DMO; (iii) to provide a tool that allows DMOs to assess the maturity of the processes they perform while managing their online communication.

2 Related Works

2.1 Maturity Models

A maturity model is a schematic framework which describes the development-growth of an organization, along specific competencies/behaviors within a time period. This evolution schema is structured in pre-defined stages of growth. The concept of maturity finds its roots in the stage theories, where the main assertion holds that there are organizational elements in a system, which move toward different levels of growth in a specific time period (Nolan, 1973).

The following maturity models have been considered: (i) Quality Management Maturity Grid (QMMG) (Crosby, 1979) was considered relevant, as this model was created in order to enable companies to understand how mature were the processes performed in relation to the management of products' and services' quality (Crosby, 1979); (ii) Capability Maturity Model (CMM) (Paulk, 1993), as it considers the quality of an organization as a unit of a maturity analysis and was used by another important model: (iii) the international standard ISO/IEC 15504. The ISO/IEC 15504 (Paulk, 1993, 1998) which harmonized the existing models adding a new two-dimensional structure for a model composition, with the objective to define a standard model for the process assessment analysis (Garcia, 1997). Among the main applications of maturity models in the domain of Information and Communication Technology (ICT) there are: eLearning, eCommerce, eGovernment, and eEnterprise. These models can be used to perform a prescriptive analysis for future development of Information Systems and Information Technology within an organization. Studies of maturity within the electronic environment related to the tourism domain see similarities with the eCommerce and eGovernment areas, both for the functional and institutional aspects similar to the tourism industry. Maturity models have been applied to the electronic Tourism domain, in particular with the Doolin, Burgess and Cooper Model built in 2000. The model was created to measure the level of development of the tourism websites of regional tourism organizations in New Zealand. Their research has made an assumption that Internet is profoundly changing business models related to tourism. The emphasis was on how to change the communication between an organization and its customers from a mere provision of information function to an active interaction with the customers via the web (Doolin, Burgess, & Cooper, 2002). However, within this model, the research interest was not on the analysis of the management processes (which is the very essence of maturity models), but on the content and site functionalities, confirming the applicability of the

staged approach to the analysis of a tourism website. The model, as well as allowing to investigate the levels of functionalities offered by a website, provided also a useful tool to assess a websites' level of development among its competitors (Doolin, Burgess & Cooper, 2002). Related work was also done by Gretzel and Fesenmaier (2004) in the field of knowledge management in the Tourism domain. Authors used a systems perspective to build the Illinois Tourism Network (ITN), an interorganizational knowledge-based tourism information system which integrated the management of information and knowledge flows in different stages of effective technology use in a tourism organization (Gretzel & Fesenmaier, 2004). Three stages of technology adoption and use by tourism organizations have been recognized: stage 1, low use of IT and knowledge management; stage 2, moderate use of IT and knowledge management; stage 3, high use of IT and knowledge management. On stage 3 functions were highly interactive and more likely to serve strategic rather than operational management needs.

However, since these models did not focus specifically on the maturity of the processes, it was decided to adapt one that allows an assessment of the processes in IT. The model chosen was the Usability Maturity Model (Earthy, 1998), which is strictly related to the ISO / IEC 15504 (Paulk, 1993, 1998). In the Usability Maturity Model a scale of maturity has been set based on human-centredness principles. As in the ISO / IEC 15504, the model is structured in levels, which contain attributes, which need to be satisfied in order to define a certain level of maturity of an organization. Moreover, each attribute also considers specific management practices, which correspond to the activities to be carried out in order to claim to possess the attribute itself. Thus, an organization in order to reach a level and claim to possess an attribute at that level, must accomplish management practices related to that. These practices are evaluated on a scale that includes the following degrees: Not Achieved (N), Partially Achieved (P), Largely Achieved (L) and Fully achieved (F). The organization is then located at the level of maturity based on the management practices accomplished.

2.2 eTourism and Destination Management Organization

As defined by Buhalis (2003), eTourism is a generic term that encompasses all the applications related to the information and communication technology in the field of tourism. From a specific point of view, eTourism is also considered the eCommerce applied to ICT to maximize the efficiency and effectiveness of a tourist organization. Tourism offers experiences that cannot be estimated in advance (Chung & Buhalis, 2008), and are not physically observable at the time of purchase. For these reasons, tourism products and services require representations and descriptions. Information and Communication Technology offers the possibility to enhance the communication and information dissemination in the tourism domain (Buhalis, 2003; Werthner & Steinbauer, 2007). With the new possibilities offered by technologies, the way the business is managed has changed. A growing number of potential travellers began to seek for tourism information on the Internet and expect to obtain benefits from the access to those contents (Gretzel, Hyan, & Yoo, 2008). Moreover, the so-called Web 2.0 (O'Reilly, 2005) and the related social media, changed the ways in which individuals create, exchange and use information (O'Connor, 2008), allowing the possibility for users to easily share their experiences online. In this way an

information is not transmitted only in one direction, but also by many to many: all the communications are available online to all users (Inversini & Buhalis, 2009). Conversations that take place within social media websites about a tourist destination, might participate on the construction of the place brand of the destination (Govers & Go, 2009). At this level, organizations have no longer control over what is said online about the destination, but they have the opportunity to monitor and participate in the conversations themselves. Given the importance of these user generated contents, changes are needed in the management of the promotion of a destination. DMOs should try to recognize the changes taking place in the web communication, and develop customized services to meet the individual needs of tourists. Moreover, constant innovation of hardware and software and the development of networks, require a new view of a tourism organization which needs to be dynamic and competitive in the tourism market (Buhalis & Law, 2008).

3 Research Design

In order to built the eTcoMM model, two models have been used as a conceptual basis: the Usability Maturity Model (Earthy, 1998) has been used for the processes management description, and the Online Communication Model (Cantoni & Tardini, 2006) for the identification of the online related operational areas. The Usability Maturity Model has been used since it describes which changes are needed within an organization in order to improve its internal processes and increase the level of awareness regarding the central role of the user (Earthy, 1998). The concepts presented in the Usability Maturity Model that have been used for the eTcoMM in order to define practices for the online communication processes management were:

i. *Practice*, it is a technical or management activity, which contributes to achieving the results obtained by the processes;

ii. *Management practice,* is the management activity leading to the implementation or institutionalization of a specific feature of the process;

iii. *Process improvement*, is the action taken by an organization in order to change a process and better meet business objectives (Earthy, 1998).

Regarding the processes specifically designed for managing online communication, it has been decided to focus on the online promotion operational areas, corresponding to: promotion via search engine and back-links; usages analysis; eWOM analysis, and training of people in charge to manage the online activities, as they represent the main activities which might drive a powerful online performance of a company (Cantoni & Ceriani, 2007). The Website Communication Model (WCM), later on renamed as Online Communication Model (OCM) (Cantoni & Tardini, 2006; 2010) was used as it allows to look at a website, and/or in general at the online communication, as a business communication act (Cantoni & Ceriani, 2007). The model provides a comprehensive structural description of both operational areas and processes that compose the online communication dynamics within an organization (with a particular focus on the online promotion). The OCM consists of four pillars and a fifth outer element. The first two pillars are related to the contents and functionalities, as

well as to the tools necessary to ensure that contents and services can be communicated/accessed. The third and fourth pillars are related to the people who manage the communication and the users. The fifth element represents the information market, were the participatory contexts within which an online communication is carried on.

3.1 The eTCoMM Model

Based on the OCM (Cantoni & Tardini, 2006; 2010), third and fourth pillars, plus the fifth element were used for defining the five operational areas, presented below, of the eTcoMM model, namely the contexts within which the online communication of a tourist destination takes place: people who manage the online communication (third pillar), analysis of the users (fourth pillar), and the online info market (fifth element). Per each operational area a study of the change in different evolutionary phases was done. The following five areas of expertise evolve based on the processes accomplished by the organization, bringing the organization itself at different levels of maturity. Each operational area can be described as follows:

1. *Staff training:* represents the area of competence regarding the third pillar of the OCM, and it concerns the management of activities related to train and update staff members on eTourism, social media, and trend in IT and Tourism. It is so important due to the continuous changes occurring within eTourism.

2. *Promotion activities using search engines*: represent the study of the organization's presence on search engines, and the activities related to the use of Search Engine Optimization and Search Engine Marketing in order to have the DMO's link on the top positions of search engines for relevant queries. This area of expertise is related to the fifth pillar of the OCM. The importance of the presence on the search engines results is related to the fact that search engines are the most efficient way to procure the first visit to the DMO site (Cantoni & Tardini, 2010; Fesenmaier et al., 2010).

3. *Promotion activities using back-links:* represent a website's promotion strategy that consists in the placement of hyper-links to the website on other websites. They are also considered by search engines as an element that contributes to evaluate the relevance of a website, determining where to place the website in the search engine results (link popularity). Several strategies can be adopted to build back-links: get a back-link for free; through payment (in various forms, such as "pay per click", "pay per lead" or "pay-per-transaction"), or via link-exchange. This area of expertise is related to the fifth pillar of the OCM.

4. *Usages Analysis:* Represents the analysis of the actual users of a websites through the log files analysis or similar strategies. This activity allows the understanding of the audience of the websites: who are the users? Where do they come from? When? Through which links? Having searched which kind of keywords on search engines? Time of visiting? How often? After a promotional campaign or another event? etc. (Cantoni & Ceriani, 2007). This area of expertise is related to the fourth pillar of the OCM.

5. *eWord of Mouth analysis*: Represents the listening and analysis of the online word of mouth about the destination. The related activity is the monitoring of the

online conversations that take place mainly on social media. This listening activity is designed to obtain knowledge and know-how in order to manage a DMO's reputation and online presence. This area of expertise is related to the fourth and the fifth pillars of the OCM.

The maturity degree that tourism organizations can achieve in the implementation of processes related to the above presented operational areas is evaluated based on five levels of maturity. The levels of eTcoMM are similar to those of the Usability Maturity Model, with the exception of the sixth level (institutionalized) which was incorporated in the fifth level. Each level requires an attribute of maturity in order to be evaluated. Attributes, such as in ISO / IEC 15504, are measurement indicators whose existence confirms that certain activities are in place within a process (El Emam & Jung, 2001). The attributes are measured based on the level of possession of each attribute. Levels of possession are four: (N) Not Held: there is no evidence or few evidence in the possession of an attribute; (P) Partially Owned: there is evidence on the achievement of an attribute possession, but some aspects of possession may not be predictable; (L) largely Held: here is a significant level of possession of an attribute, but the implementation of the related process can vary in different operating units of the organization; (F) Fully Possessed: there is evidence of a full achievement of the attribute, and there aren't differences on how to implement the process among the various operating units.

Levels, related attributes, and related activities needed to assess the five operational areas of the eTcoMM are:

I. *Unrecognized Level:* at this level, the processes for managing operational areas are not recognized as required, and the online communication variables are not a concern for the organization, there aren't processes for managing them.

   a. There are no attributes to be satisfied and no activities foreseen.

II. *Recognized Level:* the organization recognizes the need to improve processes relating to operational areas considered by the model, but has not yet implemented any process.

   b. Attribute of recognition. This attribute requires that the organization's members recognize the importance of the considered operational area. *Activities for each operational area are*: update area, no activities are foreseen but it is recognized the relevance of the employee update on regards eTourism trends. Search Engines area, no activities are foreseen but it is recognized the relevance of search engine marketing and positioning. Back-links area, no activities are foreseen but it is recognized the possibilities to use backlinks strategies for online marketing and promotion. Usages area, no activities are foreseen but it is recognized the possibilities to do usages analysis. eWom area, no activities are foreseen but it is recognized the relevance of the web2.0 and related eWord-of-Mouth

III. *Considered Level:* at this level the staff is aware of the importance of management the processes. Activities are performed informally and without a specific schedule within the organization or by third parties.

c. Attribute of consciousness. This attribute requires that the staff is aware of the possibility of implementing the activities within their operational areas and consider these activities as leading in order to get more results if guided by established processes. *Activities for each operational area are*: update area, the staff is aware of the relevance of the need of its update on eTourism trends. Search Engines area, the staff is aware about the relevance of search engine marketing activities to use in order to improve the tourism destination online communication. Back-links: the staff is aware about the relevance to have back links to their own website. Usages area, the staff is aware about the relevance to know the users and competitors online habits. eWom area, the staff is aware about the communication and marketing possibilities offered by social media websites.

d. Attribute of execution. This attribute requires that activities relating to the considered operational areas have been sporadically implemented (in some cases by third parties), without defining processes to manage them. *Activities for each operational area are*: update area, the staff participates to conference and training activities; purchase specialized magazines; participates to forum and mailing list related to the eTourism domain. Search Engines area, search engine optimization (SEO) and search engines marketing (SEM) activities are implemented but there aren't specific key performance indicators to evaluate the effectiveness of the results obtained. Back-links area, backlinks campaigns are sometimes implemented but the results are not controlled. Usages area, usages analysis is performed but not constantly. This activity can be done by a third party (e.g. external agency). eWOM area, activities performed concern the reading of the online conversations and a not regular online participation.

IV. *Implemented Level:* activities taken into consideration at this level are implemented. Guidelines are defined in order to make the processes repeatable, and are formally recognized as an important issue on which to invest budget. Activities are conducted according to predefined procedures, documented and approved as company regulations.

e. Attribute of process definition. This attribute requires that the processes for managing operational areas are implemented and produce good results. They are routinized and defined by guidelines, making them repeatable. The activities are conducted according to procedures defined, documented and approved as company regulations. This attribute also requires that roles and responsibilities are defined. Economic resources and time are committed to the management and implementation of processes that lead to quality results. *Activities for each operational area are*: update area, training update activities are periodically performed. Search engines area, SEO and SEM are regularly performed. Back-links area, backlink strategies are scheduled and monitored. Usages area, usages analysis are performed internally or by a third party. eWom area, guidelines for the online conversations monitoring are implemented with ad-hoc contents.

V. *Integrated Level:* at this level, the processes for the online communication are integrated with all other management processes within the organization. Based on the analysis of search engines results, usages and online conversations,

business decisions are made, such as the allocation of budget for marketing activities. The activities are then measured to assess the process efficiency and the quality of results.

f. Attribute of integration. This attribute requires that the processes for the management of the operational area are integrated with all other management processes within the organization. *Activities for each operational area are*: update area, training update activities are regularly scheduled and are considered a necessity for a better improvement of others online activities. Search engines area, SEO and SEM are scheduled and are considered a specific activity to be realized within the web department with a dedicated resource. Back-links area, a resource is dedicated for the implementation of the backlinks campaigns and its related monitoring. Usages area, a resource is dedicated for the implementation of the usages analysis and the results are shared among the organization. eWom area, a resource is dedicated for the implementation of the social media activities and the results are shared among the organization.

g. Attribute of improvement. This attribute requires that the results obtained by processes and activities are used to find solutions and improve the performances of the organization. *Activities for each operational area are*: update area, training activities create an improvement on the organization online performances. Search engines area, online strategies are planned based on the search engines results. Back-links area, online strategies are planned based on the back links results. Usages area, online strategies are planned based on the results from the usages analysis. eWom area: online strategies are planned based on the results from the online conversations analysis.

As in Figure 1, this five-level ordinal scale of maturity identifies an ideal path that goes from generic unstructured activities to structured and integrated activities within an organization.

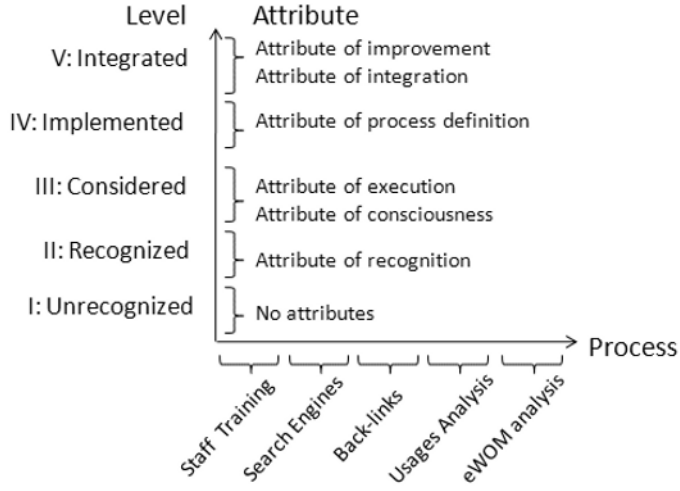


Fig. 1. eTcoMM framework

4 The case of the Canton Ticino and Lombardy

An Italian popular regional tourism destination: Lombardy, and a Swiss tourism region: Canton Ticino, have been chosen. This choice is justified particularly by the fact that Lombardy is the most advanced region considering technology in Italy, and Canton Ticino DMO is leading Swiss Cantons when it comes to eTourism. Among these two regions, seven tourism organizations have been chosen: *Italian DMOs,* 1) Province of Como - Tourism Department; 2) Province of Milan - Tourism Department; 3) Province of Varese - Tourism Agency; *Swiss DMOs*, 4) Lugano Local Tourism Organization; 5) Ascona Locarno Local Tourism Organization; 6) Mendrisiotto e Basso Ceresio Local Tourism Organization; 7) Ticino Tourism - Regional Tourism Organization. Following a desk-research preparation, semi-structured interviews with the directors of Italian and Swiss DMOs were conducted in the period May 2011 - June 2011. The interview protocol covered all the studied areas. The duration of an interview was around 45 minutes. Interviewees were asked to discuss their staff habits regarding the activities provided per each operational area related to eTcoMM. The maturity of the DMOs was assessed as the average value of the five levels of maturity resulted in each area.

4.1 Results

The overall maturity level reached by the DMOs is obtained by averaging the levels reached in each individual operating area. Results show a lack of homogeneity within the sample. While no DMO is just at the first level of maturity (Not Recognized), only one is placed at the second level of maturity (Recognized): Mendrisio. This position implies that for this organization it is important to implement and manage online processes within the operational areas considered, but there isn't at this time an investment for the online related activities. For this DMO the online communication is not currently part of the institution's internal goals, despite a willingness to be active online. From the interviews with the Swiss sample, it emerged the relevance of a third party agency: a regional center devoted to provide strategic and operational support services on online communication through new technologies. Thus a strong delegation to this third party for the outsourcing of the operational areas included in the model was accomplished for the entire Swiss sample.

Ascona Locarno and Como DMOs were located at the third level (Considered): Ascona Locarno is expecting to increase the online presence thanks to a collaboration with the Swiss regional center devoted to the online communication, as for Mendrisio. Como did not plan an immediate improvement on its online communication management due to a lack of resources available for the online communication. Milan and Lugano were located between the third and fourth level (Considered and Implemented) of eTcoMM. A high level of competence of the resources in charge to manage the online communication was presented within the Milan staff. Moreover, an improvement of the web area was also prospected for the Milan DMO in the short term. Varese DMO was at the fourth level of maturity (Implemented), the highest reached by the sample. A webmaster deals in particular with activities related to the usages analysis, promotion via search engines and back-links, these activities are integrated with other functions of the organization.

Among the analyzed DMOs, only one was found at the fifth level of eTcoMM: Ticino Tourism. Ticino Tourism is the Regional DMO of the Canton Ticino, the Italian speaking part of Switzerland. Its key role was due to its involvement with other local tourism destinations from an administrative and legal point of view. Results indicated that Ticino Tourism invested more in both the third pillar of Online Communication Model, the training of its staff in the field of eTourism, as well as in the fourth pillar of the OCM, analysis of users and of related eTourism context.

Regarding the level of maturity reached by DMOs in each operational area included in the model, the areas with a greater maturity are in this order: (i) *Usages analysis* resulted to be the operational area where the DMOs were more mature and aligned between the IV and V level. All DMOs grasp the importance of knowing who is visiting their website. Based on the studies conducted regarding the website usages, important business decisions were made within the organizations for improving their communication both online and offline. (ii) With regard to the *promotion via search engines*, DMOs seemed to act in two different ways: the more mature DMOs (Ticino, Lugano, Varese) were placed at the V level, and have processes in place for the study of their web position, competitors positioning analysis, search engines optimization (SEO) and search engine marketing (SEM) activities. The other DMOs (Mendrisio, Milano, Como, Locarno) carried out the aforementioned activities only in cases of necessity or under certain conditions, e.g. during an important event. (iii) The *promotion using back-links* presented results which show a medium/high maturity. All DMOs seemed to know and use this strategy of promotion, but in most cases they act without guidelines for dealing with it. Back-links are obtained free of charge or through exchanges, and rare were the occasions on which are carried out paid promotional campaigns due to a lack of resources. (iv) As far as for the *staff training*, the results are very different in the sample, but values show a medium/low maturity. It is unlikely that the staff participates regularly in training or conference/workshops or uses magazines, or online groups. Even where this update occurs, there are not defined processes. However, exceptions were Ticino and Milan. (v) *Study of the eWord-of-Mouth* is the operational area where DMOs, with the exception of Ticino Turismo, appeared to be less mature. DMOs were positioned between the second and third level of maturity (recognized and considered). The only activity that seemed to be implemented on eWOM was the creation of pages on specific social media, such as Facebook, Flickr and SlideShare, that contributes to the participation in the online word of mouth.

5 Discussion and Conclusions

This study allowed to build a tool for assessing the maturity level of a DMO regarding some aspects of online communication. Among DMOs, all the considered online communication operational areas were recognized as important for their successful communication. None of the seven DMOs were placed at the first level of maturity (not recognized), and the online communication was perceived as a necessity for the promotion of the destination. The differences found among the analysed DMOs happened to be solely related to the different sizes of the destinations, the resources available and the efforts to communicate the place as a brand. DMOs in fact declared to consider themselves not as a simple provider of information, but as an active

communication player and info source. The third and fourth pillars of the Online Communication Model (people and usages) were widely considered by the DMOs and their websites are becoming a real business communication activity within the organizations. In Tessin (Ticino Region), the regional center in charge of the online activities allows local DMOs to implement web activities and to exploit some of the benefits of online communication. Having a center of expertise which works for more DMOs, it appears to be a viable solution towards the inability to allocate enough budget for the online strategies. Thus, the solution implemented in the area of Canton Ticino, provides an interesting practice for other DMOs, which might think of joining forces to create regional centers of expertise to improve their online performances. DMOs reached a maturity level for four out of five online communication areas, which are: staff training, promotion via search engine, promotion via back-links, usages analysis; while the eWOM analysis is considered an asset less relevant to cover. This indicates that, despite the increasingly importance of the eWord of Mouth especially in the process of decision-making, DMOs are not yet ready to exploit the potential of Web 2.0. Regarding the test of the eTcoMM with the seven DMOs, results show that the model is capable of providing an overview of the current state of maturity of a DMO. The descriptive value of the model is confirmed.

This model could be used by organizations as a starting point for benchmarking research in order to understand the state of the art of their internal processes, and then for a subsequent improvement in their level of maturity. Best practices which are located at the top maturity level, can be taken as inspiration for a better improvement of the online communication strategies. A DMO, through a self-assessment of its online communication maturity on regards the presence or not of the activities provided per each operational area, can be more aware of which area is more competitive, or requires more efforts. Limitations of this research are related to the limited sample, and to the limited operational areas considered as variables in the model. For these reasons, future research could be carried out by referring to a larger sample, in particular dividing the DMOs based on the extension of the territory covered by them, depending on the DMO organizational and management model (e.g. public, public/private, internal organization, office size). Additional variables may also be included within the model, allowing to focus on new aspects of online communication. Moreover, it would be possible, while maintaining the same structure, to add new levels or move existing ones, e.g. level III could become the lowest level of the model, as the actual level I can become obvious and taken for granted by all DMOs. In this way the eTcoMM will be an adaptive model which changes over time, according to the changes that occur in the online communication field.

References

Buhalis, D. (2003). *eTourism: Information technology for strategic tourism management*. Harlow: Prentice Hall.

Buhalis, D. & 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* 29: 609-623.

Cantoni, L. & Ceriani, L. (2007). *Fare comunicazione online.* Roma: Comunicazione Italiana.

Cantoni, L. & Tardini, S. (2006). *Internet*, *Routledge Introduction to Media and Communication.* London: Routledge.

Cantoni L. & Tardini S. (2010). The Internet and the Web. In D. Albertazzi and P. Cobley (Eds), The media. An introduction (3rd ed.). New York et al.: Longman.

Chung, J. Y. & Buhalis, D. (2008). Web 2.0: A study of online travel community. In *Information and Communication Technologies in Tourism.* Vienna: Springer.

Crosby, P.B. (1979). *Quality is Free*. New York: McGraw-Hill.

Doolin, B., Burgess, L. & Cooper, J. (2002). Evaluating the use of the Web for tourism marketing: a case study from New Zealand. *Tourism Management* 23: 557-561.

Earthy, J. (1998). *Usability maturity Model: Human-Centredness Scale*. Report of the Telematics Applications Project IE, Version 1.2, Information Engineering Usability Support Centres.

El Emam, K. & Jung, H. (2001). An empirical evaluation of the ISO/IEC 15504 assessment model. *The Journal of Systems and Software* 59: 23-41.

Fesenmaier, D. R., Xiang, Z., Pan, B. & Law, R. (2010). An analysis of search engine use for travel planning. In U. Gretzel, R. Law & M. Fuchs (Eds.), *Information and Communication Technologies in Tourism 2010*, Proceedings of the International Conference in Lugano, Switzerland.

Garcia, S.M. (1997). Evolving Improvement Paradigms: Capability Maturity Models and ISO/IEC 15504 (PDTR). *SOFTWARE PROCESS – Improvement and Practice* 3: 47-58.

Govers, R. & Go, F. (2009). *Place Brandin. Glocal, Virtual and Physical Identities, Constructed, Imagined and Experienced*. Basingstoke Hampshire: Palgrave Macmillan.

Gretzel, U. & Fesenmaier, D. R. (2004). Implementing a Knowledge-Based Tourism Marketing Information System: the Illinois Tourism Network. *Information Technology & Tourism* 6: 245–255.

Gretzel, U. & Yoo, K. H. (2008). Use and Impact of Online Travel Reviews. In *Information and Communication Technologies in Tourism.*Vienna: Springer.

Inversini, A. & Buhalis, D. (2009). Information Convergence in the Long Tail: The Case of Tourism Destination Information. In *Information and Communication Technologies in Tourism*. Vienna: Springer.

Nolan, R.L. (1973). Managing the Computer Resource: A Stage Hypothesis. *Communications of the ACM* 16: 399-405.

O'Connor, P. (2008). User generated Content and Travel: A case Study on Tripadvisor.Com. *Information and Communication Technologies in Tourism*. Vienna: Springer.

O'Reilly, T. (2005). *What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software*. Retrieved from http://www.oreillynet.com/ pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html

Paulk, M. C. et al. (1993). *Capability Maturity Model for Software, Version 1.1*. Pittsburg: Software Engineering Institute/Carnegie Mellon University.

Paulk, M. C. (1998). Analyzing the Conceptual Relationship Between ISO/IEC 15504 (Software Process Assessment) and the Capability MaturityModel for Software. In *Proceedings of the Ninth InternationalConference on Software Quality*, Cambridge.

Steinbauer, A. & Werthner, H. (2007). Consumer Behavior in e-Tourism. In *Information and Communication Technologies in Tourism*. Vienna: Springer.

Werthner, H. & Klein, S. (1999). *Information Technology and Tourism – A Challenging Relationship*. Vienna: Springer.

Xiang, Z. & Gretzel, U. (2010). Role of social media in online travel information search. *Tourism Management* 31: 179-188.

Tourism Training: An Investigation of Virtual Learning Experience in the Context of a Virtual World

Yu-Chih Huang,
Sheila J. Backman, and
Lan-Lan Chang

Department of Parks, Recreation and Tourism Management
Clemson University, USA
{yhuang; back, lchang}@clemson.edu

Abstract

3D virtual environments offer potentials as an educational and training tool for tourism industry professionals and educators. The main goal of this study is to investigate 3D virtual learning experiences in the context of tourism. Self-Determination Theory was used to guide this study. Participants were recruited from an introductory tourism course and a tourism marketing class in the Department of Park, Recreation and Tourism Management at a university in southeast region of the United States. The results of this study found that the satisfaction of the psychological needs of autonomy, competence, and relatedness are positively related to intrinsic motivation and positive emotions while learning in the virtual worlds. Thus, in order to implement successful virtual-learning courses in the virtual worlds, tourism educators and professionals need to make efforts to foster the satisfaction of psychological needs in the virtual learning environment.

Keywords: Virtual Learning Experiences; Self-Determination Theory; Tourism Training; Intrinsic Motivation; Positive Emotions.

1 Introduction

Virtual learning is widely used in the tourism business for delivering training opportunities developed to enhance the skills and knowledge of personnel (Buhalis & Law, 2008). As Buhalis and Law (2008) pointed out, virtual learning environments can be used to support classroom teaching, stimulate discussion, and facilitate course administration. Collins, Buhalis and Peters (2003) asserted that due to the limited resources to send employees to expensive training courses, it is important for smaller tourism businesses to utilize e-learning for cost efficient training delivery and flexible work arrangements. In addition, Penfold (2008) indicated that virtual environments provide the opportunity to engage students effectively in hospitality and tourism education. Singh and Lee (2008) pointed out that tourism educational institutions need to adopt and use the next-generation technology tools for effectively preparing future tourism and hospitality students. 3D Virtual worlds are presenting opportunities for creating interactive environments in an educational context (Eschenbrenner, Nah, & Siau, 2008).

Educators in travel and tourism fields have started to explore the potential of e-learning in virtual environments and to take advantage of effective virtual environments to obtain best results in the learning process. Penfold (2008) pointed out that virtual worlds utilizing gaming concepts and virtual reality offer real-world simulations for developing creative learning spaces. Jennings and Collins (2007)

asserted that educational institutions across the globe have begun to use virtual environments for experiential learning and collaboration from different disciplines, regardless of geographical boundaries.

Currently, one of the most popular and active 3D virtual environment platforms in education is Second Life. Mason and Moutahir (2006) claimed that the emerging virtual environment development platform of Second Life offers the unique features of immersion, ease of use, wide availability and low barrier to entry for its use in education. Baker, Wentz, and Woods (2009) pointed out that in the virtual world of Second Life, not only do there exist numerous virtual campuses, but companies have also presented potential uses in training for holding lectures (Harvard University's law school), arranging meetings and conferences (IBM, InterContinental Hotels Group), conducting new employee orientation (TMP Worldwide) as well new student orientation (Hong Kong Polytechnic University), hosting recruiting events (Clemson University), and creating simulation environments (Walt Disney Internet Group).

However, the increased use of the virtual environment as an educational and training tool is not occurring without controversy. Some scholars have argued that researchers need to be aware of a number of experience and capacity issues that have arisen as well as negative effects including the following: requiring advanced technology resources, appropriate training and orientation and needing adequate time for familiarity with the environment (Eschenbrenner, Nah, & Siau, 2008); limiting certain students who prefer to learn in a regular classroom setting with face to face communication (McGrath, 1998); experiencing dissatisfaction with web based courses because of lack of interaction between instructors and students (Arbaugh, 2002); feeling isolation in online courses resulting from diminished interaction with others (Hay, Hodgkinson, Peltier, & Drago, 2004); and encountering virtual violence, assault, and sexual harassment in virtual learning environments as well as influences by disruptive players (Kluge & Riley, 2007).

In contrast, other scholars have suggested that positive effects can be derived from online learning environments such as providing opportunities for social interaction and engaging participation in virtual learning (Barab, Thomas, Dodge, Carteaux, & Tuzan, 2005); offering a unique pedagogical environment for interactive experience, engagement, interactivity, and idea generation (Eschenbrenner, Nah, & Siau, 2008); fostering students' motivation and learning of scientific inquiry skills (Dede, Clarke, Ketelhut, Nelson, & Bowman, 2005); allowing for creativity within a multi-media environment (Conway, 2007); teaching a multitude of skills of problem solving, and strategic thinking (Prensky 2006); and supporting experiential learning and global collaboration (Jarmon, Traphagan, Mayrath, & Trivedi, 2009). Given the diversity and complexity of virtual learning environments, it is evident that online learning has potential to generate both positive and negative effects to students. Regardless of such debates, it is clear that 3D virtual learning environments have tremendous appeal, and as Penfold (2008) stated, virtual worlds offer a new dimension to distance learning with enjoyable, interactive and engaging experiences for effectively preparing the new generation of learners who have been significantly influenced by technologies like i-phones, digital media and online networks.

Jennings and Collins (2007) indicated that educational institutions have established research centers for assessing the pedagogical potential of virtual worlds. Yet, as emerging virtual worlds have the potential to be used by tourism professionals and educators for providing interactive and meaningful learning experiences, little research has been undertaken to understand the motivational dynamic of learning in 3D virtual environments.

2 Literature Review

Self-Determination Theory (SDT) is an approach to human motivations and personality in understanding the importance of humans' evolved internal resources for personality development and behavioral self-regulation (Ryan, Kuhl, & Deci, 1997). According to Ryan and Deci (2000), SDT investigates people's innate growth tendencies and psychological needs that are the basis for self-motivation and personality integration. It identifies three psychological needs for facilitating optimal functioning of the natural propensities for growth, integration, social development and personal well-being. These needs are competence, autonomy and relatedness.

Research by Ryan, Rigby, and Przybylski (2006) empirically tested Self-Determination Theory in relation to gaming motivation based on the idea that players seek to satisfy their psychological needs in the context of play. In their study, they pointed out that game features of virtual environments provide opportunities for players to experience autonomy, competence and relatedness, suggesting that the SDT framework is useful and practical in understanding psychological components of gaming experience. Furthermore, Przybylski, Ryan and Rigby (2009), investigating gaming motivation, confirmed that virtual contexts of video games can engender the experiences of autonomy, competence and relatedness, indicating that SDT enhances the understanding of game play in virtual environments. The psychological needs of autonomy, competence, and relatedness as specified by Self-Determination Theory were selected for this study, because as past research suggests, these three psychological needs are important aspects of the virtual experience in 3D virtual worlds.

In the past decade, several studies have examined the relationship between evoked positive emotions and student-learning experience in web based learning environments. Based on the broaden and build theory (Fredrickson's 1998), Teixeira, Wedel and Pieters (2010) proclaimed that positive emotions of joy and surprise led to exploration and concentration of attention to Internet video ads that prolong people's retention in Web environments. In their study, positive emotions of joy and surprise directly influenced people's retention of Internet video ads and indirectly affected concentration of attention that led to exploration and play in the subsequent tasks. Exploring the experiential aspects of hedonic consumption, Holbrook, Chestnut, Oliva and Greenleaf (1984) examined the interrelationships among emotions, personality, and performance in the enjoyment of video games and postulated that positive emotions of pleasure, arousal, and dominance contributed to enjoyment of game play; they also observed that the appeal of positive emotion depends on personality-game congruity as well as perceived complexity and past experiences.

O'Regan (2003), investigating the relationships between emotion and cognition in the context of online education, identified how emotions of frustration, fear, anxiety, apprehension, embarrassment, enthusiasm/excitement, and pride are associated with online learning. He asserted that positive emotions play a crucial role in the learning experience of online environments. For instance, the positive emotions of excitement, pride and enjoyment contributed to students' engagement and motivation to learn in online learning scenarios. The study exploring positive emotion by Park and Lim (2007) examined the significance and effects of positive emotion in multimedia instructional learning, proclaiming that the experiences of positive emotion play an important role in enhancing students' learning interest, achievement, and motivation. Reviewing the current literature, this study has found the constructs of intrinsic motivations, and positive emotions, are important aspects for understanding the virtual learning experience.

3 Research Purpose

Self-Determination Theory (SDT; Deci & Ryan, 1985; Ryan & Deci, 2000) has been applied as a theoretical framework for examining human motivational behavior. Recently, a small number of studies have validated the framework of SDT in the context of video gaming, establishing that the satisfaction of psychological needs leads to sustained engagement with games and subsequently how those qualities can facilitate intrinsic motivation (Ryan et al., 2006; Przybylski et al., 2009; Tamborin et al., 2010).

Recent research by Ryan et al. (2006) applied Self-Determination Theory to understand psychological impacts and human motivations in the contexts of video gaming. Moreover, Rigby and Przybylski (2009) indicated that Self Determination Theory could be an imperative framework to bridge the gap between education and virtual worlds. These preliminary applications of the theoretical framework of SDT suggest that the quality of need satisfaction for a virtual context can have implications for explaining motivational dynamics in the virtual worlds such as Second Life.

Based on the framework of Self Determination theory (SDT; Deci & Ryan, 1985; Ryan & Deci, 2000), the main goal of this study is to investigate how the psychological needs of competence, autonomy, and relatedness can be used to understand 3D virtual learning experiences in the context of tourism. Specifically, this study investigates the impacts of virtual experience in Second Life on satisfaction of psychological needs, which in turn predict consequent motivation to learn and positive emotions. Two research questions were proposed for this study: (1) To what extent does the 3D virtual learning environment satisfy the needs of autonomy, competence, and relatedness; and (2) how is needs satisfaction related to intrinsic motivation and positive emotions toward learning while experiencing 3D virtual worlds.

4 Methodology

This exploratory study employed a convenience sampling method. The participants were recruited from an introductory tourism course and a tourism marketing class in the Department of Park, Recreation and Tourism Management at a University in the

southeast region of the United States. In-class announcements and emails about the research participation were made by the course instructors and the researcher. Students then signed up for three one-hour class sessions to take part in a study of virtual tourism in Second Life. The first two sessions focused on orientation and training and the third one involved taking part in a virtual field trip to Second Life Maasai Mara.

The Maasai Mara basecamp in Second Life was selected as the research site for this study. A virtual representation of the Maasai Mara Basecamp was created by Clemson University in Second Life and designed to provide information to create visitor awareness about the Maasai Mara tribe, promote tourism to Maasai Mara in Kenya, and preserve the culture of the Maasai Mara community.

The constructions in virtual Maasai Mara consist of the reception area of the Basecamp building, guest accommodations area (tent), local jewelry shop, bonfire area and wild animals. In creating a virtual representation of Maasai Mara Basecamp, the significance of the cultural aspects including clothing and jewelry was also established to communicate the symbolism of the Maasai culture through multi-media communication channels like implementing videos, music, and slideshows within the virtual environment.

Primary data were obtained for this study through self-administered web questionnaires (Snap software) to collect information from participants. The data collection was conducted in March and April, 2011 in a computer lab at a university in the southeast region of the United States under the supervision of two graduate students and a faculty member. Participants were scheduled in one-hour staggered time periods to reduce the risk of burdening the Second Life servers. In total, 105 usable surveys were collected.

Participants met for 3 one-hour class sessions to take part in the study. In the first two one-hour sessions, students were provided orientation and training experiences. It was important to develop their skills of movement and communication so that they would be able to complete the virtual learning experience. An introduction to the virtual environment of Second Life was provided to students within a one-hour orientation period and a one-hour training session.

During the initial orientation period, participants took part in activities that showed them how to create a Second Life account and prepare an avatar to participate in the virtual experience. A training handbook was provided for participants to practice the basics of navigation and communication skills in Second Life. After these initial preparations for participation in experiencing virtual environments, participants took part in another one-hour training period to become familiar with the control interface of Second Life.

The third session involved the virtual field trip in which participants explored different areas in Second Life Maasai Mara, interacting with multi-media information within the virtual tourism attraction, interacting with others on the site, and learning about the cultural aspects of the tourism destination. The purpose of this virtual field trip activity was to engage participants by accomplishing the following tasks: visiting the jewelry shop to look at the Maasai jewelry displayed, riding an animal to the river,

listening to Maasai Music in the guest tents, and watching a video of two Maasai warriors telling a lion killing story in front of a bonfire.

Measurements

Applying Self-Determination Theory, Ryan et al. (2006) developed Player Experience Need Satisfaction (PENS) to assess the degree to which the video game players experienced the satisfaction of the three psychological needs: in-game competence, in-game autonomy, and in-game relatedness. The PENS scale used by Ryan et al. (2006) revealed a good reliability. Thus, the measure of psychological needs, specifically regarding competence, autonomy, and relatedness, was modified from Ryan et al.'s (2006) PENS scale for use in this study to better reflect the needs of this study.

The PENS subscale of competence is comprised of three items focused on the competence, mastery and efficacy. Besides the PENS subscale, autonomy was assessed with a three-item subscale measuring the degree to which participants experienced choice, freedom and perceived opportunities to participate in activities that interested them. The PENS subscale for perception of relatedness was modified for use in this study. It consisted of four items assessing how connected participants felt to others in Second Life. Competence, autonomy, and relatedness were assessed on a 7 point Likert-type scale for each item.

The assessment of intrinsic motivation was modified from the Intrinsic Motivation Inventory (IMI) to reflect the needs of this study. Intrinsic Motivation was assessed with a four item scale. Intrinsic motivation assessment was rated on a seven-point Likert-type scale ranging from Strongly Disagree to Strongly Agree (strongly disagree =1 to strongly agree=7).

The assessment of positive emotions, based on the Positive and Negative Affect Schedule (PANAS) developed by Watson, Clark, and Tellegen (1988), was modified for use in this study. Positive emotions were measured using a 7 item scale. The seven item scale reflected the extent to which a person feels enthusiastic and active, including amusement, contentment, interest, joy, pride, cheerfulness, and delight. Participants were asked to indicate the extent to which they felt each of 7 emotions while they were experiencing Second Life Maasai Mara on a five point Likert scale (1= very little or not at all to 5= extremely or all the time).

5 Results

To examine the internal consistency of measurements used in this study, reliability tests were performed. Table 1 presents the composition of the dimensions and measures of internal consistency for all key constructs of this study including competence, autonomy, relatedness, intrinsic motivation, and positive emotions. The results of the reliability test showed Cronbach's alpha coefficients ranging from 0.83 to 0.96, indicating that all five variables had an adequate reliability (Nunnally, 1978).

Table 1. Reliability of the Scales

Scales	Items	Mean	S.D.	Cronbach's Alpha
Competence	3	5.14	0.87	0.829
Autonomy	3	4.98	1.28	0.887
Relatedness	4	3.34	1.46	0.920
Intrinsic Motivation	4	5.07	1.28	0.959
Positive Emotions	7	3.29	0.87	0.948

Pearson's correlations coefficients were used to examine the relationships among all key constructs. Table 2 presents a correlation matrix of all key constructs. Correlation analysis revealed that during learning in the virtual world of Second Life, the psychological needs of competence, autonomy, and relatedness were positively correlated with intrinsic motivation ($r= 0.651$, 0.763 and 0.588 respectively, $p<0.01$). The results of correlation analysis revealed that during learning within the virtual world of Second Life, the psychological needs of competence, autonomy and relatedness were significantly correlated to positive emotion ($r= 0.502$, 0.611, and 0.551 respectively, $p<0.01$).

Table 2. Correlation Matrix Among The Constructs

	CM	AU	RL	IM	PE
CM	—				
AU	0.713**	—			
RL	0.401**	0.550**	—		
IM	0.651**	0.763**	0.588**	—	
PE	0.502**	0.611**	0.551**	0.684**	—

CM: competence, AT: autonomy, RL: relatedness, IM: intrinsic motivations, PE: positive emotions

** P< 0.01, * P< 0.05

Further, regression analyses were conducted to examine the relationships between psychological needs satisfaction, learning motivation and positive emotions. In order to determine the impact of multicollinearity in multiple regression analyses, Tolerance and Variance Inflation Factors (VIF) were used to examine for the regression models. Examining the effect of inflation, the VIF for three independent variables are close to 1.00 and the score of tolerance showed that little variance is correlated among three independent variables. There is no sign to indicate the effect of inflation on regression models.

A linear regression model was performed to examine the relationships among psychological needs of competence, autonomy and relatedness and intrinsic motivation while learning in the virtual world of Second Life. The result of the regression showed that the adjusted r^2 of the model on intrinsic motivation is 0.628. The model of competence, autonomy and relatedness on the effect of intrinsic motivations is significant, $F=54.49$, $P<0.001$. The model explained 62.8% variance of dependent variable as shown in the following Table 3.

Table 3. Regression Analysis: Psychological Needs and Intrinsic motivation

Independent Variables	B	Beta	Sr2	F	R2	Adjusted R2
Competence	0.25	0.21**	0.022	54.49***	0.640	0.628
Autonomy	0.49	0.50***	0.102			
Relatedness	0.19	0.21***	0.032			

*** P< 0.01, ** P< 0.05, *P<0.1

The standardized regression coefficient for competence is 0.21 (β= 0.21, p<0.05), revealing a positive relationship between competence and intrinsic motivation. Sr2 is 0.022, which indicates that 2.2% variance of intrinsic motivation can be explained by competence. The standardized coefficient (β= 0.50) for autonomy was significant and positive (P<0.01), revealing positive association between autonomy and intrinsic motivation. It has a unique effect size of 0.102 (Sr2), revealing that autonomy explained 10.2 % variance of intrinsic motivation. The standardized coefficient (β= 0.21) for the effect of relatedness on intrinsic motivation was significant and positive (P<0.01), indicating relatedness positively related to intrinsic motivation. It has a unique effect size of 0.032 (Sr2=0.032), indicating that relatedness explained 3.2 % variance of the dependent variable.

The regression model for examining the effect of psychological needs on positive emotions had a good fit (F= 24.13, P<0.001) and explained variance with an adjusted R2=0.43 as seen in following Table 4. The result of regression revealed that perceived autonomy was the most important contributor in predicting positive emotion (β=0.361, p<0.001). The regression model had adjusted R^2 of 0.43, indicating the model explained about 43% of the variance. There was significantly positive association between autonomy and positive emotion while learning in the virtual world of Second Life. In addition, the results of regression analysis revealed positive relationship between relatedness and positive emotions (β=0.310, p<0.001).

Table 4. Regression analysis for positive emotions

Independent Variables	B	Beta	Sr2	F	R2
Competence	0.109	0.137		38.68***	0.40
Autonomy	0.226	0.336***	0.046		
Relatedness	0.187	0.310***	0.067		

*** P< 0.01, ** P< 0.05, *P<0.1

6 Discussion and Implications

3D virtual environments offer potential as an educational and training tool for tourism industry professionals and educators. According to Sigala (2002), virtual learning provides great flexibility that overcomes time and space barriers to match the specific working conditions within the tourism sector. Second Life provides an opportunity to develop interactive experiences to deliver employee-training programs and to teach tourism students about next generation technology tools (Law, Leung, & Buhalis, 2009; Singh & Lee, 2008).

As the applications and opportunities of virtual worlds in education continue to rapidly emerge, an understanding of the potential and influence of Second Life in the context of tourism education and training still needs to be identified and fully discovered.

Based on Self-Determination Theory, the goal of this study was to investigate the impacts of virtual experiences in Second Life on students' learning motivations and enhanced positive emotions in virtual learning. This study represents an empirical application of self-determination theory that is a theoretical and practical framework to examine how 3D virtual learning content makes a contribution to intrinsic motivation and psychological needs satisfaction of competence, autonomy and relatedness. Results of the studies revealed consistently sturdy relationships between learners' virtual experience of psychological need satisfaction and motivation for virtual learning, supporting the general principles of the SDT (Deci & Ryan, 1985). Additionally, the present study explores the role and significance of positive emotion in the context of virtual learning, specifically on the relationship between the needs satisfaction of autonomy, competence, relatedness and positive emotion.

The first question is concerned with how satisfaction of the psychological needs as specified within Self-Determination Theory is associated with students' intrinsic motivations in virtual learning. Previous studies have demonstrated the relationship between psychological needs and integration of motivations in leisure, education, and social development (Ryan & Deci, 2000). The results of present study found that the satisfaction of the psychological needs of autonomy, competence, and relatedness are positively related to intrinsic motivation while learning in the virtual worlds, indicating that greater user perception of the ability to overcome the optimal challenges fulfills the need of competence, thereby leading to enhancement of intrinsic motivations; higher levels of perception to participate in activities that interested users can enhance learning motivation; and a greater sense of feeling connecting with others corresponds with a higher level of intrinsic motivation.

The findings of this study suggest that the fulfillment of three psychological needs contributes to enhancing intrinsic motivation during learning in the virtual worlds, while the need of autonomy fulfillment is of great importance in predicting learning motivations. This study added to evidence from the works of Ryan et al. (2006) that the enhancement of intrinsic motivation is a direct consequence of psychological need fulfillment.

The second question focuses on how the satisfaction of psychological needs is related to positive emotions while learning in the 3D virtual world of Second Life. Further, the results showed that positive emotions were significantly predicted by autonomy and relatedness, suggesting a positive relationship between psychological need satisfaction of relatedness and competence and positive emotions, which is consistent with the previous work by Ryan et al. (2006). The results obtained here indicated that those who experienced relatedness satisfaction while learning in the virtual world of Second Life experienced higher levels of positive emotions whereas individuals who were generally more autonomous in their virtual learning also experienced overall higher positive emotions for virtual learning.

From a tourism educator point of view, this study can be applied to build the connection between tourism education and virtual worlds in the development of how to construct engaging and interactive learning experiences to meet the expectations of learners. With the effect of engaging learners in a 3D virtual learning environment, the psychological benefits can influence learners' motivation for learning, ultimately leading to a high level of learning achievement. In order to implement successful e-learning courses in Second Life, educators and their institutions need to make efforts to foster the satisfaction of psychological needs in the virtual learning environment.

For instance, Second Life Maasai Mara provides opportunities to participate in activities wherein interested virtual tourists could increase their sense of autonomy such as riding animated animals and listening to Maasai music. Moreover, the videos within the virtual site communicating Maasai culture and traditions of local tribes offer virtual tourists the opportunity to feel connected with the Maasai village, thus leading to enhancement of the perception of relatedness.

7 Limitations

A number of potential imitations need to be considered in this study. First, the data were collected exclusively from a university and comprised mainly of college students, and participants were determined by using convenience sampling within a major of Park, Recreation and Tourism Management. The sample may not be reflective of the broader population. For future study, it is suggested that the sample of the study should be extended to cross validate diverse populations across different cultures and age groups. In addition, due to the exploratory nature of the study, the results of the study were interpreted on the application of virtual world of Second Life and only in a particular tourism destination.

Generalizability should not be expected. Future study should include more diverse virtual world platforms such as Active World or OpenSim to better understand 3D virtual learning experience in the context of tourism. In addition, there are many other factors that can influence students' learning experience, increase motivation to learn, and enhance positive emotions in virtual worlds; future research can study the impacts of other variables such as individual differences, previous experiences with technology, student attitudes toward the use of information communication technology, and instructor teaching style to continually investigate the relative influences of learning motivation and engagement in virtual environments.

References

Baker, S. C., Wentz, R. K. & Woods, M. M. (2009). Using Virtual Worlds in Education: Second Life as an Educational Tool. *Teaching of Psychology* 36(1): 59-64.

Barab, S., Thomas, M., Dodge, T., Carteaux, R. & Tuzun, H. (2005). Making learning fun: Quest Atlantis, a game without guns. *Educational Technology Research and Development* 53(1): 86–107.

Buhalis, D. & 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* 29(4): 609-623.

Clark, L. A. & Watson, D. (1986). Diurnal *variation in mood: Interaction with daily events and personality*. Paper presented at the meeting of the American Psychological Association (August), Washington, DC.

Collins, C., Buhalis, D. & Peters, M. (2003). Enhancing SMTEs' business performance through the Internet and e-learning platforms. *Education + Training* 45 (8/9): 483-494

Deci, E. & Ryan, R. (1985). *Intrinsic Motivation and Self-Determination in Human Behaviour*. New York: Plenum.

Deci, E. L., Vallerand, R. J., Pelletier, L. G. & Ryan, R. M. (1991). Motivation and education: The self-determination perspective. *Educational Psychologist* 26: 325-346.

Dede, C., Clarke, J., Ketelhut, D. J., Nelson, B. & Bowman C. (2005). Students' motivation and learning of science in a multiuser virtual environment. In *Proceedings of the American Educational Research Association Conference*, April 11–15, Montreal, Canada.

Eschenbrenner, B., Nah, F. F. & Siau, K. (2008). 3-D Virtual Worlds in Education: Applications, Benefits, Issues, and Opportunities. *Journal of Database Management (JDM)* 19(4): 91-110.

Frederick, C. M. & Ryan, R. M. (1995). Self-determination in sport: A review using cognitive evaluation theory. *International Journal of Sport Psychology* 26: 5-23.

Fredrickson, Barbara L. (1998), What Good Are Positive Emotions? *Review of General Psychology* 2(3): 300-319.

Holbrook, M. B., Chestnut, R. W., Oliva, T. A. & Greenleaf, E. A. (1984). Play as a Consumption Experience: The Roles of Emotions, Performance, and Personality in the Enjoyment of Games. *Journal of Consumer Research* 11(2): 728-739.

Jarmon, L., Traphagan, T., Mayrath, M. & Trivedi, A. (2009). Virtual world teaching, experiential learning, and assessment: An interdisciplinary communication course in Second Life. *Computers and Education* 53(1): 169-182.

Jennings, N. & Collins, C. (2007). Virtual or Virtually U: Educational Institutions in Second Life. *International Journal of Social Sciences* 2(3): 180-187.

Law, R., Leung, R. & Buhalis, D., 2009, Information Technology Applications in Hospitality and Tourism: A Review of Publications from 2005 to 2007. *Journal of Travel and Tourism Marketing* 26(5): 599-623.

Mason, H. & Moutahir, M. (2006). Multidisciplinary experiential education in Second Life: A global approach. In *Proceedings of the Second Life education workshop*, San Francisco, CA.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

O'Regan, K. (2003). Emotion and E-learning. *Journal of Asynchronous Learning Networks* 7(3): 78-92.

Penfold, P. (2008). Learning through the world of second life - a hospitality and tourism experience. *Journal of Teaching in Travel and Tourism,* 139-160.

Przybylski, A., Ryan, R. & Rigby, C. (2009). The Motivating Role of Violence in Video Games. *Personality and Social Psychology Bulletin* 35(2): 243-259.

Rigby, C. S. & Przybylski, A. K. (2009). Virtual worlds and the learner hero: How today's video games can inform tomorrow's digital learning environments. *Theory and Research in Education* 7*:* 214-223.

Ryan, R. & Deci, E (2000). Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *American Psychologist* 55(1): 68-78.

Ryan, R. M., Kuhl, J. & Deci, E. L. (1997). Nature and Autonomy: Organizational View of Social and Neurobiological Aspects of Self-Regulation in Behavior and Development. *Development and Psychopathology 9:* 701-728.

Ryan, R. M., Mims, V. & Koestner, R. (1983). Relation of reward contingency and interpersonal context to intrinsic motivation: A review and test using cognitive evaluation theory. *Journal of Personality and Social Psychology* 45(4): 736-750.

Ryan, R. M., Rigby, C. & Przybylski, A. (2006). The Motivational Pull of Video Games: A Self-Determination Theory Approach. *Motivation and Emotion* 30: 344-360.

Sanghoon, P. & Jung, L. (2007). Promoting Positive Emotion In Multimedia Learning Using Visual Illustrations. *Journal of Educational Multimedia and Hypermedia* 16(2): 141-162.

Sigala, M. (2002) The Evolution of Internet Pedagogy: Benefits for Tourism and Hospitality Education. *Journal of Hospitality, Leisure, Sport and Tourism Education* 1(2): 29-45.

Singh, N. & Lee, M. J. (2008). Exploring Perceptions toward Education in 3D virtual Environments: An Introduction to "Second Life*". Journal of Teaching in Travel and Tourism* 8(4): 315-327.

Teixeira, T., Wedel, M. & Pieters, R. (2010). *Emotion-Induced Engagement In Internet Video Ads*. Retrieved from http://sites.som.yale.edu/nemc/files/2010/09/Teixeira-Thales-Harvard-NEMC-2010.pdf.

Vuorela, M., and Nummenmaa, L. (2004). Experienced emotions, emotion regulation and student activity in a web-based learning environment. *European Journal of Psychology of Education - EJPE* (Instituto Superior de Psicologia Aplicada) 19(4): 423-436.

Watson, D. (1988). Intraindividual and interindividual analyses of positive and negative affect: Their relation to health complaints, perceived stress, and daily activities. *Journal of Personality and Social Psychology* 54(6): 1020-1030.

Watson, D., Clark, L. A. & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology 54*(6): 1063-1070.

Potential-of-Interest Maps for Mobile Tourist Information Services

Yohei Kurata

Department Tourism Science
Tokyo Metropolitan University, Japan
ykurata@tmu.ac.jp

Abstract

To prevent tourists from being overwhelmed by a flood of tourist information, this paper proposes an alternative approach to tourist information services, namely *potential-of-interest maps*. Potential-of-interest maps visualize the sightseeing potential of each place in a tourist area, which is estimated from data about locations where previous visitors have found something impressive. Such maps provide visual aids for tourists and will save their effort at decision making. In this paper, we first compare five possible approaches for obtaining data about locations where previous visitors have been impressed, among which the use of photo geo-tags is concluded the most promising. We then construct an example of potential-of-interest maps making use of Flickr's photo data. Finally, we develop a prototype of a mobile tourist information tool which features a potential-of-interest map. In our user test, nine of twelve test users show an intention to use this tool again at other destinations.

Keywords: mobile tourist information services, potential-of-interest maps, photo-taking locations, kernel density estimation, user participation

1. Introduction

Tourism is an activity in which people need a large amount of information, as it is performed in unfamiliar environments. For this reason, many engineers and researchers have undertaken the challenge to provide information aids for tourists, making use of rapidly-advancing mobile information technologies (e.g., Zipf & Manaka 2001). In addition, AR and ubiquitous technologies, which bridge the real world and information in a more seamless way, are increasing the capabilities of mobile tourist information services. However, there is also a concern about mobile tourist information services, because they may confuse tourists by provide too much information through a tiny screen. To tackle this problem, computational techniques for selecting and reorganizing information based on user profiles, such as collaborative filtering (Resnick 1994), have been applied to the domain of tourist information services (for instance, Ricci et al. 2002). However, such techniques usually require a certain amount of time and effort to obtain the user's profile. In addition, depending on destinations and party members, a tourist's preferences may change from time to time, which makes the problem more difficult (Kurata 2011).

In this paper, we introduce an alternative approach for mobile tourist information services, called *potential-of-interest maps*. Our key idea is to visualize the sightseeing potential (or *potential-of-interest*) of places in a tourist area, based on data about locations where previous visitors have found something impressive. The map that illustrates potential-of-interest is highly useful for tourists, as they can visually identify which places will be worth visiting and on which route they can enjoy such places efficiently, without reading unnecessary text. In this paper, we discuss how we

should construct such potential-of-interest maps, as well as how we can employ these maps in mobile tourist information services. The main contribution of this paper is to propose a new layer of a tourist map which promotes the users' *visual thinking* in their tour planning, while many of existing tourist information tools depends on textual information.

The remainder of this paper is organized as follows: Section 2 explains the concept and merits of potential-of-interest maps in detail. Section 3 compares five possible approaches for determining locations where visitors have found something impressive. Section 4 constructs a sample potential-of-interest map, with which we develop a prototype of a tourist information tool in Section 5. Section 6 reports the result of our user test. Finally, Section 7 concludes with a discussion of future work.

2. Potential-of-Interest Maps: A Visual Aid for Tourists

In a potential-of-interest map, each point is painted in deeper colours as more tourists found something impressive at and around the corresponding location in the real world. Potential-of-interest maps can achieve a certain level of tourist guides without textual information. For instance, a deeply-coloured spot probably indicates the presence of popular tourist attractions there. A linearly-stretched coloured area probably corresponds to a major corridor where many tourists come and go. Isolated coloured spots may indicate hidden places-of-interest which attract adventurous tourists. With such implications, tourists can think visually about where they should go and which route they should take, without reading unnecessary text.

The absence of unnecessary text, if it is carefully designed, may stimulate users' inquisitive spirits. People who visit a high-potential location will try to find what is interesting there, if it is not obvious, by awakening their five senses. Interesting findings will not only satisfy them, but also make them perceive a sense of community with previous visitors who have found that place. Such a tourist information service will attract adventurous tourists who consider other services officious and noisy.

It is also nice, especially for novice users, to supplement potential-of-interest maps with certain textual information. For instance, we can develop a digital version of a potential-of-interest map, in which information about each tourist attraction appears when the user clicks the corresponding point on the map. An important feature of such text-enhanced potential-of-interest maps is that the painted pattern of potential-of-interest helps the users to decide which tourist attractions they should consider in their tour planning; in other words, they can visually identify which attractions are not worth reading their information. In this way, potential-of-interest maps can be used as base maps for other tourist information.

3. Calculation and Visualization of Potential-of-Interest

To make a potential-of-interest map for a tourist area, we first need to know *visitor-interested locations* (the locations where previous visitors found something impressive) in this area. Here we consider the following five possible approaches for obtaining visitor-interested locations:

i. Annotation. This is a simple and inexpensive survey-based approach, in which tourists are asked after their sightseeing to point or encircle the locations where they have found something impressive.

ii. On-site manual recording. This is also a survey-based approach, in which tourists are asked to carry a GPS-equipped device and record their locations when they find something impressive. One possible method is to prepare an application for smartphones, which records the geo-coordinates when the user taps its display, analogous to giving applause. Another possible method is to apply existing micro-blogs, such as Twitter, in which the user can post a short message with location data.

iii. Use of photo geo-tags. In addition to GPS-equipped digital cameras, recent smartphones also allow users to record the locations where they take photos. When photos are taken by tourists with these devices, their photo-taking locations can be regarded as visitor-interested locations, because in many cases tourists take photos when they find something impressive there.

iv. Detection from GPS logs. During a walking tour, tourists typically reduce their speed when they find something impressive. Thus, by analyzing the trajectory of tourists recorded by GPS-equipped devices, we can detect visitor-interested locations with some degree of accuracy.

v. Text mining. In this approach, we extract the data of visitor-interested locations from a large amount of texts in travel blogs and reviewing sites. The extraction of such location data from text on web pages can be easily performed when the page shows geo-coordinates or mailing addresses (Kurashima *et al*. 2005).

In the next step, we compare these five approaches from the following seven viewpoints:

i. On-site workloads. It is not desirable to impose certain workloads on tourists during their sightseeing, because these workloads distract them and may affect their activities.

ii. Off-site workloads. In *annotation*, the tourists have to recall the locations that have impressed them and then find them on a map. This task is difficult and error-prone, especially for people with low map-reading skills.

iii. Accuracy. Spatial accuracy is a concern with both *annotation* and *text mining*, since these approaches rely on human memory.

iv. Possibility of commission errors. In *detection from GPS logs*, such locations as traffic intersections and toilets may be mis-detected as visitor-interested locations, unless a certain data refinement process using facility location data is applied.

v. Possibility of omission errors. In *use of photo geo-tags*, visitor-interested locations in photography-prohibited areas are never detected. Similarly, in *detection from GPS logs*, visitor-interested locations where tourists cannot stop, such as scenery spots on a highway without lookouts, cannot be detected. In *text mining*, visitor-interested locations without any reference to their addresses are difficult to detect. Finally, *on-site manual recording* is also risky because the subjects may forget to record their locations, especially when they are devoted to sightseeing.

vi. Technical challenge. At this moment there are no well-established techniques that realize detection of visitor-interested locations from tourists' GPS trajectories or their writings. Thus, *detection from GPS logs* or *text mining* are more technically challenging than others.

vii. Device necessity. Survey-based approaches are costly, especially when we have to ask the subjects to carry certain devices. In addition, the devices' weight may affect their activities. This, however, becomes a less important issue, because many tourists nowadays carry their own smartphones and they can be used for experiments.

Table 1 summarises the above discussion. Which approach we should take depends on which viewpoints we decide to emphasizes. We, however, concluded that *use of photo geo-tags* is the most promising approach, considering its smallest number of problems (note that *device necessity* seems not critical as stated above). In addition, it is also nice that *use of photo geo-tags* imposes less workload on tourists than *annotation* and *on-site map recording*, as photo-taking is a tourists' natural activity.

Table 1. Comparison of five possible approaches for obtaining visitor-interested locations from seven viewpoints

	On-site Workload	Off-site Workload	Accuracy	Possibility of commission errors	Possibility of omission errors	Technical challenge	Device necessity
Annotation		×	×				
On-site manual recording	×				×		×
Use of photo geo-tags					×		×
Detection from GPS logs				×	×	×	×
Text mining			×		×	×	

Now let us assume that we have already obtained a large dataset of visitor-interested locations in a tourist area by one of the above five approaches. The next mission is to calculate the potential-of-interest of each location in this area. For this calculation, we can use a technique called *kernel density estimation* (Parzen 1962). Roughly speaking, kernel density estimation is a computational operation in which we assign a small hill for each data point, called a *kernel*, and measure the total height of all hills on each location. As a result, a location with a larger number of points around it has a higher value. Kernel density estimation is, in the original sense, a technique to estimate the probability distribution of a stochastic process from a large set of samples. If we take

the *use of geo-tags* approach, for instance, tourists' photo-taking activities are considered a stochastic process that follows a certain two-dimensional probability distribution. This probability distribution should reflect the distribution of sightseeing potentials.

The estimated probability distribution can be represented by a continuous surface map. This surface map is sometimes called a *heat map* in analogy to a thermograph. Intuitively speaking, this surface map represents the density of sample points (in our case, visitor-interested locations) at each location. Thus, the surface map, constructed via kernel density estimation from a large dataset of visitor-interested locations, is considered a potential-of-interest map.

4. Demonstration

As a case study, a potential-of-interest map was made by the *use of photo geo-tags* approach in Section 3. Our target was Yokohama, a portside city near Tokyo. Yokohama's central area is a popular tourist destination, which attracts more than forty million tourists in a year.

In order to ease the labour of collecting data about photo-taking locations from tourists, we used Flickr photos for this study. Flickr (http://www.flickr.com/) is one of the most popular photo-sharing services on the web and is currently owned by Yahoo! USA. Flickr has many photos taken by tourists. In addition, Flickr has an open and powerful API with which we can perform geo-searching of Flickr photos in our own program. Therefore, Flickr is a suitable data source for collecting data about tourists' photo-taking locations. Indeed, there are already many researches which employ Flickr's photo data. For instance, Girardin et al. (2007) and Zheng et al. (2011) inferred tourists' movement patterns from spatio-temporal sequences of photo locations of individual photographers, while Hollenstein and Purves (2010) analyzed people's perception of areas in a city from the combination of location data and user-generated tags of Flickr photos.

In our study, we first developed a JavaScript program which automatically collects the data of Flickr photos that have been taken within a certain rage of a specified address. The data include the Photo ID, the location (geo-coordinates) where the photo has been taken, the owner's ID, and the owner's place of residence (country and city names). With this program, we collected the data of 1061 photos of Central Yokohama (the photos taken within 3km from Yokohama City Hall). We then removed 38 photos taken by local residents, who were probably not tourist photos. The remaining 1023 photos were taken by 95 people (on average 10.8 photos per person); 34 live in Japan, 20 in other countries, and the place of the remaining 41 people were unknown. Then, we imported the location data of these 1023 photos into ArcGIS 8.1 with Spatial Analyst and carried out kernel density estimation. The resulting map (Fig. 1) shows higher values at and around major tourist attractions in Central Yokohama, such as Minato-Mirai Waterfront Area, Akarenga Park, Ohsanbashi Wharf, and Yokohama Chinatown. This supports our intuition that the location data of Flickr photos, which include many tourist photos, can be employed as the data source for potential-of-interest maps.

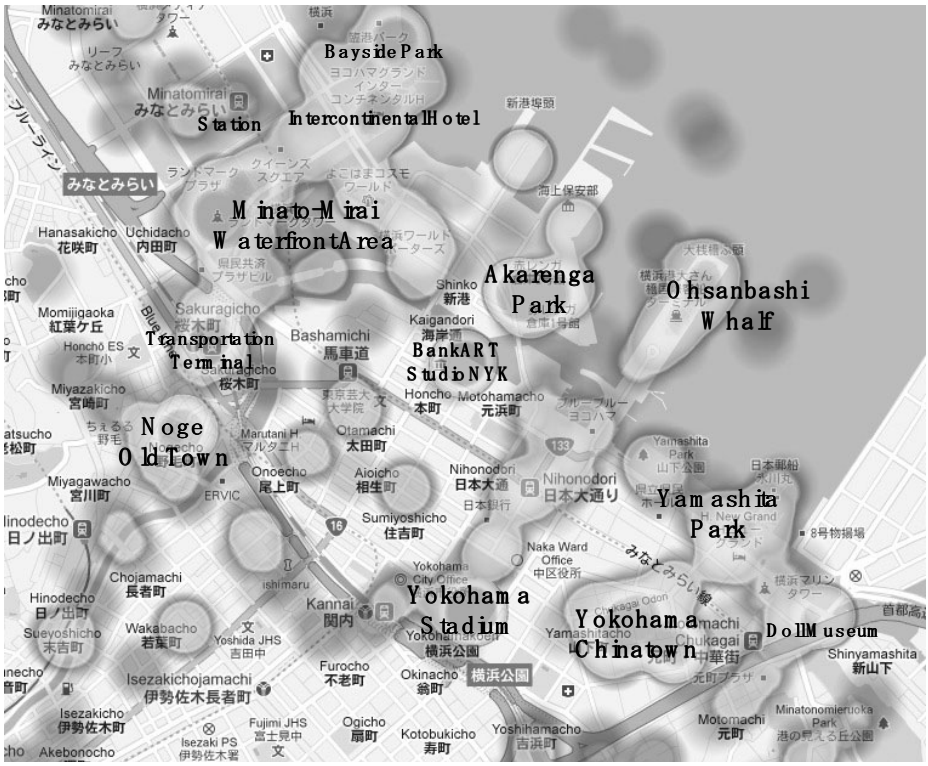


Fig. 1. A potential -of-interest map of Central Yokohama, made from photo-taking location data of Flickr photos

5. Potential-of-Interest Map-based Tourist Information Tool

Now that we have an example of a potential-of-interest map, we implemented a prototype of the mobile tourist information tool discussed in Section 2, which features a point-of-interest map. This tool first shows a Google map of Central Yokohama, which is enhanced with a layer showing the potential-of-interest map (Fig. 2-left). In this map, the darker areas represent the areas with higher potential-of-interest (i.e., the areas where many tourists have taken photos).

This map also contains blue markers, which represent the locations of major tourist attractions and train stations. In total 34 markers are prepared for this area. If you click a marker, a short description and a picture of the corresponding facility are shown in a balloon (Fig. 2-right). The description contains hyperlinks to the official site and Wikipedia page of this facility, if they exist.

The screen also shows a button named ‘Start Tracking.’ If this button is on, the map automatically scrolls such that the user’s current location comes to the centre of the map, together with a circle at the centre which shows possible measurement error of the current location.

Since this tool was developed with JavaScript and the Google Maps JavaScript API v3, it can be used on various platforms through web browsers. Its screen layout and user interface depends on the device, due to the specification of the Google Maps JavaScript API v3. For instance, on Android smartphones and iPhones, the user can scroll the map using one finger. In addition, on iPhones (but not Android smartphones at this moment) the user can change the map scale with two fingers.

The most remarkable feature of this prototype is its simplicity. The initial screen shows almost no text except labels. Nevertheless, the map is informative, as it visually tells the user which attractions are possibly worth visiting and which are not. If the user becomes interested in a certain area, he may browse the descriptions of tourist attractions in that area. These descriptions are valuable to judge quickly whether the attraction is interesting or not for him. If he needs more information for his decision, the user can obtain it from external sites via hyperlinks. We considered that tourist information tools should not be relying entirely on proprietary information, as it is difficult to keep updating the information up-to-date within individual tools.

Fig. 2. Screenshots of our mobile tourist information tool which features a potential-of-interest map. The balloon in the right screenshot introduces *Kishamichi Promenade*, saying "A trans-bay promenade which was built on a former freight railroad. Visit here if you want to enjoy a nice view of Minato-Mirai Waterfront Area over a bay".

6. User Evaluation

In order to evaluate the developed tool which featured a Flickr-based potential-of-interest map, we conducted a user test. In this test, we asked its participants to visit Central Yokohama and do sightseeing freely for at least two hours using our tool. After that, we asked them to answer an online questionnaire. We had twelve participants (eight males and four females; average age was 27.2). Nine participants used their own smartphones, while other three borrowed the Android smartphones we prepared.

Table 2 summarizes the major result of our user test. Unfortunately, the average scores of six evaluation items were not glorious. The major reason was that, as implied by the large standard deviation (S.D.) values, the participants were divided into sympathetic and anti-sympathetic groups. Particularly, the three participants who reported their poor map-reading skills were anti-sympathetic. For instance, the average score of Item 1 (ease-of-use) by these three participants was only 2.0. On the other hand, three participants who borrowed our Android smartphones gave a score of 3.4 for Item 1, even though they were not used to the smartphones.

Among the six evaluation items, the average score of Item 2 (new findings) was the lowest. This was probably because most participants have already visited Yokohama many times. Item 3 (validity of potential-of-interest map) was given low scores by the participants who visited Yokohama on rainy days. Our potential-of-interest map seems unsuitable for rainy days, because it was made from the geo-tags of tourist photos mostly taken on sunny days. Lastly, Item 5 (use intention in other destinations) and Item 6 (use intention at home) are given high scores by nine and seven of twelve participants, respectively. This makes us confident of the presence of consumer needs for our new tool.

Table 2. Test users' evaluations of our mobile tourist information tool featuring a potential-of-interest map (5: highest and 1: lowest, $n = 12$)

Question	Ave. Score	S.D.
1. Did you feel that it was easy to operate the system?	3.5	1.2
2. Did you find any new attractive spots while using the system?	3.1	1.1
3. Did you think that potential-of-interest is reliable?	3.3	1.1
4. Did you think that the information about each point-of-interest, shown in a balloon, was sufficient?	3.5	0.9
5. Do you want to use this tool on smartphones again when you visit other destinations?	3.7	1.2
6. Do you want to use this tool again on a PC when you plan a trip at home?	3.4	1.4

We also asked the participants to comment freely on our tool. Some participants gave positive comments on such features of our tools as the absence of unnecessary textual information, the presence of potential-of-interest data on unfamiliar districts, and easy access to the information around the user's current location. We also found complaints about slow data communication speed, inaccuracy of estimated current locations, and small displays, which were actually device-side problems (although there is still room for improvement in our side).

Several ideas for improvement were proposed by some participants. One nice idea was to prepare several potential-of-interest maps which correspond to different situations (e.g., day and night, sunny and rainy days, dating and solo trip) and to enable their switching or to represent them with different colours.

7. Conclusions and Future Work

It is difficult for tourists to process a large amount of text-base tourist information for finding which attractions, streets, and areas are worth visiting. This problem becomes more serious when the tourists are already at the destination and have to rely on their mobile device with a tiny screen. This paper, therefore, proposed the use of potential-of-interest maps in mobile devices as a visual aid for tourists. Potential-of-interest maps allow the tourists to understand the distribution of sightseeing values in a tourist area and accordingly, help the tourists to select the targets (attractions, streets, and areas) whose information should be looked carefully for their tour planning.

In this paper, we made a potential-of-interest map of Central Yokohama from the location data of Flickr photos. Geographic visualization of Flickr's photo-taking locations itself is not new, but the use of it as a source of tourist information seems a new idea. Zanker et al. (2010) also discussed the application of user-generated geo-tagged information on the Web to the tourism domain, although their focus was to use such information in extending the framework of information retrieval. On the other hand, we pursued a new presentation style of tourist information.

One remaining issue in the current approach is that we cannot assure that the photos we have used are taken by ordinary tourists. Thus, we are planning to revise our potential-of-interest map based on photo location data collected from actual tourists. In addition, in order to keep the potential-of-interest map up-to-date, we are planning to introduce a 'user participation' mechanism, in which photo location data are collected seamlessly from the users of our tool by building a camera function into our tool. A benefit of this mechanism is that the result of tourists' natural actions is utilized for guiding future tourists. Ultimately, we would like to evolve the tool into a dynamic system in which its users can see where other users are taking photos at each moment, because many season- and time-dependent attractions, such as festivals, flowers, and street performances, take place in a tourist area. Another merit of the user participation mechanism is that we can consider personalization by collaborative filtering (Resnick, 1994); that is, we can customize potential-of-interest maps for each user, based on the photo location data of other *similar* users who have taken photos at similar locations.

References

Chenoweth, R. (1984). Visitor-Employed Photography: A Potential Tool for Landscape Architecture. *Landscape Journal* 3(2): 136-143.

Girardin, F., Fiore, F. D., Blat, J. & Ratti, C. (2007). Understanding of tourist dynamics from explicitly disclosed location information. In G. Gartner, W. Cartwright and M. Peterson (Eds.), *4th International Symposium on Location Based Service and Telecartography*.

Kurashima, T., Tezuka, T. & Tanaka, K. (2005). Blog Map of Experiences: Extracting and Geographically Mapping Visitor Experiences from Urban Blogs. In A. Ngu, M. Kitsuregawa, E. Neuhold, J. Chung and Q. Sheng (Eds.), *6th International Conference on Web Information Systems Engineering*, Lecture Notes in Computer Science 3806.

Hollenstein, L. & Purves, R. (2010). Exploring Place through User-Generated Content: Using Flickr Tags to Describe City Cores. *Journal of Spatial Information Science* 1: 21-48.

Kurata, Y. (2011). CT-Planner2: More Flexible and Interactive Assistance for Day Tour Planning. In R. Law, M. Fuchs,and F. Ricci (Eds.), *ENTER 2011, Information and Communication Technologies in Tourism*.

Parzen E. (1962). On estimation of a probability density function and mode. *Annals of Mathematical Statistics* 33: 1065-1076.

Ricci, F., Arslan, B., Mirzadeh, N. & Venturini, A. (2002). ITR: A Case-Based Travel Advisory System. In S. Craw and A. Preece (Eds.), *6th European Conference on Advances in Case-Based Reasoning*, Lecture Notes in Computer Science 2416.

Iacovou, N., Suchak, M., Bergstorm, P. & Riedl, J. (1994). GroupLens: An Open Architecture for Collaborative Filtering of Netnews. In *Conference on Computer Supported Cooperative Work*.

Zanker, M., Jessenitschnig, M. & Fuchs, M. (2010). Automated Semantic Annotations of Tourism Resrources based on Geospatial Data. *Information Technology & Tourism* 11: 341-354.

Zheng, Y., Li, Y., Zha, Z. & Chua, T. (2011). Mining Travel Patterns from GPS-Tagged Photos. In K. Lee, W. Tsai, H. Liao, T. Chen, J. Hsieh, and C. Tseng (Eds.), *Advances in Multimedia Modeling*, Lecture Notes in Computer Science 6523.

Zipf, A. & Malaka, R. (2001). Developing Location Based Services for Tourism—The Service Providers View. In *ENTER 2001*.

Acknowledgements

We appreciate our students who have participated in our user test and gave us precious feedbacks. This work was supported by JST RISTEX's research program on Needs-oriented Service Science, Management and Engineering (title: Architecting service with customer participation based on the analysis of customer experience and design processes: sophisticating tour design processes as a case study) and JSPS Grant-in-Aid for Young Scientists (B) 23701030 (title: Construction and application of potential-of-interest maps making use of tourists' photo-taking location data).

Territoriality and Consumption Behaviour with Location-Based Media

Iis P. Tussyadiah

School of Tourism & Hospitality Management
Temple University, USA
iist@temple.edu

Abstract

The development in location-based mobile media has led to the popularity of its use for place experiences. This study explored the concept of territoriality, which is suggested as the underlying human behaviour that influences consumers' mobility and experience stimulated by the social gaming feature of location-based media. From an exploratory investigation with a series of focus group discussions with users of location-based media, this study observed the activities of territorial tagging for the purposes of territorial claim and defence to gain and maintain the perceived territorial control over resources and rewards attached to certain places. The ability of location-based media to make the physical territory to interact with informational devices enables territorial behaviour to manifest in the consumption of local establishments, making location-based media a powerful tool for marketers and managers to transform people–place experiences. Managerial implications are provided.

Keywords: territoriality; mobility; location-based media; mobile technology.

1 Introduction

Tourism and hospitality industries are witnessing the continuous development in mobile computing and location-based services and how the use of such technology influences the ways people experience places. Recently, location-based applications on smart phones, such as *Foursquare*, *Gowalla*, and *SCVNGR*, have emerged. These applications combine location-based services, social networking, and social gaming (i.e., playing games by means of social interactions) to encourage the consumption of places. Using such applications, people are expected to experience cities and destinations in a more fun, playful way, leading to a higher degree of mobility and a pervasive social influence.

Indeed, with more than 10 million users worldwide and three million *check-ins* daily (Foursquare, 2011), *Foursquare* has proven to be an attractive mobile application for tourism marketing. The City of Chicago and the State of Pennsylvania in the US are partnering with *Foursquare* to encourage visitors to uncover the history and culture of the cities and unlock special badges associated with their lifestyle (Van Grove, 2010a; 2010b) and more destination marketing organisations (DMOs) are following their footsteps. The growing interest in using location-based media for tourism and hospitality promotion indicates the importance of understanding how the technology transforms consumer behaviour in order to strategize such approach.

The social gaming feature of location-based media brings about the consequence of social competition through mobility. Since these applications offer rewards for accomplishments of specific tasks associated with the consumption of places, users are competing with their peers to achieve a certain status while experiencing cities

and destinations. Specifically, playing a social game with location-based media gives users the opportunity to lay claim to a specific place (e.g., by becoming a *Mayor*) ahead of others, indicating the basic behaviour of human territoriality. More importantly, the use of location-based media enables territoriality behaviour to be bound to consumption. Hence, it is posited in this study that the ability to use location-based media for territorial marking could be leveraged further for marketing and management strategies. While the role of social interactions in influencing individual decision making has been a topic of interest in tourism literature, the aspect of territoriality in the consumption of places is yet to be explored.

Therefore, the aim of this study is to investigate the intersection of human territoriality and technology within the context of tourism and everyday life. Particularly, the study explores the use of location-based media for territorial claim and defence behaviour in the process of territorial production. Further, this study provides managerial implications for tourism and hospitality industries.

2 Literature Review

2.1 Human Territoriality and Tourism

Human territoriality is defined as an act of laying claim to and defending a territory (Hall, 1959; Sack, 1986) to secure a set of spaces for performances of various activities (Rivano-Fischer, 1987), and to affect, influence, or control access, actions, and interactions (Kärrholm, 2007). This behaviour has been attributed to the creation of status and self-image as well as the concept of privacy and intimacy (Brown, 1987). Territoriality can be seen as the social production of space, as the creation and transformation of territory establish the framework of social life (Soja, 1989). Indeed, territorialisation establishes rules and relationships (Brown & Capdevila, 1999) and, in order for them to remain effective, territories need to be constantly produced and reproduced by way of control, socialized behaviours, etc. (Kärrholm, 2007).

Within the context of tourism, traveling to a tourist destination means entering the territory of (and governed by) others. Based on the French term *terroir*, referring to a district of certain geographical quality (Gottman, 1975), the term *touristic terroir* was coined to indicate the unique combination of attributes at a tourism region that defines the regional flavour of tourism experiences (Hall & Mitchell, 2002). *Touristic terroir* is an expression of regional identity that is unique and difficult to replicate, making each tourism destination a unique geographic territory characterised by the natural and cultural landscapes that are formed and reformed by its people as a result of social processes. Consequently, tourist experiences are bound by the rules and relationships established by the 'owners' (e.g., in the form of entrance permit, access to objects, people, and information, etc.). In tourism research, tourists' spatial behaviour has been linked to itinerary models (i.e., tourists' spatio-temporal mobility) and host−tourist relationships (i.e., socio-cultural interactions). Studies focusing on the basic territorial behaviour in tourism are rather limited.

The discussion on tourists' territorial behaviour in a destination has been focusing on territorial functioning in various tourism settings. Andereck (1997) identifies territorial behaviours among tourist groups at tourism attractions. Perceiving an

attraction as a public space, tourist groups hold it a territory, form an attachment to it, albeit in a short period of time, and exhibit negative responses to territorial invasion. However, Andereck (1997) also identified that the behaviour is limited in small territorial layer, such as marking and claiming a bench, rather than occupying the entire attraction, which is consistent with the characteristic of territoriality behaviour in public spaces. Here, territoriality is seen as the perceived "ownership" of a place at a particular time period, where a place is appropriated but not necessarily defended (Altman, 1975), making human territoriality different from that of aggressive-defensive zoological territoriality.

2.2 Technology and Territoriality

The discussion of territorial behaviour is further enriched by the development in information and communication technology (ICT), chiefly because ICT added another layer to the definition of space. As noted by Adams (1997), the vocabulary of space is broadly used to explain the virtual environment, such as *cybercafé*, *virtual office*, *chat room*, *blogosphere*, etc. ICT articulates space and place in different ways, giving rise to *hybrid ecologies* (Crabtree & Rodden, 2008) that support new forms of encounter and interaction based on social and virtual proximity and presence (Licoppe & Inada, 2008). Further, these space metaphors imply spatial behaviour over time, including exploration, settlement, and virtual habitation (Adams, 1997), allowing for the basic territorial claim and defence behaviour comparable to that of the physical nature. For example, people claim their territory on the *blogosphere* by maintaining a *weblog*, where they can control other internet users' access to information and develop relationships with certain groups of desired users.

Most importantly, the development in ubiquitous computing adds a new dimension to users' territorial behaviour. Lemos (2010) argues that mobile technology enables new means of mobility, communication and sociability, creating new ways of territorialisation, based on the convergence between physical space and cyberspace. With the new locative media, particularly location-based services on mobile phones, places and material objects are able to interact with informational devices (Lemos, 2010), giving ways for users (e.g., marketers, artists, tourists, etc.) to experience and attach new meanings to space.

Location-based services, for example, enable *mediated co-proximity* (Licoppe & Inada, 2009) (i.e., two users mutually recognize that they are close to each other), which can further stimulate face-to-face interaction and/or knowledge sharing and collaboration in cyberspace (e.g., by leaving online recommendations, playing social games, unlocking bonuses and deals, etc.). Hence, places, including tourism destinations, can be composed of several different territorial, physical and virtual, layers.

The concept of privacy and personal space explains the different territorial layers of places. According to Proxemics Theory suggested by Hall (1966), personal space can be divided into several distance zones: intimate, personal, social, and public zones. Each of these zones relates to one's preferred social, interpersonal distance. Additionally, in an attempt to conceptualize territorial privacy within the context of ubiquitous computing, Könings and Schaub (2011) define three categories of human territory: physical territory, extended territory, and private territory (See Fig. 1).

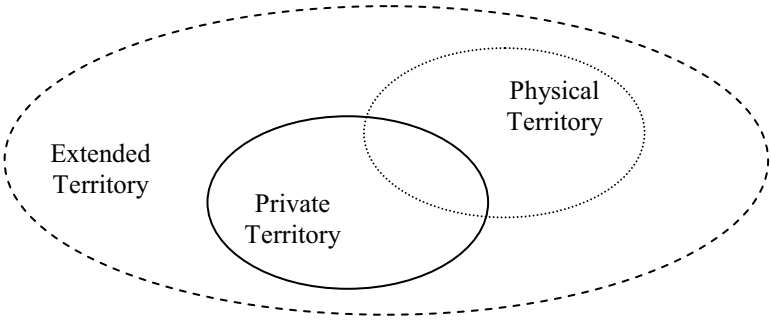


Fig. 1. Human Territory (source: adapted from Könings & Schaub, 2011)

Physical territory refers to the environment characterized by material objects and physical boundaries. Extended territory encompasses the physical territory as well as the remote entities connected via ICT. This suggests the cyberspace as meta-territorial domain whose online characteristics entangled with the physical properties. Lastly, private territory is a subset of extended territory, but not necessarily a superset of physical territory (Könings & Schaub, 2011).

The concept of private territory is central to this study as it relates to the claim and defence of personal and shared environment in one's social life. Although Hall's (1966) Proxemics Theory was initially conceptualized for physical space, it can be suggested that private territory is similar to or having the characteristics of intimate and personal zones.

2.3 Territorial Production

Although territoriality can be considered universal as human behaviour, the forms that it takes can be varied enormously (Delaney, 2005). "Territories are produced everywhere, in different ways, in different contexts, and by different means, and encompass a wide range of phenomena" (Kärrholm, 2007, p. 441). Bringing together the research on Actor−Network Theory and human territoriality, Kärrholm (2007) suggests four different forms of territorial production: territorial strategies, tactics, association, and appropriation (See Table 1).

Table 1. Forms of Territorial Production (Kärrholm, 2007)

	Impersonal Control	**Personal Control**
Intended Production	Territorial Strategy	Territorial Tactics
Production through Use	Territorial Association	Territorial Appropriation

Territorial strategies and tactics are intentional attempts to claim a territory. Territorial tactics are personal; they are directed explicitly toward the ordering of a

certain area. On the other hand, territorial strategies are impersonal, planned and mediated control. Territorial association and appropriation represent territorial productions that are not planned but are consequences of regular practices. Territorial appropriation is typically based on a repetitive and consistent use of an area by certain individuals and/or groups, while territorial association characterizes a place with a certain usage and specific conventions and regularities that underpin this usage (Kärrholm, 2007).

In a tourism setting, partitioning attractions in conjunction with sequences of a tour program to allow for a group of tourists to occupy certain areas for themselves within a period of time can be considered a territorial strategy. A tourist marking a bench by the hotel swimming pool with a towel is a form of territorial tactic. In a tourism destination, a certain park can be appropriated for tourists' use during the day and associated with homeless people after dark. Kärrholm (2007) further argues that the "different forms of territorial production often operate at the same place, mobilizing different sorts of artefacts, rules, and so forth" (ibid., 2007, p. 441). The convergence of material and informational space made possible by location-based media sets forth new possibilities of territorial production, which include ways of territorial production (i.e., how to mark and defend territories), combinations of different territorial layers to mark and defend, etc. A straightforward example of territorial marking in cyberspace is to purchase land and/or islands in a virtual reality environment (e.g., *SecondLife*).

Furthermore, Garner, Rashid, Coulton and Edwards (2006) presents how people use mobile devices and RFID technology as *digital spraycan*, making the technology a means to mark their environment by creating digital graffiti. This implies the form of *territorial tagging* (i.e., in the form of *geo-tagging*, a process of adding geographic information to metadata) as the new way of territorial marking. In fact, territorial tagging is a common practice in mobile social networking today as tourists are leaving their digital footprints everywhere by publishing retrievable *geo-tagged* information.

More recently, location-based services on mobile phones let users to *check-in* from places, allowing them to claim certain venues and access their benefits. To be able to *check-in* using location-based technology, people need to be physically at the venue, which most of the time also means consuming the venue (e.g., patronage to restaurants or attractions). In other words, consumption behaviour becomes a form of territorial production. Therefore, territorial behaviour using such technology has a significant impact for marketers and planners to influence and turn the technology users into consumers.

3 Exploratory Investigation

The ways location-based media allow territorial behaviour to manifest in consumption makes technology-assisted territorial production important to explore, especially for tourism and hospitality businesses. Not only that these forms of territorial production important in terms of deepening our understanding on the convergence of mobility in material space and the cyberspace, it is also important to derive implications of these behaviours for planning and marketing purposes. Therefore, the chief goal of this

study is to explore the use of location-based technology applications for various forms of territorial production to explain territorial behaviour in the intersections of physical space and cyberspace. Specifically, the focus of this study was the use of *Foursquare* applications on smart phones, which integrates the aspects of location literacy, social network, and competition (through social gaming), to establish experiential territories at home and at tourism destinations. An exploratory qualitative study was undertaken to examine territorial behaviour using location-based media.

Specifically, focus group discussions with location-based social network application users were conducted to gain valuable insights into the following inquiries:

(1) how users define and establish personal and experiential territories through the consumption of tourism and hospitality venues while using location-based technology, and

(2) the common forms of personal territorial production with location-based media.

Five moderated focus group discussions, averaging in six participants, were conducted from October 2010 to February 2011. A metropolitan area in the Eastern coast of US was chosen for the focus group discussions considering the fact that these applications were targeted for use in urban areas. Invitations to join the focus group discussions were posted on a *Facebook* group page that was left open to all viewers. Interested participants were screened and they identified themselves as avid users of *Foursquare*. Applications such as *Foursquare* and *Facebook Places* allow different types of businesses to register their venues, ranging from restaurants and bars to beauty salons and medical centres. For the purposes of this study, tourism and hospitality venues (e.g., restaurants, attractions, shops, movie theatres, etc.) were highly emphasized in the discussions. The scope of the discussion includes users' patronage to local venues at home and those visited at tourism destinations while traveling.

Several researchers suggest focus group discussions to be composed of homogeneous respondents (Bellenger, Berhardt, & Goldstucker, 1976) for a shared perspective to emerge. However, Calder (1977) also suggests that heterogeneous respondents may yield rich information for exploratory research. Therefore, considering the exploratory nature of this investigation, the first discussion was composed of heterogeneous respondents to obtain rich information from diverse user experiences. To further gain new ideas and confirm the collected information, homogenous respondents were allocated for the remaining discussions. Based on the characteristics of the respondents, two discussions were composed of only students and the other two of only working professionals, each in similar age groups (student groups were in their 20s and professional groups were in their 30s). The discussions were recorded into sound files (i.e., roughly five hour long) and later transcribed into textual data. All respondents received a $25 dining certificate upon completion of the discussion.

4 Result and Discussion

Territoriality behaviour was observed from the use of location-based applications (Table 2). It was identified that these applications presented the opportunities for social competition among users. Driven by the motivation to compete with others,

people strategize their mobility by patronizing different places to collect rewards offered by the application (e.g., points, badges, status) and by merchants (e.g., discounts, bonuses), as well as to gain recognition from other users. Indeed, it was identified that even though all participants did not perceive the use of such application as "playing games" *per se*, they agreed that using it makes their daily and touristic experiences more playful and fun.

4.1 Territorial Tagging through *Check-In*s

Check-in is the activity whereby users announce their physical location to the location-based mobile system and allow the system to make it visible to select friends/contacts. In return, the system will allow users to identify other people nearby for further interactions and to retrieve recommendations. In other words, location-based media allow physical proximity to be transformed into mediated co-proximity (Licoppe & Inada, 2009). This shows how the physical space and cyberspace converge with the assistance of location-based media, illustrating what Lemos (2010) suggests as the new mobility and sociability.

From the focus group discussions, it was identified that all participants *checked-in* from places for different purposes, ranging from social connection (e.g., to share their experience with friends, to screen the social environment at places, etc.) to social competition (see Excerpts a and b). The differences in purposes of check-ins, however, did not correspond to their age or occupation. One of the main purposes of *check-in* identified from the discussion was to achieve the rewards offered by the venues and the gaming feature in the system (see Excerpt c). The reward-seeking behaviour manifested in *check-in* activities, which also enforced social competition, requires users to apply different strategies of personal control, making territorial tagging a mechanism for territorial production.

4.2 *Mayorship*: Territorial Claim and Defence Behaviour

Foursquare Mayorships are awarded to users with the most *check-in*s (i.e., more days than anyone else) at a specific venue over the last 60 days. Besides gaining recognition through the mobile clients, users crowned as *Mayor*s are typically eligible for special rewards provided by the venue, such as free merchandises, discounts, or special arrangements for social recognition (e.g., bars displaying their *Mayors* on a digital jukebox monitor for all patrons to see). Hence, from the point of view of consumers, a *Mayorship* title is seen as an outcome of territorial production, in that becoming a *Mayor* brings in some forms of territorial control through appropriation. This leads to the social competition through territorial claim and defence behaviours.

Table 2. Excerpts: Location-based Technology and Territory

	Narratives	Respondents
a.	*"I am not from around here, my family and close friends are somewhere else. I have to show them all the exciting things that I do here (by* checking-in*)…"*	Male, Student
b.	*"…it's a way of being interactive with the world. To show all my friends that I am building a history… that I am a cool person,* you know*… and it becomes a competition too because it's rewarding."*	Female, Professional

cont. Table 2

c.	*"For me* [using the application] *is a way to pass the time. It's not necessarily a game… it's just something I do when I walk down the streets… but then you get some* stuff *back. Because I* checked-in*, it shows me there's a special over here… or if I* checked-in *ten times I can unlock a coupon somewhere…"*	Male, Professional
d.	*"I'm into competition… Once your friends are following, you got to compete… who got more points and who got more* Mayorships*."*	Female, Professional
e.	*"…it became more about competitions. That's when the* Mayor *thing starts coming into play. If I keep on checking-in, in five days I can become a* Mayor*, so I want to check-in five days in a row… especially in places like* [Bar Name] *where they show the* Mayor *on the jukebox."*	Male, Professional
f.	*"…I would go to regular places more often to become a* Mayor*."*	Female, Student
g.	*"*[When traveling] *It's used to be very easy to become a* Mayor *in* [pause] *remote areas, where not so many people use* Foursquare*... So, I would make sure to drive further from the city to* check-in *at random places and try to become a* Mayor*."*	Female, Professional
h.	*"I'm the* Mayor *of a couple of places around where I work, mostly coffee places, lunch places... if I'm the* Mayor *and I haven't been there for a while I definitely want to go there just because I don't want to lose my* Mayorship*. So I make sure to go there to check-in and maintain my* Mayorship*."*	Male, Professional
i.	*"If I'm ousted as a* Mayor*... I would be very angry!* [Laugh]*. The challenge is not over… I would try to get the* Mayorship *back. I would check-in* like *three times a day!"*	Female, Professional
j.	*"*[Using the application] *makes me think about different things to do in the city… where to go… what yet to be discovered. It forces me to check out other places I've never been to… check out things larger than your own places."*	Female, Professional

From the social competition context, it was identified from the discussions that users were competing with others not only to get the *Mayorship* title at a specific venue, but also to get the most *Mayorships* (i.e., expansion of territories claimed). Hence, territorial tagging through *check-in*s can be seen as a territorial tactic (i.e., intended production) that people employ within a social network. From the discussions, most participants indicated that they visited places more often since they used *Foursquare* to become a *Mayor* and enjoy the rewards (see Excerpts e and f). Some participants would travel to places in remote areas where there were not many *Foursquare* users so they could get the *Mayorship* title easily (see Excerpt g), which indicates an expansion of territory. The system requirement for *Mayorship* also dictates the possibility of users being ousted as a *Mayor* due to an absence of *check-in*s; other users who frequent the venues might take over the title. In other words, *Mayors* are prone to territorial invasion. This indicates the consequence of territorial defence by continuing the consumption of venues after the achievement of a *Mayorship* title.

From the discussions, territorial defence behaviour was identified from participants' perceived necessity to maintain their status and the perceived threat from invasion by others (see Excerpts h and i). Therefore, the territorial production identified from the

discussions mostly encapsulates what Kärrholm (2007) suggests as territorial tactics through appropriation, in that most participants, especially those emphasized the importance of social competition, visited venues with the intent to become the *Mayor*. The use of these location-based mobile media encourages territoriality behaviours where places are consumed and appropriated for various social benefits.

4.3 Mobility and Experience Territory

The important consequence of territorial behaviours mediated by location-based mobile applications is the creation of mobility patterns and the establishment of private experience territories. The social competition challenges users to strategize their mobility within the city or in tourism destinations. Being a *Mayor* at a specific venue requires territorial appropriation, and users have to be physically at the venue to do so. In other words, territorial tactics through appropriation necessitate mobility. As a result, territorial behaviour causes users' patterns of mobility to develop. Some users visit regular places such as restaurants and cafés in the city more frequently (see Excerpts f. and h.), others travel to different places (see Excerpts g. and j.) to mark, appropriate, and communicate their territory. Either way, territorial behaviour (and mobility) manifests in the consumption of local establishments, making location-based media a powerful tools for marketers and managers alike to transform people–place experiences, particularly in tourism and hospitality industries.

Further, the mobility and consumption of places while using location-based media enable users to set the boundaries of their experience territory, both for everyday experience and touristic experience. For example, by *checking-in* to cafés and lunch places surrounding their workplace, people mark a network of venues as their everyday experience territory. While traveling, a tourist *checks-in* from different attractions and venues in a tourism destination to convey their obtained privilege of "claiming" another *terroir* for a period of time. This implies a process of privatization of experiential territory by personalizing and appropriating the 'public' space (Goffman, 1963; Licoppe & Inada, 2008).

From the discussions, the fact that others would be aware of their locations and patterns of consumptions (e.g., where they have lunch, where they went to on vacation, etc.) did not deter most participants to share their *check-in*s through the pervasive, location-based mobile media. For them, the establishment of personal experience territories is seen as a way of conveying their self-image to the world. Consequently, assisted by mobile media, a city (and a tourism destination) is comprised of networks of places of which many people perceive as their own experience territories and for which many people compete to gain territorial control.

5 Conclusion and Implication

This study observed consumer behaviour as it is influenced by the basic human territoriality behaviour mediated by the social gaming feature of location-based technology. Location-based media allow people to express their territorial behaviour through the consumption of places with an intention of gaining control over the resources offered by these places, albeit being different from claiming ownership to the places. In the world of location-based social gaming, the media assist users with

digital territorial tagging (e.g., *check-in*s), which can be employed to claim a territory and defend it from territorial invasion by way of consumption. Since territorial tagging activity requires users' exact geographic locations to be reported to the system, these applications have the potential to influence and change people's mobility within cities and tourism destinations. Consequently, perusing such applications for tourism marketing, by integrating the persuasive power of perceived territorial control, rewards, and social recognition, is considered beneficial.

For marketers, conveying the notion of territorial control with the achievement of a status after completing certain tasks will result in consumers' motivation to compete with each other to gain the perceived control. Marketers targeting consumers' variety behaviour (i.e., making consumers explore and visit more places in the area) should pay more attention to facilitate the activities of territorial claim.

On the other hand, marketers targeting loyalty behaviour (i.e., increasing the frequency of visits) should emphasize the importance of territorial defence. For example, destination marketers wanting their visitors to explore their area should create a mechanism of rewarding some forms of territorial control (e.g., privilege to certain information, discounts, etc.) after the achievement of certain number of *check-in*s within a period of time. This requires collaborations and partnerships among different venues sharing similar goals and themes within the tourism destination. Meanwhile, hospitality business owners targeting loyal customers should pay more attention on rewarding their *Mayor* to keep the social competition alive for both territorial claim and defence behaviour.

This exploratory study offers a deeper understanding of human territoriality behaviour manifested in the consumption of places with the use of location-based media. The results illustrate the convergence of physical and cyberspace and how people set the territorial boundaries of their experience territory by appropriating and personalizing areas within cities and tourism destinations.

Further, this study contributes to tourism practitioners by providing the implications to capitalize the use of location-based media for marketing and promotion. Further research can be extended to assess the role of territoriality behaviour in the success of location-based marketing by developing and testing a measurement model, which can be generalized for different consumption situations.

References

Adams, P. (1997). Cyberspace and virtual places. *Geographical Review* 87(2):155-171.

Altman, I. (1975). *The environment and social behavior.* Monterey, CA: Brooks/Cole.

Andereck, K. (1997). Territorial functioning in a tourism setting. *Annals of Tourism Research* 24(3): 706-720.

Bellenger, D. N., Bernhardt, K. L. & Goldstucker, J. L. (1976). *Qualitative Research in Marketing*. Chicago: American Marketing Association.

Brown, B. B. (1987). Territoriality. In D. Stokols & I. Altman (Eds.), *Handbook of Environmental Psychology* (pp. 505-531). New York: John Wiley.

Brown, S. D. & Capdevila, R. (1999). Perpetuum mobile: Substance, force and the sociology of translation. In J. Law & J. Hassard (Eds.), *Actor-Network Theory and After* (pp. 26-50). Oxford, UK: Blackwell.

Calder, B.J. (1977). Focus groups and the nature of qualitative marketing research. *Journal of Marketing Research* 14: 353-364.

Crabtree, A. & Rodden, T. (2008). Hybrid ecologies: Understanding interaction in emerging digital-physical environments. *Personal and Ubiquitous Computing* 12: 481-493.

Delaney, D. (2005). *Territory: A Short Introduction*. Oxford, UK: Blackwell.

Foursquare (2011). *About Foursquare*. Retrieved July 19, 2011 from https://foursquare.com/about.

Garner, P., Rashid, O., Coulton, P. & Edwards, R. (2006). The mobile phone as a digital SprayCan. In *ACE 06*, Hollywood, California, USA, June 14-16, 2006.

Goffman, E. (1963). *Behavior in Public Places*. New York: The Free Press.

Gottman, J. (1975). The evolution of the concept of territory. *Social Science Information* 14(3-4): 29-47.

Hall, E. T. (1959). *The Silent Language*. New York: Doubleday.

Hall, E. T. (1966). *The Hidden Dimension*. New York: Anchor Books.

Hall, C.M. & Mitchell, R. (2002). The touristic terroir of New Zealand wine: The importance of region in the wine tourism experience. In A. Montanari (Ed.), *Food and Environment: Geographies of Taste* (pp. 69–91). Rome: Societa Geografica Italiana.

Kärrholm, M. (2007). The materiality of territorial production: A conceptual discussion of territoriality, materiality and the everyday life of public space. *Space and Culture 10*: 437-453.

Könings, B. & Schaub, F. (2011). Territorial privacy in ubiquitous computing. In *The 8th International Conference on Wireless On-demand Network Systems and Services (WONS 2011)*, Bardoneccia, January 2011.

Lemos, A. (2010). Post—mass media functions, locative media, and informational territories: New ways of thinking about territory, place, and mobility in contemporary society. *Space and Culture* 13(4): 403-420.

Licoppe C. & Inada Y. (2008). Geolocalized technologies, location-aware communities and personal territories: The Mogi case. *Journal of Urban Technology* 15(3): 5-24.

Rivano-Fischer, M. (1987). Human territoriality: Notes on its definition, classification systems and micro territorial behavior. *Psychological Research Bulletin* 27: 1–18.

Sack, R. D. (1986). Human territoriality: A theory. *Annals of the Association of American Geographers* 73(1): 55-74

Soja, E. W. (1989). *Postmodern Geographies: The Reassertion of Space in Critical Social Theory*. London – New York: Verso.

Van Grove, J. (2010a). *How 5 Brands are Mastering the Game of Foursquare*. Retrieved October 10, 2010 from http://mashable.com/2010/04/02/foursquare-brands/.

Van Grove, J. (2010b). *Pennsylvania Partners with Foursquare to Inspire State Tourism*. Retrieved October 10, 2010 from http://mashable.com/2010/05/26/pennsylvania-foursquare/.

Evaluation of Mobile Augmented Reality Applications for Tourism Destinations

María Teresa Linaza[a], David Marimón[b], Paula Carrasco[c], Roberto Álvarez[d], Javier Montesa[e], Salvador Ramón Aguilar[f], and Gorka Diez [g]

[a] Department of eTourism and Cultural Heritage
Vicomtech-IK4, Spain
mtlinaza@vicomtech.org

[b] Telefonica I+D, Spain
marimon@tid.es

[c] Universidad Politécnica de Valencia, I3BH/LabHuman, Spain
pcarrasco@labhuman.i3bh.es

[d] Invitation Digital, United Kingdom
r.alvarezceballos@gmail.com

[e] Brainstorm Multimedia, Spain
jmontesa@brainstorm.es

[f] Indra Software Labs, Spain
sraguilar@indra.es

[f] Fomento de San Sebastian, Spain
gorka_diez@donostia.org

Abstract

Every city contains interesting places and stories to be discovered. Mobile Augmented Reality provides the means to enrich tourists through precise and tailored information about the surroundings of the area they are visiting. MobiAR is an AR platform based on Android, which assists users who need tourist information about a city. When users observe reality through the MobiAR application via their mobile devices, they can experience events that took place at their location through multimedia content, and can access useful information to plan their routes in the city. This paper describes the MobiAR platform and presents the evaluation process that has been applied to the MobiAR application, in order to gather the opinion of real users.

Keywords: Mobile Augmented Reality; evaluation; usability analysis; tourism destination.

1. Introduction

Mobile Augmented Reality (AR) has recently become very popular thanks to the convergence of SmartPhones, faster networks and cloud computing. Consumer-based mobile AR application development has grown quickly over the past years. In March 2011, there were nearly 650 applications with descriptions listing AR as a feature in Apple's iTunes store, including city guides and sightseeing applications for the tourism sector (itunes.apple.com [March 11, 2011]). This sector presents a compelling scenario for such technologies, as tourists usually require comprehensive, precise and updated information while discovering a destination.

The mixture of reality and properly positioned multimedia helps tourists to make the most of the destination and to better understand its tourist resources. The majority of published research in the field of AR has focused on the development of new technologies and improved application prototypes. Only few studies have analysed the experience of the user and have evaluated mobile applications, although such knowledge should benefit developers to move forward from the initial surprising factor. If mobile AR is to become a viable medium for delivering visual information, empirical research must be conducted to discover the perception and usage of commercially available applications. For example, the mobile AR industry must address the monetary and technical limitations of existing devices if this technology should achieve any significant market penetration.

Thus, this paper presents the results obtained during the user evaluation process of MobiAR, a mobile AR guide for urban tourism. The main goals pursued by the user validation were twofold: assessing whether users find the developed application suitable or not, and gathering suggestions about how to improve the application. Thus, aspects such as the usability of the MobiAR application, its acceptance in the market, as well as the suitability of the content provided by the application have been analysed.

This paper is organised as follows. Section 2 briefly describes the current status of mobile AR platforms, with special attention to the application in the tourist sector. Section 3 presents the MobiAR platform from a technical point of view, whereas Sections 4 and 5 focus on the description of the evaluation of the MobiAR application with the pilot users. Both the methodology that was followed and the outcomes of the process are explained. The article is concluded by a short discussion.

2. State of the art

2.1. Defining Mobile Augmented Reality

Augmented Reality (AR) is a relatively recent computer science field considered as a subfield of the broader concept of Mixed Reality (MR) (Milgram et al., 1994). Unlike Virtual Reality, AR attempts to enhance the personal environment of the user, rather than replace it. Although many researchers have broadened the definition and scope of AR, Azuma (1997) states that AR systems share the following three characteristics: combination of real and virtual objects in the real world; interactivity and interaction in real-time; and registration and alignment of virtual and physical objects with each other in 3D.

From both definitions (Milgram et al., 1994; Azuma, 1997), the main components of a typical mobile AR system can be induced. Wearable or portable input (camera) and interaction devices to interact with the augmented world are needed, as well as displays in order to incorporate the virtual data in the physical world; data storage and access, and a computational platform for the coordination of the tracking and the 3D registration of the real scene.

As mobile devices and services have proliferated during the last years, several groups of researchers around the world are creating mobile AR applications that are low cost and delivered to self-owned mobile devices of the users, such as cell phones,

SmartPhones or PDA. Typically, the user points the device in the direction of an item of interest and the camera output augments the display with additional information about the environment.

Several commercial applications have appeared in the market and reached users massively, namely Wikitude, Layar, Sekai Camera, Acrossair, and Junaio. Most of these applications for mobile AR have a similar user experience. Users access a list of different layers of information usually structured by content type or by proximity to the user. By choosing one of those content layers, users can browse their surrounding area by pointing with the device. Some of these applications also enable users to contribute with their content.

2.2. Existing Mobile AR Applications in DMO

AR technology has been applied frequently to the field of tourism, as it is well suited to assist both tourists who seek to explore a destination during their stay, and providers of goods and services (merchants, transportation and cultural heritage sites) which have information to share, promotions to announce and advice to offer.

Regarding worldwide mobile AR applications for tourism destinations, one of the pioneers was the region of Tuscany that launched Tuscany+ (http://www.turismo.intoscana.it/allthingstuscany/aroundtuscany/tuscany-the-first-augmented-reality-tourism-application/ [Oct. 11, 2011]). The application offers an interactive, real-time guide in order to enhance the trip with four categories of information (sightseeing, accommodation, dining, and entertainment). Nowadays, in Europe, cities like Rotterdam and Amsterdam in the Netherlands give the opportunity to tourists to discover city areas as they were in the past and as they will be in the future. Other cities such as Dublin are planning to use AR application for promoting tourism.

Asia is also very active in the development of innovative AR applications. For example, the Korea Tourism Organization has launched tourism services using AR technology in close collaboration with mobile service operators. Tourists can access real-time information about area attractions using their SmartPhones. Moreover, the Hong Kong Tourism Board (HKTB) in collaboration with Cathay Pacific Airways have developed the "DiscoverHongKong AR" application (http://www.discoverhongkong.com/eng/jsp/mobile-app/iphone.jsp [Oct. 11, 2011]), which features over 100 attractions, 5,000 retail outlets and 2,000 dining establishments, as well as the shopping malls and open-air markets all over Hong Kong.

Emerging destinations such as Dubai have launched a new iPhone application for visitors to promote the tourist attractions of the Emirate as part of its new IT strategy (http://itunes.apple.com/us/app/dubai-travel-guide/id354106773?mt=8 [Oct. 11, 2011]). The new AR App includes maps, deals and restaurants offering a particular cuisine or attractions in the city.

Finally, although it cannot be considered as a DMO, TripAdvisor (http://tripadvisor.wordpress.com/2011/02/17/tripadvisor-adds-virtual-tours-to-its-ipad-app/ [Oct. 11, 2011]) has launched a new functionality for travellers to take a virtual walk through their travel destination with nearby places superimposed over

Street View from Google, including TripAdvisor's traveller reviews and opinions, to discover great hotels, restaurants and attractions.

2.3. Current Practices in AR Evaluation

Although AR technologies have been developed during the last three decades, research about evaluation and Human Computer Interaction (HCI) issues in AR has just started (Dunser et al., 2008; Swan & Gabbard, 2005; Dunser et al., 2007; Grasset et al., 2007b). Standard design guidelines and interaction metaphors have scarcely been established due to the variety of devices and interfaces used for AR applications (Anastassova et al., 2007a; Anastassova et al., 2007b; Gabbard & Swan, 2008; Haller et al., 2007).

Anastassova et al. (2007a) examined 48 articles related with user-centred design and evaluation in AR industrial applications. Although most of the articles treated usability issues, only 17% of them presented a deep analysis of the needs of the user. The number of participants in the examined experimentations ranged from 1 to 75, with an average of 15. Another important remark is that usually experimental tasks are simple and of short duration, with only a 18% of the studies that have tried to place the potential users in real conditions (Damala, 2009).

The need for the implementation of user-based analysis regarding the creation of AR applications and interactive experiences has also been underlined by Gabbard and Swan (2008). They have proposed a scheme for the classification of user-based experimentations related with AR (Damala, 2009).

Finally, Dunser et al. (2008) have also proposed a evaluation taxonomy with five categories which has been partially followed within the MobiAR project: objective measurements (task completion time, accuracy, error rates); subjective measurements gathered by the use of questionnaires; qualitative analysis including formal user observation, formal interviews and coding of the behaviour of the user; usability evaluation techniques based on task analysis; and informal evaluations including observation.

To conclude, it must be mentioned that although the literature related with AR applications and experiences evaluation is not very wide, the research community seems to become more aware of the need for evaluation and user-based studies. When evaluation is carried out, it is often done at an advanced stage, in the laboratory and for short duration tasks. Thus, the work presented in this paper tries to contribute to the existing literature with a real scenario case.

3. The MobiAR Project

The goal of the MobiAR project is to develop a mobile service platform for tourist information based on mobile AR technologies (Marimon et al., 2010). MobiAR targets users who are willing to discover or better know a destination with enhanced experiences by enriching their visit with multimedia and location-based information. Users can be either tourists who need a guided tour, or local citizens eager to discover more about emblematic places in a destination. In both cases, users can generate their comments and reviews, so as to share them with others.

The main use cases covered by the MobiAR platform include the user registration; the configuration of the user profile; the visualisation of the city map that illustrates the Points of Interest (PoIs); the access to the description and the browse of multimedia content attached to particular PoIs; the sharing of pictures taken by the users to the system; and the update of the comments about a particular PoI.

The client side consists of a mobile application that orchestrates the interaction of the user with the content. The application determines the location of the user through the GPS antenna of the mobile phone, and reads the data provided by the digital compass and the accelerometers to dynamically present the geo-located AR information about the PoIs on the screen. This AR view can be seen using either the standard or the 3D view (Fig. 1 (a)), where PoIs are identified by their 3D models.

(a)

(b)

Fig. 1. MobiAR prototype. (a) 3D models of the PoIs; (b) Multimedia content associated with each PoI.

The use of 3D information enhances the tourist experience while discovering specific buildings in a tourist urban destination. This mode is enabled by the 3D visualisation module of the application, which is able to capture the motion of the real world and display 3D objects in real-time. This module has been implemented as a multipurpose 3D engine based on the OpenGL ES library.

Textual descriptions and multimedia content (images, audiovisual files, 3D models) about the PoIs provide the real value to users in a tourist context (Fig. 1 (b)). The content and user management server manages all the information related to a tourist destination, as well as the information about the users that allows the system to offer the right data at the appropriate moment. The server, published as a Web service, provides personalised information about the tourist resources taking into account both the position of the user and the set of user preferences determined by the language of the content, the radius, and the number and type of resources to be shown. Moreover, the server receives and stores all the content generated by the users (such as pictures taken through the MobiAR application and textual comments about PoIs).

Even though the mobile device initially calculates the user position, this information is corrected by the platform. The content and user management server stores a query image captured by the client, and makes a HTTP request to the image-assisted global positioning module, which corrects the GPS data based on visual matching of geo-referenced images. The visual recognition engine is the core technology that allows the identification of all similar images and their spatial relationship with respect to the image captured by the mobile device (Marimon et al., 2010). The matching is performed in two stages: (1) initial ranking of reference images using visual words of keypoint descriptors and voting in the limited pose space, and (2) re-ranking of the initial results using more complex spatial verification stage. The most probable location of the user is calculated combining data acquired from the device (namely GPS and orientation information) with the spatial relation calculated between the reference images and the query.

The database of reference images has been populated from two sources: Panoramio and Google Street View. During the assessment tests, the former has demonstrated that manual geo-tagging by users is not reliable enough for the targeted enhancement. For the latter, the precision is higher and hence tests showed the viability of the proposed solution. However, since the places where the user evaluation was conducted where mainly pedestrian and no images from Street View were available, qualitative studies on user experience could not be conducted and deferred for future work.

4. Evaluation Methodology and Experimental Setup

4.1. Description of the Methodology

As remarked by Damala (2006), three main types of evaluation can be distinguished: front-end evaluation that occurs during the very early stages of a multimedia project; formative evaluation during the development and production of a project; and summative evaluation that follows the completion of a project. MobiAR has conducted a formative evaluation in order to improve the development of the project.

Regarding the evaluation methods, two main approaches can be distinguished: qualitative and quantitative methods. Each method has different strengths and weaknesses which have been taken into account before choosing the methods best suited for MobiAR. In this case, questionnaires for quantitative, and interviews and observations as qualitative evaluation methods have been selected. A taxonomy of evaluation points candidate to be assessed has been established. First, technology-related points are strongly linked with choices made regarding the software and hardware used to design, implement and deliver multimedia applications in MobiAR. Under this category, the issues classified are the usability and ease of use of the application, and the overall user interface, including navigation and interaction design. Evaluation points related to information are focused on content quality, structure and design, usability and user satisfaction. Finally, regarding the tourist professionals, evaluation points are related to commercial exploitation.

4.2. Validation Scenario

The city of San Sebastian has been the validation scenario of the MobiAR application during the last week of November 2010. The field trials were carried out in the surroundings of the Town Hall and the Alameda Boulevard, making use of the free Wi-Fi network supplied by Fomento de San Sebastian. A set of 25 PoIs was selected, including accommodation, restaurants, places or buildings of interest, and monuments. It is worthwhile mentioning that the experience can be scaled to other PoIs of the city by introducing new content.

4.3. Recruiting the Candidates

Tourist offices of San Sebastian and Bilbao, the School of Tourism of the University of Deusto, companies involved in the development of Information and Communication Technologies (ICT) and other representative local institutions have been contacted. The system was tested and evaluated by 15 participants from both sexes (9 men and 6 women). Regarding their age, 7 participants of the sample were users between 36 and 60, 6 were users between 19 and 35, and the remaining 2 were under 18. The vast majority was familiar with computers, Internet, email or traditional mobile phones, but not with SmartPhones and mobile games. It must be mentioned that the ability of participants with ICT was not a selection criteria for the sample.

4.4. Description of the Experience

Before the beginning of each experimental session, participants were welcomed and introduced to the research team who explained the goals and the context of the project to the users. In addition, a short introduction about Android devices was provided to make the experience easier, especially for participants who did not have any experience of using such devices. Given their availability, users tested the MobiAR application individually or in pairs. Once located at the starting point in the real scenario, users interacted freely with the MobiAR application during nearly one hour.

After the experience, users received a complementary explanation of the features of the application that were not tried out. The last step consisted of completing a questionnaire, which was defined in order to gather the opinion that users had about the prototype and an informal interview with the observers in order to gather further information about their experience.

5. Results and Discussion

5.1. Questionnaires

Data has been collected on the basis of one specific questionnaire developed for the usability study of MobiAR, including aspects related to different constructors such as the perceived usefulness of the application, its perceived ease of use or the perceived added-value. The main difficulty has been related to the design of the questions which could bias the answers.

Usability of the AR Guide

This section, composed of four "closed" questions, was one of the most crucial regarding the main AR aspects of the prototype tested. More particularly, the statements included in this section tried to examine the overall easiness of use of the guide, the easiness of identification of the commented works and the easiness of navigation through the content of the guide.

The first statement was: "Most usable visualization option for PoIs" and found that 3 out of 15 participants (20%) to "3D" and 12 (80%) to "Normal". The following statement, related with the interaction with 3D icons of each PoI, was: "Have you been able to interact with 3D icons of each PoI?", with 1 participant (7%) that could not interact at all, 2 (13%) that found it quite difficult to interact, 8 (53%) for whom it was quite easy and 4 (27%) that interacted perfectly.

The statement "Have you been able to interact with the PoIs and 3D representations?" found 3 participants (20%) that had difficulties interacting, 8 (53%) that interacted easily and 4 (27%) that found it very simple. Finally, participants were also asked to position themselves on the statement "The PoIs were correctly placed in the AR view". The answers were nearly equally distributed between perfectly located (47%) and quite near to their real location (53%).

Measuring the Content Effectiveness of the Guide

This part of the survey was formed by questions that examined the content effectiveness. The first statement was: "Was the content interesting?" and found that it was "Quite interesting" for one third of the participants (5) and "Very interesting" for the remaining two thirds (10). Further possible answers were "Not interesting at all" and "Scarcely interesting".

The question "Which type of PoI did you find more attractive?" found that monuments and buildings (5.71 over 6) were the most attractive ones, followed by tourist offices (5.35 over 6), accommodation and restaurants (4.9 over 6), and pharmacies (4.71 over 6).

Questions about an Hypothetical, Future Use of the Guide

Lastly, a set of questions regarding a hypothetical future intention of using such a guide was included in the survey. The first question was "Would you use such a guide in other cities?". Participants could choose between "yes" or "no". Only one participant would not use the guide in the future. Another subject on which we wanted to have an estimation was the price participants would pay for such a guide. 8

participants agreed on a price between one and five euro, 2 (13.3%) less than one, 2 (13.3%) between six and ten euro, another 2 (13.3%) more than ten euro, and only one participant (6.6%) would only use it if it was free.

5.2. Semi-Structured Interviews

The interviews conducted with the participants straight after the visit, provided interesting feedback regarding the use of the guide. They were asked in an informal environment, as a conversation between the users and the observers.

On the Usability of the Application

In general, users found the tested mobile phone quite easy to use, and they browsed through the application with almost no problems. However, the lack of experience in using the Android operating system prevented people from discovering different functionalities that were only accessible through the menu button. In addition, the back button of the Android device was often pressed by accident due to its sensitivity.

Several users agreed on the fact that sometimes it was difficult to interact with the AR view icons because of their oscillation. The problem affected mainly to users with less experience with Android. Furthermore, some comments showed that the display of 3D illustrations was not clear enough, because users did not discover the 3D view button. Despite those problems, the evaluation of the usability was very satisfactory. Comments provided by users during the evaluation stated that the application was very intuitive, the selected icons for tourist resources were very suitable and the application had an easy-to-use interface. It is worth mentioning that almost all the users found the map view more functional than the AR view during navigation.

On the Content and Services Provided

The comments gathered about the contents have been very positive, especially those related to additional services and the possibility of including more user-generated multimedia content. However, the way in which the option for uploading photographs was shown - it only appeared after selecting a particular PoI and visualizing all the contents attached to it - seemed to be not intuitive enough for some users, because some did not discover the button for this purpose. Finally, it was noticed that users gave high importance to the quality and quantity of the multimedia content available at each PoI. For example, one of the users said that period pictures helped to discover unknown aspects of resources.

On New Functionalities

Apart from the comments about particular features of the application, users made some suggestions for improvements and new possible features of the MobiAR application (mainly regarding the usability of the application). One of the most frequent suggestion was to include a route navigation service in the application. Once the user selects a PoI, the system would show how to reach it, either through an audio-guide system or visual signs on the map.

Regarding the content, several users would like to include other type of tourist resources like parking and more information about the existing PoIs (i.e. opening hours, or address). Moreover, users considered the addition of audio files (besides

pictures) very interesting. Finally, some users suggested the possibility of storing all the visualized contents in the mobile device in order to be able to send them by e-mail. Some users suggested the inclusion of a help menu in order to give explanations about how the application should be used, by means of either textual or audio descriptions. Other users proposed changing the colours of some of the icons and screens, so that menus become more intuitive and accessible.

Finally, it is worth pointing out that content providers considered that managing their own PoI would help them to keep contents under control. Having such a management application would allow the owners of tourist resources to register and update the information about their own PoI. This tool has already been developed within the MobiAR platform, but it was out of the scope of the user validation. Thus, some users suggested the possibility of validating this tool.

6. Conclusions

Current research in the field of mobile AR focuses mainly on the development of improved tools, while only few researchers have looked at how users perceive this technology. Such knowledge will benefit developers to move forward from the initial surprising factor. Thus, this paper presents the results obtained during the user evaluation process of MobiAR, a mobile AR guide for urban tourism. The goal of the MobiAR project is to develop a mobile service platform for tourist information based on mobile AR technologies.

The main goals pursued by the user validation that took place in San Sebastian were twofold: assessing whether the users find the developed application suitable or not, and gathering suggestions about how to improve the application. Thus, aspects such as the usability of the MobiAR application, its acceptance in the market, as well as the suitability of the content provided by the application were analysed.

According to the results, MobiAR is a compelling medium to display visual information. Participants in this evaluation think that AR technologies are an engaging tool for the provision of personalized context-based information. However, it is also clear that users are used to 2D views for searching and browsing and they do not feel so comfortable with 3D icons. This aspect should be further analysed before launching commercial applications.

Furthermore, authors would like to strength the importance given to the quantity and quality of the multimedia contents available at each PoI. It can be concluded that content will be a critical aspect for the acceptance of such applications. Finally, it should be mentioned that tourism applications are among the most-logical ways an augmented reality app can create a business model. By definition, tourists are exploring places, looking for interesting places to stay, places to eat and things to buy. That, in turn, means lots of opportunities for marketing partnerships and advertising specifically designed for SmartPhones.

Despite limitations, the participants in the survey agreed that mobile AR has potential although it needs time to mature. As mobile AR is relatively young, there are many areas for future research. New applications are introduced on a weekly basis and mobile technologies continue to improve every month. Future studies would benefit

from the results of this evaluation by examining how users perceive and interact with other forms of mobile AR applications besides information browsers.

References

Anastassova, M., Burkhardt, J.M. & Megard, C. (2007a). User-Centered Design and Evaluation of Augmented Reality Systems for Industrial Applications. In S. Richir (Ed.), *VRIC Laval Virtual Proceedings 2007,* Laval, France.

Anastassova, M., Megard, C. & Burkhardt, J.M. (2007b). Prototype Evaluation and User Needs Analysis in the early Design of Emerging Technologies. In *Human Computer Interaction 2007*. Beijing, China*:* Springer-Verlag.

Azuma, R. (1997). A Survey of Augmented Reality. *Presence: Teleoperators and Virtual Environments* 6: 355-385.

Cobos, Y., Toro, C., Sarasua, C., Vaquero, J., Linaza, M.T. & Posada, J. (2008), An Architecture for Fast Semantic Retrieval in the Film Heritage Domain, In *Proceedings of the 6th International Workshop on Content-Based Multimedia Indexing (CBMI)*. London: IEEE.

Damala, A. (2006) Evaluation strategies for mobile museum guides: a theoretical framework. In *Third International Conference of Museology & Annual Conference of AVICOM.* Mytilene.

Damala, A. (2009) *Interaction Design and Evaluation of Mobile Guides for the Museum Visit: A Case Study in Multimedia and Mobile Augmented Reality*. PhD Thesis.

Dunser, A., Grasset, R. & Billinghurst, M. (2008). *A Survey of Evaluation Techniques Used in Augmented Reality Studies,* TR-2008-02. New Zealand: HIT Lab.

Dunser, A., Grasset, R., Seichter, H. & Billinghurst, M. (2007). Applying HCI Principles in AR Systems Design. In *2nd International Workshop on Mixed Reality User Interfaces: Specification, Authoring, Adaptation (MRUI2007),* Charlotte, North Carolina, USA.

Gabbard, J. & Swan, E. (2008). Usability Engineering for Augmented Reality: Employing User Based Studies to Inform Design. *IEEE Transactions on Visualization and Computer Graphics* 14: 513-525.

Gabbard, J., Swan, E., Hix, D., Langazorta, M., Livingston, M., Brown, D. & Julier, S. (2002). Usability Engineering: Domain Analysis Activities for Augmented Reality Systems. In A. Woods, J. Merritt, S. Benton & M. Bolas (Eds.), *SPIE, Vol 4660, Stereoscopic Displays and Virtual Reality Systems IX*. San Jose, California, USA: SPIE.

Haller, M., Billinghurst, M. & Bruce, T. (2007). *Emerging Technologies of Augmented Reality, Interfaces and Design.* Hershey, USA: Idea Group Publishing.

Höllerer, T, H. & Feiner, S, K. (2004). Mobile Augmented Reality. In H. Karimi and A. Hammad (Eds.), *Telegeoinformatics:Location-Based Computing and Services.* London: Taylor and Francis Books Ltd.

Linaza, M.T., García, A., Torre, I. & Torres, J.I. (2008). Interacting with Augmented Assets in Cultural Tourism. In *Transactions on Edutainment I.* Berlin – Heidelberg: Springer.

Marimon, D., Sarasua, C., Carrasco, P., Álvarez, R., Montesa, J., Adamek, T., Romero, I., Ortega, M. & Gasco, P. (2010). Tourist Experiences through Mobile Augmented Reality. In *Proceedings of 2010 NEM Summit*, Barcelona, Spain.

Milgram, P., Takemura, H., Utsumi, A. & Kishino, F. (1994). Augmented Reality. A call for displays on the Reality-Virtuality continuum. *SPIE, Telemanipulator and Telepresence Technologies* 2351: 282-292.

Nielsen, C.M., Overgaard, M., Pedersen, M.B., Stage, J. & Stenild, S.. (2006). It's worth the Hassle! The Added Value of Evaluating the Usability of Mobile Systems in the Field. In A. Morch, K. Morgan, T. Bratteteig, G. Glosh and D. Svanaes (Eds.), *Proceedings of the 4th Nordic Conference on Human-Computer Interaction: Changing Roles*. NordiCHI'06, vol. 189. New York, NY: ACM.

Sandor, C. & Klinker, A.G. (2007). Lessons Learned in Designing Ubiquitous Augmented Reality User Interfaces. In M. Haller, M. Billinghurst and B. Thomas (Eds.), *Emerging*

Technologies of Augmented Reality, Interfaces and Design. Hershey, USA: Idea Group Publishing.

Swan, E. & Gabbard, J. (2005). Survey of User Based Experimentation in Augmented Reality. In *1st International Conference on Virtual Reality*. Las Vegas, Nevada, USA, July 22-27, 2005.

Acknowledgements

This project has been done within the MobiAR project with the funding support of the Ministry of Industry, Tourism and Commerce inside the Avanza program. Telefónica I+D participates in Torres Quevedo subprogram (MICINN), cofinanced by the European Social Fund, for Researchers recruitment.

Digitalizing Loyalty Cards in Tourism

Wolfram Höpken[a], Philipp Deubele[a], Gerhard Höll[a], Jörn Kuppe[a], Dominik Schorpp[a], Roger Olides Zapata Licones[a], and Matthias Fuchs[b]

[a]Business Informatics Group,
University of Applied Sciences Ravensburg-Weingarten, Germany
{wolfram.hoepken, deubelph, hoellge, kuppej, schorpdo, zapatalr}@hs-weingarten.de

[b]The European Tourism Research Institute (ETOUR),
Mid-Sweden University, Sweden
matthias.fuchs@miun.se

Abstract

Loyalty cards play an important role in modern customer relationship management, especially in tourism. The most common way to implement loyalty card systems is to use plastic cards, e.g. in form of magnetic cards or chip cards. However, considerable obstacles of current loyalty card systems are high hardware investments for participating retailers, which are even more noticeable in tourism due to the great diversity of potential participants in tourism destinations. This paper presents an approach for digitalizing loyalty cards based on quick response (QR) codes. Both, customer and retailer make use of their mobile phone to present and validate the QR code as digital customer card and offer corresponding discounts and (e.g. tourism or information) services, without the need for additional hardware investments. A prototypical implementation has been evaluated in a real life scenario in a German tourism destination.

Keywords: digital loyalty card, QR code, mobile application, CRM

1. Introduction

Loyalty cards offer companies a popular opportunity to improve customer relationship management (CRM) (Mauri, 2003). Currently, loyalty cards are mostly available as plastic cards, e.g. in the form of magnetic cards or chip cards. These cards are particularly used to increase customer loyalty and to attract new customers by offering discounted services in case of repeated purchase (discount card), special benefits or services for registered customers (bonus or club card) and payment services. Next to these customer benefits, loyalty card programs typically generate a huge amount of customer transaction data, constituting a valuable input to further CRM activities and market intelligence, like customer profiling and consumer behaviour analysis (Zilliani & Bellini, 2004). Due to raising customer expectations and decreasing customer loyalty a strong need for intensified customer relationship management emerged (Pan & Lee, 2003), thus, especially tourism destinations intensified the use of loyalty card programs in recent years, e.g. *Zell am See-Kaprun card* (Stärz, 2010), *Freizeitticket Tirol* (Schroll, 2010) or *Mieminger Plateau card* (Lampe, 2010). However, particularly in a tourism context, traditional loyalty card systems exhibit serious obstacles. First, magnetic cards or chip cards typically lead to high investment costs for the technical infrastructure, especially card reading devices at the variety of points of sale (POS) (Wittbrodt, 1995). Second, the usage of separate plastic cards by different loyalty card systems leads to inconveniences for customers and, consequently to reduced user acceptance.

This paper presents an approach to overcome above obstacles by digitalizing traditional loyalty cards through the combined use of quick response codes (QR codes) and mobile phones. Both, customer and retailer make use of their mobile phone to (1) present the QR code as digital customer card to the retailer and (2) validate the QR code and offer corresponding discounts and services. The digital customer card does not need any hardware investments and significantly improves customer-side handling. Based on a mobile application on the retailers' mobile phone, a flexible management and redeeming of various types of coupons is supported.

The paper is structured as follows. Section two gives a brief description of loyalty cards in tourism as well as QR codes and typical applications. The third section presents the concept and the architecture of the digital loyalty card system. It shows the general work flow, the concept of coupons and the technical infrastructure. Section four is about the prototypical implementation and evaluation of the digital loyalty card system in a tourism destination. The last section summarizes the results and gives an outlook on future research activities.

2. Background

2.1 QR Codes

Traditional (one-dimensional) bar codes evolved to 2D codes in order to increase reading speed, accuracy and data capacity (Lenk, 2002). Different types of 2D codes are *composite codes*, *stacked codes*, *dot codes* and *matrix codes*. QR codes are a kind of matrix code developed by Denso Wave in 1994 and first used by the Japanese automobile industry for identifying car parts in the production line. Key features of QR codes (established as an ISO standard ISO/IEC18004) are high capacity encoding of data, small printout size, resistance to dirt and damage and 360° readability (Denso, 2010).

The QR code structure consists of three patterns for position detection, a pattern for alignment and a data area. The patterns for position detection and alignment allow rapid identification and omnidirectional readability of QR codes (Soon, 2008). QR codes can be divided into up to 16 data areas, increasing the data capacity by a small printing size. The data area itself consists of control zones for tact, version and format and reserved areas for user data and error correction. Based on the format and version identifiers, the code type and size are recognized. The actual data block is divided into sections for user data and error checking and correction (ECC), depending on the chosen ECC level. Error checking and correction is typically based on the Reed-Solomon algorithm (Lenk, 2002). Examples of QR codes with one (left) or several data areas (right) are shown in Fig. 1.

Fig. 1. Examples of QR codes

2.2 Applications of QR Codes in Tourism

Typically, QR codes serve the purpose of providing encoded information for authentication (mobile couponing) or linking internet resources with physical objects (mobile tagging). The encoded information can be captured with handheld scanners, fix-placed charge couple devices (CCD) and especially mobile phones. The use of the latter grounds the variety of application scenarios for QR codes especially in tourism (Canadi et al., 2010; Chen & Weng, 2010).

For example, the city of Alberta (CAN) uses QR codes to link to videos that bring to life various Canadian destinations, cultural experiences and vacation packages (McBurney, 2011). The Fort Smith National Historic Site in Arkansas uses QR codes to link printed advertising material or points of interest with additional information on the Web (Woerner, 2010).

3. Digital Loyalty Card System Architecture

3.1. General Concept

With the digital loyalty card system, a technical solution is offered, which allows a group of retailers to form a marketing alliance and, by offering bonuses and benefits, acquire new customers and increase customer loyalty. The digital loyalty card system replaces existing *plastic cards* by a QR code provided on the customer's mobile phone and corresponding specific reading devices by a mobile application on the retailer's mobile phone (cf. Canadi, 2011).

Fig. 2 illustrates the general usage scenario of the digital loyalty card.

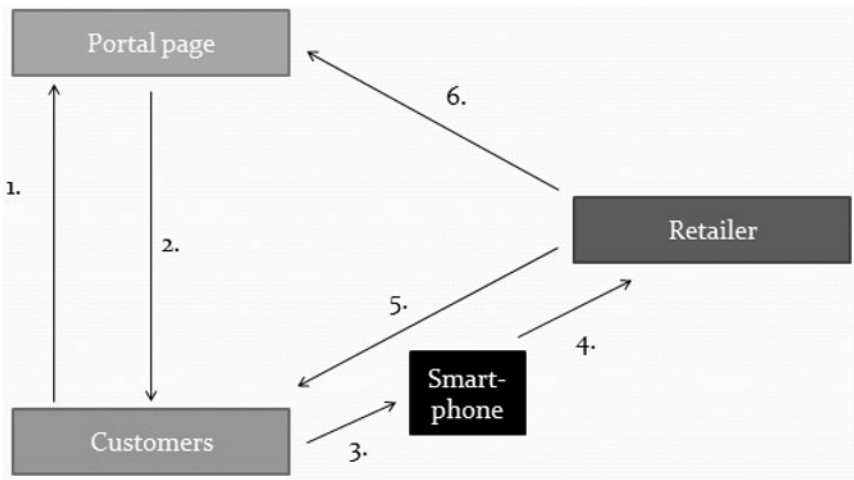


Fig. 2. Digital loyalty card usage scenario

(1) Customers join the loyalty program by registering on a portal page.

(2) At any time, customers can get information on available bonuses and benefits.

(3) Customers download a personal QR-code onto their mobile devices, which represents the digital loyalty card.

(4) At the POS, customers issue their digital loyalty card for authentication by presenting the QR code (as a stored picture or MMS) on their mobile phone.

(5) The retailer authenticates the customer by scanning the QR code using his mobile phone, offers the available bonuses and benefits to the customer and books the bought products or used services.

(6) The collected purchase data is sent to the portal back-end system as input to update available bonuses and benefits or for consecutive analysis of customers' purchase behaviour. As the backend system centrally collects purchase data of all customers, several or all suppliers can co-create certain bonuses or benefits, e.g. in the case of a points system.

The proposed approach of the digital loyalty card overcomes the limitations of traditional loyalty cards described in the introduction and, moreover, offers significant additional advantages. First, plastic cards are replaced by QR codes and only marginally costs occur for their production and distribution to the customer. Additionally, handling (especially multiple) loyalty cards by the customer is simplified and, thus, will significantly increase user acceptance.

Second, no hardware investments into specific reading devices at the POS are necessary, while, at the same time, securing a comfortable and robust authentication process based on QR codes. Finally, especially compared to offline loyalty programs, the digital loyalty card system enables a flexible and dynamic management of bonuses and benefits by each single retailer.

3.2. Technical Architecture

The technical architecture of the digital loyalty card system is shown in Fig. 3.

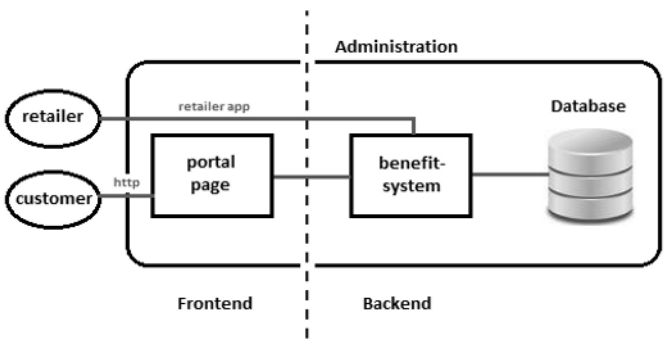


Fig. 3. Technical architecture of the digital loyalty card system

The *frontend layer* consists of the *online portal page* (cf. section 0) used by the customer to register to the loyalty program and gain access to information about available bonuses and benefits, and the *retailer client (app)* (cf. section 0), a mobile application for retailers to scan customers' QR codes, inform about available bonuses and benefits and book purchased products and services. The *backend layer* consists of the *benefit system* (cf. section 3.3), enabling the flexible management of bonuses and benefits, and the *database*, storing operational data as well as non-operational data for consecutive analyses.

Online Portal

As outlined, the online portal serves as customers' main entrance point to the loyalty card system, offering functions, like maintain customer profile, view available bonuses and benefits and provide information on special events and news. The online portal is implemented by object-oriented PHP5 and offers a customized layout structure for mobile devices identified by their user agent. After registration, based on a double-opt-in mechanism, the personal QR code is generated and displayed as part of the customer profile.

The customer can transfer the QR code to his mobile phone by (1) simply taking a picture and storing the QR code picture on the mobile phone, (2) requesting the QR code to be send by MMS or email or (3) access the online portal directly via mobile phone.

Retailer Client

Each retailer participating in the loyalty program needs a device to authenticate customers by scanning and verifying the customer's QR code, get information about bonuses and benefits available for a specific customer and book purchased products and services. In order to avoid any additional hardware investments, this functionality is provided by a mobile application executed on the mobile phone of the retailer.

Fig. 4 illustrates the architecture of the retailer client, implemented as an Android application. The *controller* manages the communication between the graphical user interface (GUI), the service module and the QR code reader.

As QR code reader an external application from the Android Market is used, which is automatically downloaded if necessary. The access of the retailer client to the corresponding functionality of the backend system for managing bonuses and benefits is realised by web services, implemented based on a REST architecture (managing services via HTTP GET requests) and JSON as platform-independent message format. The *service* module provides the web services' functionalities within the client environment. The following services are available:

- setup: Initialization of the client on the back-end system using a client identifier (mobile device) and a retailer's identifier.
- getVouchers: Request all vouchers (bonuses and benefits) available for a specific customer and retailer.
- redeemVoucher: Redeem a voucher after consumption by the customer.

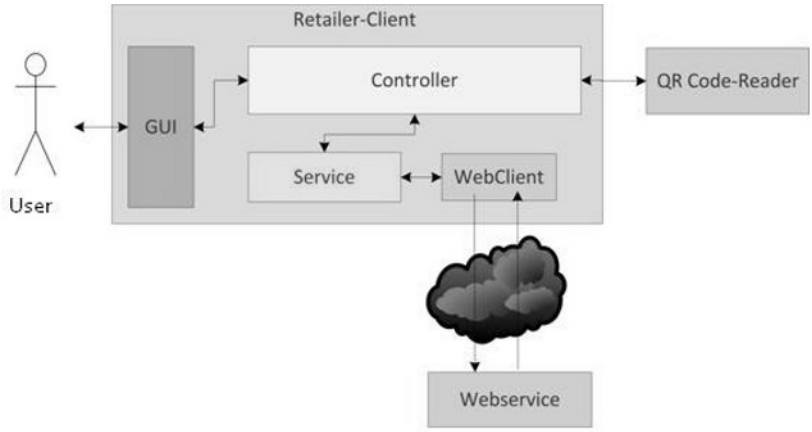


Fig. 4. Architecture of the retailer client

3.3. Benefit System

The benefit system manages the different types of benefits and offers its functionality to the different components of the frontend layer (cf. Fig. 3). Benefits are represented as rules, following the concept of ECA rules (Beer, 2007), enabling their flexible and powerful definition. Rules for benefits consist of different components, illustrated in Fig. 5.

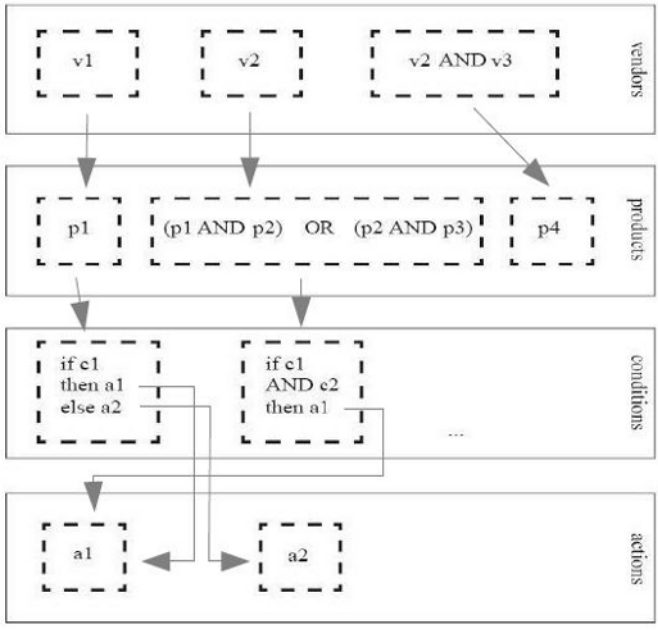


Fig. 5. Components of rules for bonuses and benefits

The first component of a rule specifies the *vendor* (i.e. retailer) who provides the benefit. Usually a benefit is provided by one vendor. However, it is also conceivable that several vendors participate in providing a benefit. For example, a points system, offering a benefit based on points consecutively collected over several purchases at a tourism destination, would typically involve multiple vendors. The second component of a rule specifies the *products*, triggering the benefit, in form of single products, a range of products as well as complex Boolean expressions. *Conditions* of a rule enable to specify pre-conditions in form of "if - then - else" expressions, enabling the vendor to restrict the availability of benefits to certain context parameters, e.g. time or customer characteristics. The *action* of a rule specifies the benefit the customer is granted by the vendor, like a discount, free service or points within a point system. Actions are triggered by conditions and can be arranged in a timely manner.

The presented rule-based approach for managing bonuses and benefits enables the benefit system to be highly flexible and powerful. Consequently, a variety of different types of benefits can be modelled to fit the retailers' needs. Benefits can be defined and adjusted easily and at any time.

4. Evaluation

4.1. Focus Group Interview

As first step of evaluation the presented digital loyalty card system has been tested and discussed in the course of a focus group interview. Focus group interviews are a well-known method in usability engineering to obtain subjective but valuable knowledge on the users' perception of an IT application, based on an interactive group discussion (Nielsen 1993) The focus group consisted of 5 students (three male, two female) enrolled in a course in information technology in tourism, with an age between 22 and 25 and fairly good technical backgrounds and, thus, represents the target segment for the digital loyalty card system (i.e. early adopters) well.

The focus group interview started with an introduction to the general approach of the digital loyalty card system and its prototypical implementation. The focus group members then experienced the prototype by executing several real-life usage scenarios, both from a customer as well as a retailer perspective. The experiment has been observed by the project team to gain first insights into usability and possible usage barriers. Afterwards, the focus group members evaluated the system based on a prepared questionnaire (cf. Table 3). Finally, a moderated focus group discussion has been executed to identify major strengths and weaknesses of the presented approach and gain insights into reasons and motivations for the qualitative judgements concerning usability and usefulness, already quantitatively expressed within the questionnaire.

Table 3 presents the obtained questionnaire results. For each role, retailer (R) and customer (C), questions are segmented into three topics: user registration and log-in (T1), main work flow (T2) and, finally, general questions and individual judgements (T3). Answers were given on a five-point scale ranging from 0 (i.e. strong disagreement) to 5 (i.e. strong agreement). Although the gathered quantitative results cannot be considered as representative due to the small number of probands typically

involved in focus group discussions, results clearly indicate that usability and usefulness are consistently judged as high and satisfactory. Q17 (R-T2) achieves an average value of 4.0 and, although still high, marks the relatively lowest average result throughout the questionnaire, indicating that a potential disturbance or retardation of the selling process at the POS may constitute the likely biggest threat to the system's acceptance by the retailers.

Table 3. Results of focus group questionnaire

Roles & Topics	Statements	Av. Rating
C – T1	Q1: I had no problems to register on the portal.	4.8
C – T1	Q2: I had no problems to log-in and log-out.	4.8
C – T1	Q3: I could insert my master data without any problems.	4.8
C – T1	Q4: I had no problems to view available benefits	4.5
C – T1	Q5: I had no doubt to insert my personal data.	4.4
C – T2	Q6: I had no problems to photograph the QR code with my smartphone.	4.6
C – T2	Q7: I had no problems to find the QR code on my smartphone.	5.0
C – T3	Q8: The mobile digital loyalty card is a reasonable alternative to the old plastic cards.	4.8
C – T3	Q9: I can imagine registering to other digital loyalty card programs.	4.5
C – T3	Q10: I had no problems with the storage of the transactions.	4.3
C – T3	Q11: I had no problems to get along on this website.	4.8
C – T3	Q12: I had no problems with the identification with my QR-code.	4.8
R – T1	Q13: I had no problems to find and start the app quickly.	4.8
R – T1	Q14: I had no problems to login with my device.	4.6
R – T2	Q15: The scanning process of customer's QR code did not make any problems.	4.4
R – T2	Q16: I had no problems to redeem the coupons on my smartphone.	5.0
R – T2	Q17: The scanning process did not bother the selling process.	4.0
R – T3	Q18: Mobile digital loyalty cards are a reasonable alternative to old plastic cards.	4.6
R – T3	Q19: I can imagine using the digital loyalty card system permanently.	4.4

The moderated discussion, as main part of the focus group interview, revealed further interesting insights into the perception of the digital loyalty card system, thus, explaining or underpinning the quantitative results in Table 3. In general, the overall application (online portal and retailer client) is judged very positive and as comfortable to use by all focus group members. Specifically from a customer's perspective the following judgements were most frequently stated:

- The registration and log-in process as well as the download and handling of QR codes for authentication at the POS have been judged as reasonable and very easy to use.
- The application is perceived as usable even on older mobile phones. Thus, this specific form of QR code usage is not restricted to modern or expensive smart phones.
- At the same time, the application is considered as impractical for older people or users without a minimum of technical background. Although a print function for the QR code to be used without a mobile phone is offered, the focus group members doubt that older people are willing to use such a system. However, and qualifying the statement, this is not viewed as a specific problem of the presented system but of most technical innovations.

From a retailer's perspective, the focus group members stated the following judgements:

- As no additional hardware investments for retailers are needed (as one of the main advantages of the proposed system), no costs for using the system should arise for retailers, which is especially important in the context of tourism destinations with a variety of small- and micro-sized retailers as potential participants (although the retailer client is currently only available as Android application, it can be provided for a wide variety of even older mobile phones in principal).
- The overall effort for setting up and maintaining the central system is viewed as problematic for a single retailer but as reasonable for a group of retailers forming a marketing alliance, what typically occurs in a tourism destination context.
- Specifically, for restaurants or bars, the loyalty card system should be integrated into mobile order terminals, as using two different devices is viewed as impractical.

4.2. Pilot Application

The presented digital loyalty card system has been instantiated and implemented for the German tourism destination "Bierkulturstadt Ehingen" and is in use since 1st of July 2011. Bierkulturstadt Ehingen is a theme-oriented marketing alliance, supporting local breweries and associated service providers, like restaurants, beer gardens or gift shops and linking them with cultural activities. Collaborating retailers have been introduced to the digital loyalty card system and corresponding benefits (in form of rules presented in section 3.3) have been defined. Fig. 6 shows the instance of the online portal developed for Bierkulturstadt Ehingen with the dialog displaying the customer's QR code. Fig. 7 shows the main screen of the developed retailer client. The participating retailers successfully use the retailer client application on different types of Android-based mobile phones. The retailers either use a direct Internet connection via a cellular network or an existing WLAN hotspot to connect to the central system. The retailer client and the underlying work flow proofed feasible for the retailers and has not been perceived as retarding the sales process so far.

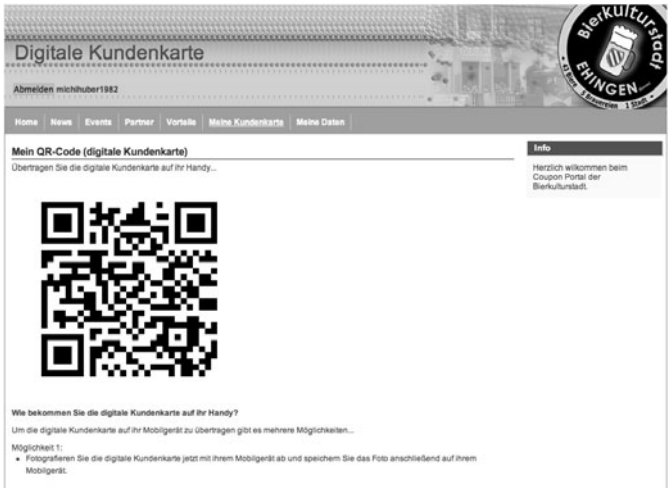


Fig. 6. Online portal with QR code

Fig. 7. Retailer client main screen

5. Conclusion and Outlook

5.1. Conclusion

The paper presented an approach for digitalising loyalty card systems by making use of QR codes and mobile phones instead of traditional plastic cards, like magnetic cards or chip cards. A technical framework, consisting of an online portal and a mobile retailer client as front-end components and a benefit system and database as back-end components has been developed and a pilot system for the German tourism destination "Bierkulturstadt Ehingen" has been instantiated.

It could be shown that the proposed concept of using QR codes and mobile phones is suitable in a real-world scenario. A focus group interview revealed that the handling of all system components is easy and fast and the process of customer authentication and redeeming of benefits is smooth and without any difficulties. However, the risk of retarding the sales process is seen as most dangerous obstacle to retailers' system acceptance. A pilot phase in the destination of Ehingen, starting at the 1st of July

2011, confirmed the positive results of the focus group interview. Unfortunately, a lack of marketing activities led to a limited number of users, so far, which can be identified as a general obstacle of introducing technical innovations to tourism destinations. Participating stakeholders, like retailers in this case, tend towards a more observant and reactive behaviour due to a lack of IT affinity (Fuchs et al., 2010). Local authorities, like DMOs, often don't have the power and access to the market to sufficiently attract potential customers.

Overall, the approach of using QR codes and mobile phones for digitalizing loyalty card systems has proofed as a reasonable alternative to traditional plastic cards. No hardware investments at each POS are necessary, giving even small retailers the chance to participate as overall costs for implementing and operating the digital loyalty card system are comparably low. Finally, from a customer perspective the approach convinces by a comfortable handling of the digital loyalty card.

5.2. Outlook

Currently, a full functional digital loyalty card system is available and in operation as pilot application in Ehingen (Germany). Of course, after completion of the pilot test, more information on usage behaviour and user acceptance will be available for further analyses. In parallel, the digital loyalty card system will be further improved by extended features. (1) Although, the online portal adapts itself to mobile devices, a specific mobile application (as Android app) for customers is planned, to simplify customers' interaction with the loyalty card system and offer additional functionality, like a location-based service, visualising surrounding participating retailers in a map (Höpken et al. 2008; Chen & Weng, 2010). (2) The benefit system will be extended by a rule engine, encapsulating the handling of rules for bonuses and benefits, and enabling an easier and more powerful definition and maintenance of rules. The application of extended versions of the digital loyalty card system is planned for further tourism destinations, like the region of Upper Swabia (Germany).

References

Beer, Th., Rasinger, J., Höpken, W., Fuchs, M. & Werthner, H. (2007). Exploiting E-C-A Rules for Defining and Processing Context-Aware Push Messages. In *International Symposium on Rule Interchange and Applications (RuleML)*, Orlando, pp. 199-206.

Canadi, M., Höpken, W. & Fuchs, M. (2011). Virtualisation of Customer Cards with 2D Codes. In R. Law, M. Fuchs and F. Ricci. (Eds.). *Information and Communication Technologies in Tourism.* New York: Springer, pp. 215-227.

Canadi, M., Höpken, W. & Fuchs, M. (2010). Application of QR Codes in Online Travel Distribution. In U. Gretzel, R. Law, and M. Fuchs (Eds.). *Information and Communication Technologies in Tourism*. New York: Springer, pp. 137-148.

Chen, C. & Weng, T. (2010). Research and Development of Applications of Mobile Barcode to Mobile Sightseeing Guide on Mobile Phone. *Transactions on Information Science and Applications* 7 (1): 16-25.

Denso (2010). *QR Code.com. Denso Wave Incorporated*. Retrieved from http://www.denso-wave.com/qrcode/qrfeature-e.html.

Fuchs, M., Höpken, W., Föger, A. & Kunz, M. (2010). E-Business Readiness, Intensity, and Impact – An Austrian Destination Management Organization Study. *Journal of Travel Research* 49 (2): 165-178.

Höpken, W., Scheuringer, M., Linke, D. & Fuchs, M. (2008). Context-based Adaptation of Ubiquitous Web Applications in Tourism. In P. O'Connor, W. Höpken, and U. Gretzel (Eds.), *Information and Communication Technologies in Tourism*. New York: Springer, pp. 533-544.

Lampe, M. (2010). Tourismuspress.at: *Mieminger Gästekarte als touristischer Joker im Winter 2009/10*. Retrieved August 12, 2009 from http://www.tourismuspresse.at/ presseaussendung/TPT_20100811_TPT0001, Obsteig, Austria.

Lenk, B. (2002). Handbuch der automatischen Identifikation: *2D Codes, Matrixcodes, Stapelcodes, Composite Codes, Dotcodes.* Vol. 2, Monika-Lenk Fachbuchverlag, Kirchheim unter Teck.

Mauri, C. (2003). Card Loyalty - A New Emerging Issue in Grocery Retailing. *Journal of Retailing and Consumer Services* 10(1): 13-25.

McBurney, R. (2011). *Mobile Discovery's QR Code Campaign for Canada Tourism Makes Headlines Across the U.S.* PRWeb. Retrieved from http://www.prweb.com/ releases/2010/06/ prweb4120084.htm.

Nielsen, J. (1993). *Usability Engineering*. Boston: Academic Press.

Pan, S. & Lee, J-N. (2003). Using e-CRM for a Unified View of the Customer, *Communications of the ACM* 46(4): 95-99.

Schroll, T. (2010). Tourismuspress.at: *Freizeitticket Tirol jetzt mit den 3 Innsbrucker Hallenbädern, Bergbahnen Hochkössen und Kreithlifte Leutasch*. Retrieved September 21, 2009, from http://www.tourismuspresse.at/presseaussendung/ TPT_20100916_TPT0001, Innsbruck, Austria.

Soon, T. J. (2008). *The Synthesis Journal:* Section three – the QR Code. iTSC Information Thechnology Standard Comitee, Singapore, pp. 59-78.

Stärz, O. (2010). Tourismuspress.at: *Sattes Plus dank der neuen Zell am See-Kaprun Karte*. Retrieved September 11, 2009, from http://www.tourismuspresse.at/presseaussendung/ TPT_20100812_TPT0002, Zell am See-Kaprun, Austria.

Wittbrodt, E. J. (1995). *Kunden binden mit Kundenkarten; Kundenbindungssysteme entwickeln, einführen, steuern*. Köln: Luchterhand, pp. 77-81.

Woerner, D. (2010). QR Codes Debut at Fort Smith National Historic Site. Arkansas Department of Parks & Tourism, Retrieved from http://www.visitmyarkansas.com/dena-woerner/ default.aspx?id=622.

Ziliani, C. & Bellini, S. (2004). From Loyalty Cards to Micro-Marketing Strategies. *Journal of Targeting, Measurement and Analysis for Marketing* 12(3): 281-289.

Evaluation of Intelligent Routes for Personalised Electronic Tourist Guides

Ander Garcia[a],
Maria Teresa Linaza[a], and
Olatz Arbelaitz[b]

[a] Dept. of eTourism and Cultural Heritage
Vicomtech, Spain
{agarcia, mtlinaza}@vicomtech.org

[b] Dept. of Computer Architecture and Technology
University of the Basque Country, Spain
olatz.arbelaitz@ehu.es

Abstract

Personalised Electronic Tourist Guides (PETs) provide an integrated solution for route generation based on the profiles and constraints of tourists, up-to-date Points Of Interest (POIs) and destination information. In this paper we present the result of an evaluation of a PET prototype that applies an advanced algorithm to generate personalised tourists routes including public transportation. The prototype has three main functionalities: recommendation, personalised route generation and route customisation. The validation scenario of the prototype has been the city of San Sebastian. The result of the validation has been positive and it showed that PETs are perceived as interesting tools by tourists. Moreover, both the personalised route generation and the inclusion of public transportation are perceived as valuable functionalities.

Keywords: Personalised Electronic Tourist Guide, public transportation, evaluation

1. Introduction

The main objective of Personalised Electronic Tourist Guides (PETs) (Garcia et al., 2009) is to provide an integrated solution for personalised route generation based on the profile and constraints of tourists, and up-to-date information about Points Of Interest (POIs) and destination information. Currently, this is a time consuming task that is often done by the Local Tourist Office (LTO) staff.

However, routes generated by the staff do not take into account circumstances that may occur during the visit (longer time spent visiting an attraction, changes in the weather, transportation delays, etc.) (Dunlop, Ptasinski, & Morrison, 2004). Moreover, due to physical (office location, available space, etc.) and temporal constraints (timetables, travel times, etc.), it is not possible for all tourists to visit a LTO. PETs help tourists in these tasks, offering added-value functionalities (Vansteenwegen & Oudheusden, 2007). The personalised tourist route generation process of PETs can be described by three basic steps (Fig. 1) (Garcia et al., 2010).

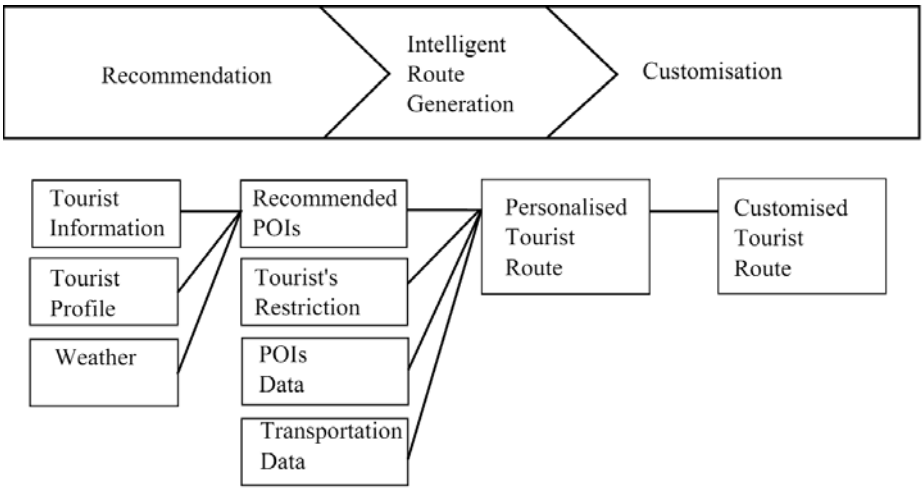


Fig. 1. Steps of the personalised tourist route generation process (Garcia et al., 2010)

- **Recommendation**. A list of recommended POIs can be generated combining information about the destination and tourist profiles. Thus, each POI should have a different score and visit duration for different tourist profiles. For each tourist, these values are stored in their Personal Interest Profile. Interested readers can find a comprehensive review of tourist recommendation systems in a recent paper (Kabassi, 2010).
- **Intelligent Route Generation**. Once the system has determined which the most appealing POIs for the tourist are, an intelligent routing engine applies an algorithm combining this information with the restrictions of the tourist (available time, duration of the route, budget, start POI); POI data (location, opening hours, ticket price); destination context data (weather, special events); and transportation data (travel times, public transportation network data) to generate personalised tourist routes.
- **Customisation**. Finally, tourists can customise the proposed personalised route to better fit their needs. Inserting new visits and removing or reordering visits on the route are the basic functionalities of a customisation engine.

This paper summarises the validation of a PET prototype that has been developed to evaluate route generation functionalities. The prototype applies an advanced algorithm to generate personalised routes including public transportation. The objective of the validation is to analyse if tourists perceive these functionalities offer an added value, encouraging their inclusion in next generation travel guides.

The paper is organised as follows. First, we present a summary of related work. The next Section introduces the prototype of the validation. Finally, we present the results of the validation. The last Section remarks the main conclusions and some future work lines.

2. Related work

2.1 Route Generation Functionalities of PETs

During the previous decade, advances both in hardware and mobile networks have fostered the development of new PET prototypes, which have received different names such as Mobile Tourist Guides, Personal Navigation Systems for Tourism, Electronic Tourist Guides or even Tourist Decision Support Systems. Existing PETs have been thoroughly reviewed in the literature (Souffriau et al., 2008; Kenteris, Gavalas, & Economou, 2010; Garcia, 2011).

Although requirements and the main functionalities of PETs were detected while evaluating initial prototypes, not all of the functionalities have received the same attention. Specifically, there are mature examples of recommendation functionalities available (Kabassi, 2010), while the route generation functionality has been relegated to a secondary category, not applying advanced algorithms from other fields.

The generic personalised tourist route generation problem has been defined as the Tourist Trip Design Problem (TTDP) (Vansteenwegen & Oudheusden, 2007). The TTDP presents a general model that serves as a reference for the implementation of PETs. The TTDP describes the characteristics that a perfect PET should meet regarding route generation, integrating the selection and planning of visits to POIs. Thus, its solution selects the best combination of interesting POIs for a tourist and schedules a feasible route.

As the TTDP presents several properties of an ideal PET, it is difficult for real PETs to implement all of these properties. Thus, researchers have proposed different problems to model the route generation problem as simplified versions of the TTDP. One of the simplest problems that can be used as a basic model of the TTDP is the Orienteering Problem (OP) (Tsiligirides, 1984). The OP is based on the orienteering game, in which several locations with an associated score have to be visited in order to obtain a total route score. Each player can visit each location only once.

Extensions of the OP have been successfully applied to model the TTDP. One of the most advanced extensions is the Team Orienteering Problem with Time Windows (TOPTW) (Savelsbergh, 1985), which includes multiple routes and time windows that can be applied to model opening hours of the POIs and multiple day routes. The prototype of this paper introduces a solution to the Time Dependent Multi Constrained Team Orienteering Problem with Time Windows (TDMCTOPTW) (Garcia et al., 2011), an extension of the OP that can be applied to model opening hours of the POIs, multiple day routes, multiple tourist constraints (for example a maximum budget), and public transportation.

Focusing on existing PETs, the oldest examples of prototypes generating routes have proposed approaches based on models of the TTDP and algorithms that are either too simple or not efficient. Two approaches, P-Tour (Maruyama et al., 2004) and DTG (Hagen et al., 2005) propose advanced models but fail to implement an efficient solution. However, the most recent example, City Trip Planner (Vansteenwegen *et al.*, 2010), models the TTDP as the TOPTW and proposes and efficient solution algorithm to solve it in real-time. Finally, m-Trip is a commercial PET (www.mtrip.com., [July

10, 2011]) that applies a custom engine, called Genius, to generate personalised tourist routes for the main cities of the world. The PET evaluated in this paper (Garcia et al., 2009; Garcia et al., 2010) applies a solution algorithm for the TDMCTOPTW that generates routes in real-time. This is the first example of both a solution algorithm for the TDMCTOPTW, and a PET generating personalised route in real-time and including public transportation, which was identified as one of the most valuable functionalities of PETs (Schmidt-Belz et al., 2003; Stroobants, 2006; Beer et al., 2007).

Although there are algorithms (Delling et al., 2009; Bauer et al., 2010) to generate routes between two locations (for example a route from the hotel to the cathedral) and including public transportation, the problem they solve is different. This problem, known as The Time Dependent Shortest Path Problem (TDSPP) (Cooke & Halsey, 1966) or the Earliest Arrival Problem (Pyrga et al., 2008) is a time-dependent problem for individual queries between locations. Thus, this problem does not model the selection of the locations that should be included in the route and its solution algorithms are not suitable for the generation of personalised tourist routes.

2.2 Evaluation of Existing PETs

Researchers have conducted evaluations of some existing PETs focusing on their tourist features and applying methodologies based on direct observation of tourists, semi-structured interviews, questionnaires and on-line surveys. For example, Chevers et al. (2000) presented their experiences developing and evaluating the GUIDE prototype, concluding that tourists appreciated the benefits of their PET. Some years later, Schmidt-Belz et al. (2003) remarked the importance of offering new services (such as services related to transportation) as a conclusion of their evaluation of the CRUMPET prototype.

More recently, Kramer et al. (2007) focused on the behaviour of tourists following routes generated by the DTG prototype, suggesting PETs could help tourists enjoy the full potential of destinations. Finally, Rasinger, Fuchs, and Höpken (2007) presented a survey on usage intention of different tourist mobile information services in tourism, identifying transportation information as one of the top-six services.

However, the recent mobile revolution (smart-phones, mobile Internet ...) has opened new development possibilities for PETs that were not available some years ago. For example, none of the existing PET prototypes could generate personalised tourist routes including public transportation, which is one of the functionalities tourists find more valuable.

Thus, after having successfully developed a PET prototype that could generate personalised tourist routes including public transportation in real-time, we have validated it in a real scenario. The evaluation presented in this paper analyses the importance of personalised route generation and public transportation functionalities for PETs, fostering further research and developments that take advantage of the new technical possibilities available nowadays.

3. Description of the Prototype

3.1 Main Functionalities of the Prototype

Although PETs can include several functionalities, in order not to deviate from the main objective of the validation, we have focused on the route generation functionalities. Thus, we have avoided the development of a complete PET, which would have required a considerable development effort in areas already covered by prior research. The prototype (Garcia et al., 2009; Garcia et al., 2010) is a Web application that includes the following functionalities (Fig. 2):

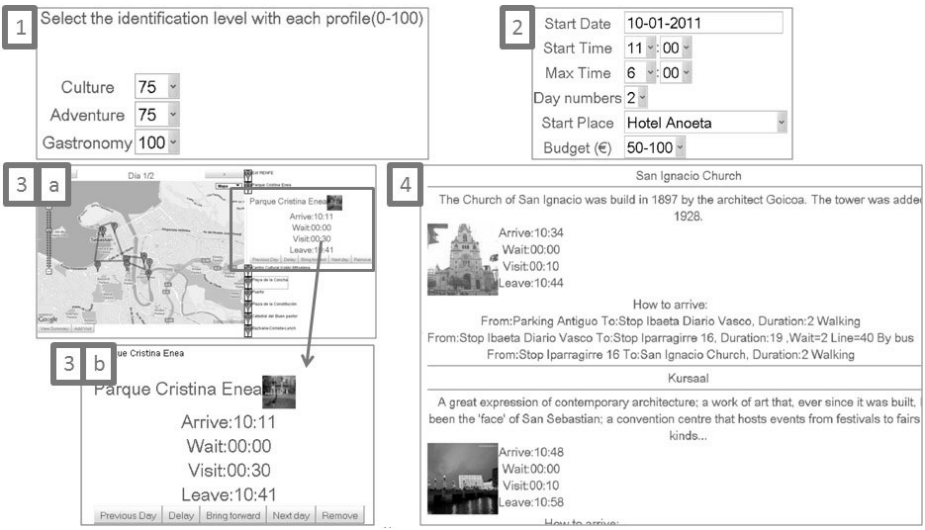


Fig. 2. Functionalities of the prototype

- Generation of personalised tourist routes. The prototype allows tourists to generate personalised tourist routes that maximise their enjoyment when visiting a city. First, tourists have to enter some data about their preferences (Fig. 2-1) and the constraints of their route (Fig. 2-2).

  As it has been remarked in previous sections, recommendation is not the main functionality of the prototype. The recommendation functionality of the prototype is based on a matrix that relates tourist profiles defined by the LTO and POIs. Tourists using the system have to define their profile, introducing to what degree they feel identified with the defined profiles. More advanced recommendation techniques are summarised by Kabassi (2010) and are not within the scope of this paper.

  Then, a personalised route is generated and showed on a map with the details of each visit (Fig. 2-3a). Tourists are allowed to customise the proposed route to better meet their requirements (Fig. 2-3b). Finally, tourists can generate a detailed summary and a PDF of the route (Fig. 2-4).

- Inclusion of public transportation. The generated routes should encourage the use of public transportation, when convenient. The summary of the route shows how to move between visits, including the details about the public transportation stops and services.
- Customisation of the generated routes. Tourists can refine the proposed route (Fig. 2-3b) applying six basic operations: add a visit; remove a visit; move a visit towards the beginning of the route; move a visit towards the end of the route; move a visit to the previous day; and move a visit to the next day.
- Real-time execution. The prototype runs in real-time, generating and visualizing the personalised route in less than five seconds to provide an interactive experience.

3.2 Implemented Algorithm

The prototype applies an advanced algorithm to generate personalised tourist routes in real-time including public transportation (Garcia et al., 2011). We have modelled the TTDP as the Time Dependent Multi Constrained Team Orienteering Problem with Time Windows (TDMCTOPTW). A typical TDMCTOPTW contains a number of POIs with a fixed location, opening hours (time windows), and a given score; and several constraints. Movements between POIs can be done on foot or by public transportation. The public transportation network is defined by a number of fixed stops and different lines between these stops, each of them with a given frequency.

The algorithm applies a real calculation approach modelling public transportation transfers as direct connections. The approach is based on a fast evaluation of the possible insertion of an extra POI to a route, evaluating each insertion locally and efficiently. Our algorithm is the first solving this type of problems in real-time. Interested readers can access a thorough description of the algorithm (Garcia et al., 2011).

We have already validated the efficiency of the algorithm using test instances based on real data about the city of San Sebastian and it is able to generate routes in less than one second for 50 POIs and two day routes. In this paper we focus on the validation of the PET prototype integrating this algorithm from the tourist point of view.

4 Validation

4.1 Objectives of the Validation

The objectives of the validation of the prototype are twofold. On the one hand, we want to check the real world viability of the approach to generate personalised tourist routes within the tourism domain. On the other hand, we have used the validation to discover more aspects on the opinion of tourists relating to route generation functionalities and PETs: what they expect and what they would like a perfect PET to include.

4.2 Methodology of the Validation

We used a quantitative analysis approach for the validation. Data was collected on the basis of a questionnaire developed for the usability analysis, including aspects related to different constructors such as the perceived usefulness (PU) of the prototype, its perceived ease of use (PEOU) or the perceived added-value.

The first part of the questionnaire focused on the profile of the user and was composed of demographic data (age, gender, nationality, country of residence), general information (first time visitor, languages spoken, prior technological and mobile knowledge), and tourist preferences (information sources, tourist interests, transportation means). The second part focused on parameters of the Technology Acceptance Model (TAM) and the assessment of the prototype, the ease of use of the functionalities and the perceived value of the prototype. The last part focused on the assessment of PETs, asking tourists about the services they would expect from an ideal PET and their willingness to pay for them.

The validation process started by giving tourists a short explanation of the prototype. Then, we asked them to generate a route for themselves and to customise it until they were satisfied with the route. Finally, we asked them to complete the questionnaire. The whole process took around 30 minutes for each tourist. The answers were processed and analysed to form a judgement on the technical performance of the system, usability aspects of the prototype and overall aspects related to PETs.

4.3 Validation Scenario

The prototype has been validated in the city of San Sebastian, which is a medium size city of approximately 200,000 inhabitants and 50 POIs distributed throughout the city. Although most of the POIs are located around the three beaches of the city and the city centre, it is not desirable to visit all POIs on foot. Therefore, most of the tourists visit the city combining public transportation with short walks.

4.4 Sample Data

Twenty tourists and professionals from the tourism sector (researchers and Information Technology company employees) took part in the validation (11 men and 9 women) in November 2010. 15 of them were between 20 and 39 years old; three between 40 and 59; and two of them between 10 and 19. 18 of them were Spanish (half of them from outside the region of Gipuzkoa); one was Irish; and another was French. Regarding their previous knowledge about the city, most of them lived in, or near, San Sebastian (11 of them); six were frequent visitors to the city; and three of them had prior knowledge of the city.

Regarding their technological background, most of them were familiar with new technologies. All of them often used PCs, the Internet and a mobile phone; and had at least infrequently used digital cameras (both stand alone and mobile phone cameras). GPS units were not so familiar, as five tourists had never used them. Video cameras were also not used often; ten tourists had never used a stand-alone video camera and four, a mobile phone video camera. Tourist applications were quite unknown. 14 tourists had never used any type of PET and 12 of them had never used a traditional audio guide.

These results confirm that tourist applications are not widespread among users, not even in a sample with an average technological knowledge as the one taking part in the validation. The most popular tourist information sources (used by nearly all tourists) were Internet (20); Web pages of the Destination Management Offices (DMOs) (20); friends (19); and Local Tourist Offices (LTOs) (19). Printed tourist guides were used by nearly 15 people. On the contrary, around half of the tourists never used hotels (8); hotel Web sites (9); social networks (8); specialised forums (7); or travel agencies (9); as sources of tourist information.

Participants used different transportation means to arrive to San Sebastian. Nine of them came by car; four by train; four by bus; and three by other means of transportation. Regarding the use of public transportation while visiting the city, only one tourist did not want to use it. This highlights the importance of providing information about public transportation, which has also been identified by previous studies (Schmidt-Belz et al., 2003; Stroobants, 2006; Beer et al., 2007).

4.5 Results of the Validation

Focusing on the prototype, most tourists (around 80% for most offered functionalities) found it easy to perform different tasks (Fig. 3). Customising the profile, which consisted of updating the score and visit time for each POI, was not considered easy by three tourists; and customising the generated route by two. Only one person found difficulties with the generation of a profile, the generation of a personalised route and the navigation through the map.

Nearly all tourists found the prototype easy to use (19) and considered it was clear what the system expected from them at each step of the interaction (14) (Fig. 3). The time required to create a profile (17) and to generate a route (15) was considered acceptable. Most tourists considered the proposed route sufficiently satisfying (14) and the proposed POIs met their interests (16). This general feedback about the prototype was very good and encourages further research and development.

Participants were quite satisfied with the prototype. Most of them would recommend it (17), and would use it again in San Sebastian (15) or in another city (18). 15 of them would like to use the prototype on their mobile phone. Most of them (18) also considered the prototype was suitable to discover a city and its Cultural Heritage more effectively. Moreover, more than 75% of them (16) perceived a city offering a similar experience as a more technologically advanced city. These results are very positive, as they confirm the perceived utility of the personalised route generation with public transportation is high amongst tourists.

The relationship between the perceived value of the prototype and the amount of money tourists would pay to use it has been diverse. Nearly half of the sample would pay nothing (7) or less than one euro (2) for the prototype.

The remaining tourists would pay between one and three euro (5); between three and five (3); and between five and ten (3). No one would pay more than ten euro. These quantities are in the range of values of the tourist applications available on the market, with both free applications and some applications that usually do not cost more than ten euro.

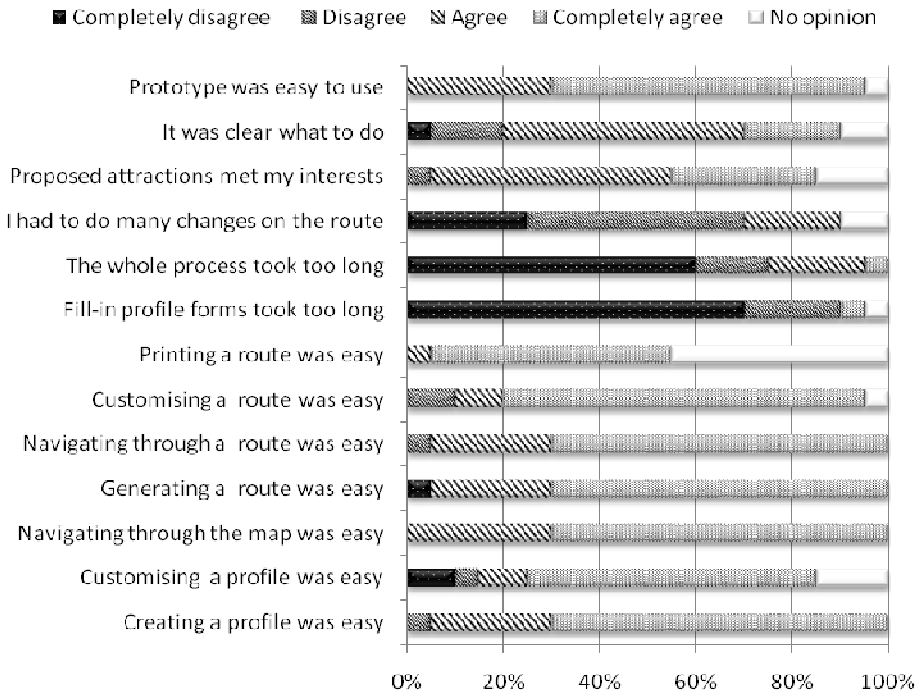


Fig. 3. Opinion about the prototype of the participants of the validation

Participants were asked about their ideal PET (Fig. 4). All tourists considered lists of POIs, public transportation information, and maps as important functionalities. Some tourists were reluctant about some of the functionalities: promotions and coupons (7); practical information (7); shopping guides (4); restaurant guides (1); and door-to-door navigation (1). Again, participants considered both personalised route generation and public transportation important aspects of an ideal PET.

The monetisation of the ideal PET presents differences from the prototype. Nearly half of tourists would pay nothing (6) or less than a euro (3) for it. The remaining tourists would pay between one and three euro (3); between three and five (2); and between five and ten (5). One tourist would pay more than ten euro for an ideal PET. Participants would pay more for the ideal PET, which includes all the functionalities they consider important, than for the prototype, which includes only a subset of these functionalities.

The validation showed that PETs are perceived as interesting tools by tourists. Moreover, both the personalised route generation and the inclusion of public transportation are perceived as valuable functionalities. This encourages the implementation of these functionalities in a real and fully functional PET.

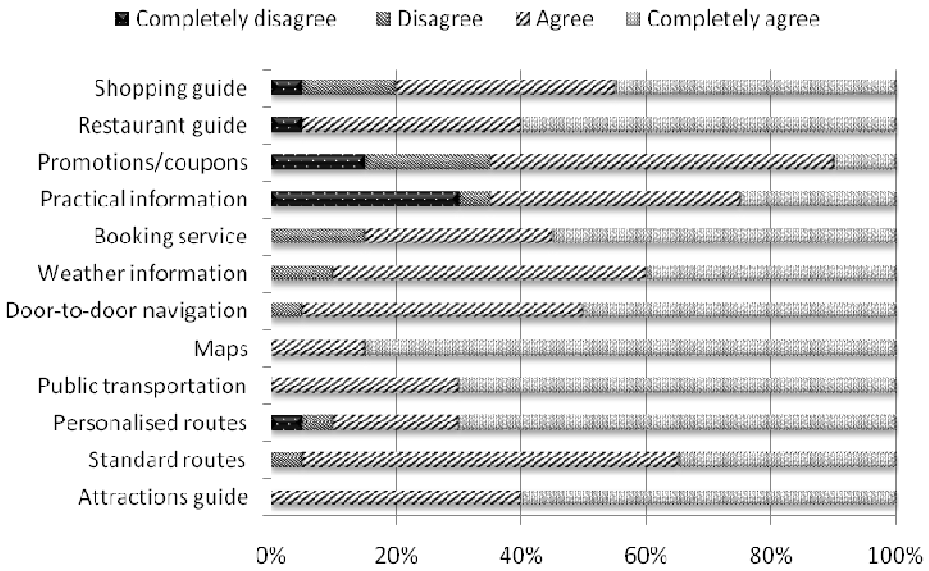


Fig. 4. Opinion about the functionalities of an ideal PET of the participants of the validation

5 Conclusions

This paper presented the validation of a PET prototype focused on the route generation functionality including public transportation. The validation scenario has been the city of San Sebastian, with 200,000 inhabitants; around 50 POIs distributed through the city; and a dense public transportation network.

Although the prototype has three main functionalities (recommendation, personalised route generation and route customisation), the main focus of the prototype and the validation is the route generation functionality. The personalised route generation applies an advanced algorithm to solve problem instances in real-time and including the public transportation of the city. Once tourists obtain a personalised route, they can customise it to better fit their requirements.

The result of the validation with twenty participants has been positive and encourages the inclusion of the personalised route generation functionality in PETs. Most tourists would recommend the prototype (85%) and would use it again on their next visit to San Sebastian (75%). Nearly all of them would like to use it in another city (90%). Tourists found the prototype easy to use and think PETs have many interesting functionalities to offer, including public transportation and personalised routes.

The difficulties of monetising PETs was reflected by the fact that half of the tourists would pay nothing, or less than a euro, even for a perfect PET. However, a third of the tourists would pay more than five euro for it. This situation is reflected by the market, where some PETs are offered for free (including advertisement for example) and some PETs have to be paid for (for example, existing PETs for iPad such as the

DK Eyewitness Paris Travel Guide (itunes.apple.com/us/app/paris-dk-eyewitness/id403266639 [July 10,2011])). As future work, we propose to integrate the personalised route generation functionality, including public transportation, in a fully functional PET in order to validate it in different cities.

References

Bauer, R., Delling, D., Sanders, P., Schieferdecker, D., Schultes, D. & Wagner, D. (2010). Combining hierarchical and goal-directed speed-up techniques for Dijkstra's algorithm. *Journal of Experimental Algorithmics* 15(2): 1-31.

Beer, T., Fuchs, M., Höpken, W., Werthner, H. & Rasinger, J. (2007). CAIPS: A Context-Aware Information Push Service in Tourism. In *Information and Communication Technologies in Tourism 2007* (pp. 129-149). Ljubljana, Slovenia: Springer Wien NewYork.

Cheverst, K., Davies, N., Mitchell, K., Friday, A. & Efstratiou, C. (2000). Developing a context aware electronic tourist guide. In *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '00* (pp. 17-24). New York, USA: ACM Press.

Delling, D., Sanders, P., Schultes, D. & Wagner, D. (2009). Engineering Route Planning Algorithms. In *Algorithmics of Large and Complex Networks, LNCS 5515, volume 2*. Berlin Heidelberg, Springer-Verlag.

Dunlop, M. D., Ptasinski, P. & Morrison, A. (2004). Design and development of Taeneb City Guide - From Paper Maps and Guidebooks to Electronic Guides. In A. Frew (Ed.), *Information and Communication Technologies in Tourism 2004* (pp. 58-64). Cairo, Egypt: Springer Berlin - Heidelberg.

Garcia, A. (2011). *Intelligent Personalised Tourist Route Generation*. PhD Thesis. University of the Basque Country.

Garcia, A., Arbelaitz, O., Linaza, M. T., Vansteenwegen, P. & Souffriau, W. (2010). Personalized Tourist Route Generation. In F. Daniel and F.M. Facca (Eds.), *10th International Conference on Web Engineering 2010, ICWE2010, LNCS 6385* (pp. 486-497). Heidelberg: Springer.

Garcia, A., Linaza, M. T., Arbelaitz, O. & Vansteenwegen, P. (2009). Intelligent Routing System for a Personalised Electronic Tourist Guide. In *Information and Communication Technologies in Tourism 2009* (pp. 185-197). Amsterdam, The Netherlands: Springer Berlin - Heidelberg.

Garcia, A., Vansteenwegen, P., Arbelaitz, O., Souffriau, W. & Linaza, M. T. (2011). Integrating Public Transportation in Personalised Electronic Tourist Guides. *Computers & Operations Research, Special issue: Transport Scheduling*: accepted.

Hagen, K. T., Kramer, Ronny, Hermkes, M., Schumann, B. & Muelle, P. (2005). Semantic matching and heuristic search for a dynamic tour guide. In *Information and Communication Technologies in Tourism* (Vol. 2, pp. 191-206). Innsbruck, Austria: Springer Vienna.

Kabassi, K. (2010). Personalizing recommendations for tourists. *Telematics and Informatics* 27(1): 51-66.

Kenteris, M., Gavalas, D. & Economou, D. (2010). Electronic mobile guides: a survey. *Personal and Ubiquitous Computing*, 15(1): 97-111.

Kramer, R, Modsching, M., ten Hagen, K. & Gretzel, U. (2007). Behavioural Impacts of Mobile Tour Guides. In *Information and Communication Technologies in Tourism 2007* (pp. 109-118). Ljubljana, Slovenia: SpringerVienna.

Maruyama, A., Shibata, N., Murata, Y., Yasumoto, K.,& Ito, M. (2004). A Personal Tourism Navigation System to Support Traveling Multiple Destinations with Time Restrictions. In *18th International Conference on Advanced Information Networking and Applications (AINA)* (Vol. 2, pp. 18-22). Los Alamitos, CA, USA: IEEE Computer Society.

Rasinger, Joerg, Fuchs, M.,& Höpken, W. (2007). Information Search with Mobile Tourist Guides - A Survey of Usage Intention. *Information Technology and Tourism* 9(3): 177-194.

Savelsbergh, M. W. P. (1984). Local search in routing problems with time windows. *Annals of Operations Research* 4(1):285-305.

Schmidt-Belz, B., Laamanen, H., Poslad, S. & Zipf, A. (2003). Location-based mobile tourist services-first user experiences. In *Information and Communication Technologies in Tourism 2003* (pp. 115-123). Ljubljana, Slovenia: Springer Wien New York.

Souffriau, Wouter, Vansteenwegen, Pieter, Vertommen, J., Berghe, G.V. & Oudheusden, Dirk Van. (2008). A personalized tourist trip design algorithm for mobile Tourist guides. *Applied Artificial Intelligence* 22(10): 964-985.

Stroobants, R. (2006). *Mobile Tourist Guides*. Technical report. Leuven, Belgium: Katholieke Universiteit.

Tsiligirides, T. (1984). Heuristic methods applied to orienteering. *Journal of the Operational Research Society* 35(9): 797- 809.

Vansteenwegen, P. & Oudheusden, D. V. (2007). The Mobile Tourist Guide: An OR Opportunity. *Operations Research Insights* 20(3): 21-27.

Vansteenwegen, P., Souffriau, W., Berghe, G. V. & Oudheusden, D. V. (2010). The City Trip Planner: A Tourist Expert System. *Expert Systems with Applications* 38(6): 6540-6546.

Vansteenwegen, P., Souffriau, W., Berghe, G., Vanden & Oudheusden, D. Van. (2009). Iterated Local Search for the Team Orienteering Problem with Time Windows. *Computers & Operations Research* 36(12): 3281-3290.

Acknowledgements

The authors would like to thank the Basque Government for partially funding this work through the neurebide and etourgune projects and to the Centre for Industrial Management of the Katholieke Universiteit Leuven for hosting Ander Garcia as a guest researcher during 2008 and 2009. Finally, authors would like to thank dbus ("Compañía del Tranvía de San Sebastián") for providing real data about the public transportation network of the city.

Mobile Tourism Services and Technology Acceptance in a Mature Domestic Tourism Market: The Case of Switzerland

Adrian Bader[a], Markus Baldauf[a], Sandra Leinert[a], Matthes Fleck[b], and Andreas Liebrich[c]

[a]Lucerne School of Business
{adrian.bader; markus.baldauf; sandra.leinert}@stud.hslu.ch

[b]Lucerne School of Business
Institute of Marketing and Communication
matthes.fleck@hslu.ch

[a]Lucerne School of Business
Institute of Tourism
andreas.liebrich@hslu.ch

Abstract

This paper presents a literature-based and survey-based investigation of a technology acceptance model (TAM) in order to better understand users' acceptance of mobile services in a mature tourism market (Switzerland). Data from a survey (n=588) were used to estimate a conceptual model using structural equation modelling. Findings show that perceived usefulness followed by perceived ease of use, self-efficacy and social influence drive the behavioural intention to use mobile tourism services. The behavioural intention to use mobile tourism services strongly drive mobile tourism services usage, whereas cost decrease the effective usage. Based on the findings the paper concludes that service providers should overcome the negative impact of costs in order to provide the tourists a more convenient and efficient stay in their destination.

Keywords: Technology Acceptance Model, Mobile Services, Swiss Domestic Tourists

1 Introduction

From a historical point of view the tourism industry is considered as an early adopter of new technologies and is in many ways an example of an industry profoundly transformed by technology (Pääjärvi, 2004). The use of information technology in the tourism industry began in the early days of computing, and has become increasingly intense, up to the point where it is probably the strongest driving force for change today (ibid., 2004). However, there are only few early adopters that act as pacemakers (Schegg et al., 2008). The current distribution of mobile devices in Switzerland, Germany and Austria shows that 69% of Germans have a mobile phone with Internet access compared to 68% of Swiss and 76% of Austrians. Standard mobile phones (without mobile Internet access) are out of date and the spread of mobile phones with user friendly operator interface and navigation are dramatically increasing, e.g. iPhone +50%, Smartphones with touchscreen +44%. This paper will present a study undertaken to analyse the technology acceptance of domestic tourists in Switzerland. Thereby, the general usage of mobile tourism services, the development of the conceptual model based on the general technology acceptance model (TAM) is outlined. Moreover, recommendations for the tourism industry will be presented.

2 Literature Review

2.1 Mobile Communication and Tourism Industry

Types of used mobile tourism services. Most people use mobile Internet to check e-mails (44%), to inform about the weather (43%) and news (38%), for navigation (35%) and to track travel itineraries (32%) (Accenture, 2010). Studies about the usage of social networking indicate that only a minority of travellers is using such services. Referring to a study from HSHN (Brözel & Hiller, 2011), travellers spend 28 minutes per day on Twitter compared to 9 minutes per day on Facebook. As 36% of the 1000 interviewed holiday makers took advice from social media sites before making a booking, mostly using trip advisor (two third) and facebook (one third) it can be assumed that this number will rise in the future similar to the mobile device usage (O'Connell, 2011). To summarize, mobile Internet usage increased dramatically in daily life. The Internet is often used for travel purposes.

Trends outlined by experts. The assessment of mobile tourism services is extended by the expected trends according to expert opinions. Three major trends were identified. First, destination managers need to engage with the social online community (Schegg, Liebrich, Scaglione, & Syed, 2008). Second, the trend to mobile apps (Applications for mobile phones) is still important and especially apps about travel information, travel communication and travel distribution are expected to increase in importance in the future. Tourism companies will face the challenge of an `app world` and hence, the development of an app-strategy will become more important for tourism companies and destination managers (Leo, 2010). The third trend refers to social media monitoring (Wolfgang, 2010). Millions of users share thoughts and opinions about travelling online and hence, it is a major source of global knowledge. Destination managers should use this stored knowledge, such as ratings about a hotel, restaurant and transportation, in order to add value to their customers and to re-integrate into the tourism value chain (Wolfgang, 2010).

Additionally, Leo (2010) concluded that smartphones have forced the efforts to develop location-based services, crowd sourcing applications, payment systems and "smart systems" that will be used during the stay in a tourism location. As the literature review outlines the usage of mobile tourism services is increasing. Therefore, a study analysing the technology acceptance of domestic tourists in Switzerland was undertaken.

2.2 Modelling Technology Acceptance

TAM. The technology acceptance model (TAM) is one of the most widely used models to explain user acceptance behaviour. The model is based on social psychology theory in general and the Theory of Reasoned Action (TRA) in particular (Fishbein & Ajzen, 1975) to specify the causal linkages between the two key features: perceived usefulness and perceived ease of use, and users' attitudes, intentions and actual computer adoption behaviour. TAM was originally designed to apply to computer usage behaviour (Davis, Bagozzi, & Warshaw, 1989), but it can be readily extended to apply to any type of technology. Many researchers proposed additional variables in order to provide a stronger model (Legris, Ingham, & Collerette, 2003). Venkatesh and Davis (2000) and Taylor and Todd (1995) proposed an extension,

TAM2 including social influence processes and omitted the variable attitude towards use due to weak predictors of either behavioural intentions or actual system use.

Despite the criticisms, Venkatesh et al. (2003) integrated the elements of TAM, the diffusion of innovation theory, and six other prominent acceptance models to formulate the unified theory of acceptance and use of technology (UTAUT). The UTAUT model sets out to integrate the fragmented theory and research on individual acceptance of information technology into a unified theoretical model found to outperform each of the individual models (Venkatesh, et al., 2003).

As the latest development of TAM, the UTAUT suits for analysing the acceptance of mobile services, the characters of mobile service application are different from other information systems. Therefore, some factors must be included such as cost and speed of connection. For discovering the nature of acceptance of mobile services, some factors in theory of innovation diffusion theory (IDT) and task technology fit (TTF) are considered as well. He & Lu (2007) proposed an integrated mobile service acceptance framework revised from UTAUT, IDT and TTF.

Conceptual Model. The study introduces a conceptual model (Fig. 1) for measuring and predicting user adoption, i.e. technology acceptance of mobile tourism services built upon the convergence and modification of the unified theory of acceptance and use of technology (UTAUT) (Venkatesh, et al., 2003), trust research (Gefen, Karahanna, & Straub, 2003), self-efficacy studies (Thatcher & Perrewe, 2002; Venkatesh & Davis, 1996), and mobile service, in particular, mobile tourism service research conducted in this study, to explain user adoption decisions. The conceptual model as a modified UTAUT with additional variables (self-efficacy, costs, speed of connection, perceived trust) controls for age, gender, education, residence, duration and type of the journey as well as for the types of mobile devices and its access to the services.

2.3 Hypotheses

Perceived usefulness - In terms of mobile tourism services perceived usefulness is understood as the user's perception to the extent that the technology, mobile tourism services, will improve the user's journey, i.e. the user has a perception of how useful the technology while being on a journey. *Perceived ease of use* - Perceived ease of use refers to the degree to which a person thinks that using a certain technology will be free of effort (Davis, et al., 1989). *Self-efficacy* - Perceived self-efficacy refers to "beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments" (Bandura, 1997). *Social influence* - Venkatesh, Morris, Davis & Davis (2003) defined social influence as the degree to which an individual perceives that important other persons believe he or she should use the new system. In the conceptual model social influence is implemented with the assumption that determined perceptions are generally communicated through interpersonal channels of communication – important as well as less important peers.

Literature indicates that women are more sensitive to others' opinions. Therefore, it was found that social influence tends to be more salient when forming an intention to use new technology (Miller, 1986; Venkatesh, Morris, & Ackerman, 2000). Rhodes' (1983) review of age effects concluded that affiliation needs increase with age,

suggesting that older people are more likely to place increased salience on social influences, i.e. subjective norms (Venkatesh, et al., 2003).

Costs - According to behavioural decision theory, the cost-benefit pattern is significant to both perceived usefulness and ease of use. In terms of mobile tourism services the costs refer to equipment costs, access cost, and possible transaction fees (Constantinides, 2002). *Speed of connection* - The speed of connection of mobile tourism services describes the extent to which service accessibility is fast and reliable according to customer requirement. The speed of connection has significant effects on the usage of mobile tourism services. *Perceived trust* - According to Jones (2002) trust is somebody's assurance that he or she may predict actions of a third party, may rely upon his actions, and that his actions will follow a foreseeable pattern in the future, especially under risky circumstances and when no explicit guaranty is provided (Serenko & Bontis, 2004). Trust is seen as a major enabler of wireless interaction involving the exchange of personal and private information. It is important for a mobile tourism service user to believe in the integrity, credibility, security, authenticity, reliability, and honesty of a service provider (Serenko & Bontis, 2004).

Control variables - The conceptual model also takes into account possible differences between different mobile devices and accesses to the mobile tourism services. To grasp moderations of the conceptual model's constructs the study controls for gender, education as well as age. The hypotheses resulting from the literature review are demonstrated in table 1.

Table 1. Hypotheses

H1	Perceived usefulness of a mobile tourism services has a positive direct effect on behavioural usage intentions toward mobile tourism services.
H2	Perceived ease of use of a mobile tourism services has a positive direct effect on the perceived usefulness of mobile tourism services.
H3	Perceived ease of use of a mobile tourism services has a positive direct effect on behavioural usage intentions toward mobile tourism services.
H4	The degree of an individual's self-efficacy in the domain of mobile devices will have a positive direct effect on the perceived ease of use of mobile tourism services.
H5	Self-efficacy has a positive direct effect on the intention to use mobile tourism services.
H6	The positive influence of social influence on behavioural intention to use mobile tourism services will be moderated by gender and age, such that the effect will be stronger for women, particularly older women.
H7	Cost has a negative direct effect on the usage of mobile tourism services.
H8	Speed of connection has a positive direct effect on the behavioural intention to use mobile tourism services
H9	Perceived trust towards mobile tourism services has a positive direct effect on behavioural usage intentions towards mobile tourism services.

3 Methodology

The primary goal of this study is to better understand users' acceptance of mobile services in domestic tourism with help of a model estimation resulting in recommendations for the mobile tourism services market. As a preliminary step to the

survey an expert interview using a semi-structured questionnaire took place in order to update the literature with the latest information about mobile tourism services.

3.1 Sample Selection

Due to the unknown basic population of incoming and domestic tourism in Switzerland it was not possible to do a random sampling. Therefore, a non-probability sampling was used in form of a quota sampling asking tourists aged 15 and above. Although, the sample of the study represents a wide range of interviewees aged 17 to 67, the young sample indicates a biased distribution. It might be a sub-sample of young people and it can be questioned whether it is representative or not. On a cautionary note, it is emphasized that the current sample is not a student sample (Stevens, 2011). However, as mobile tourism services are mostly used by young people, the sub-sample is a valid representation of this young generation (Tan, Goh, Theng, & Foo, 2007).

As the study will focus on the technology acceptance of the Swiss domestic tourism market, the definition of a tourist is given according to BFS (Swiss Statistics, Federal Statistical Office - Baumann & Schiess, 2008). Thus, tourists are travellers who sojourn out of their daily life - i.e. the principal residence, the work and study space as well as all places lying close-by them. Those tourists visiting a place in Switzerland can be further differentiated using two criteria: the purpose and the duration of the travel. As a result, the study distinguishes three types of domestic tourists: domestic tourist on a private journey with accommodation, domestic tourist on a private journey without accommodation, domestic tourist on a business journey.

3.2 Data Collection Method

The survey has been undertaken using quantitative research methods. A standardized questionnaire was provided. The questionnaire was pre-tested within a small sample size of ten people in order to identify gasps and to modify questions. Afterwards, the survey was distributed within the researchers and the University of Applied Science Lucerne, Ms Sc BA Tourism class 2012's network. Furtheron, the researchers collaborated with the University of Lucerne and several hotels in Lucerne, Bern and Zurich and environment to meet tourists on the go.

4 Results

The descriptive results regarding sociographic characteristics can be withdrawn from table 2.

Table 2. Sample – sociographic characteristics

Type of journey	N	Age	Female
private journey without overnight stay	291	25.39	43%
private journey with overnight stay	251	25.35	39%
business journey	46	26.63	41%
Total	588	25.76	41%

Analysing the usage of mobile tourism services points out that information services (47.2%) and communication services (55.83%) are mostly used by the domestic tourists. In contrary, booking services (12.88%), social share (20.16%), support tools (14.41%) and location guides (26.81%) are less used by the domestic tourists. Most of the interviewees (55.57%) used a smartphone on their recent journeys within Switzerland, whereas 54.13% used a simple mobile phone and 45.1% a notebook. Thus, they are the most used mobile devices on recent journeys of domestic tourists within Switzerland. Evaluating the access, most respondents got access via a Swiss provider (78.13%) and via a free Wi-Fi (53.23%). Additionally, the most accepted channels for sharing and receiving information of mobile tourism are SMS (83.17%), followed by telephone (75.13%), mobile Internet (70.57%) and Wi-Fi (68.90%).

4.1 Qualitative Feedbacks from Respondents

Beside the analysis of the quantitative data, a short analysis of the 37 interviewees (6.3%) stating their personal opinion about mobile tourism services was undertaken. Most respondents would prefer easier applications in order to know how to best use them. Moreover, most domestic tourists argued that costs are a major barrier of using mobile tourism services. On the contrary some respondents stated that they do not perceive mobile tourism services as useful, because the world is moving fast enough and those services make people more helpless and lazy.

Moreover, arguments against mobile tourism services are the loss of the social interaction with people during travelling and the trust in oneself to find a certain place without a mobile device. Additionally, the decrease of service quality and the reduction of working places were assumed negative outcomes of mobile tourism services.

4.2 Model Estimation

The conceptual model is estimated with Structural Equation Modelling (SEM) using the Maximum Likelihood (ML) algorithm. The General Least Square estimates differ only in the coefficients values. Figure 1 shows that the conceptual model has a relatively small fit. In particular, RMSEA is not satisfactory for this model. According to Browne and Cudeck (1993), a RMSEA less than 0.08 is acceptable. Similarly, the RCI and NFI fit indexes are not close enough to 1 to be considered as a good model fit. However, the model fit cannot be increased significantly by a stepwise integration of the constructs. With respect to the suggested conceptual model, several conclusions can be made.

First, the behavioural intention to use mobile tourism services has a strong positive impact on its use. Although, costs do not influence the intention to use mobile services they are reducing the service use significantly. The intention to use mobile tourism services is driven mostly by usefulness. Further on, it is moderated by self-efficacy, i.e. ease of use and social influence. Neither the speed of connection nor trust show significant coefficients.

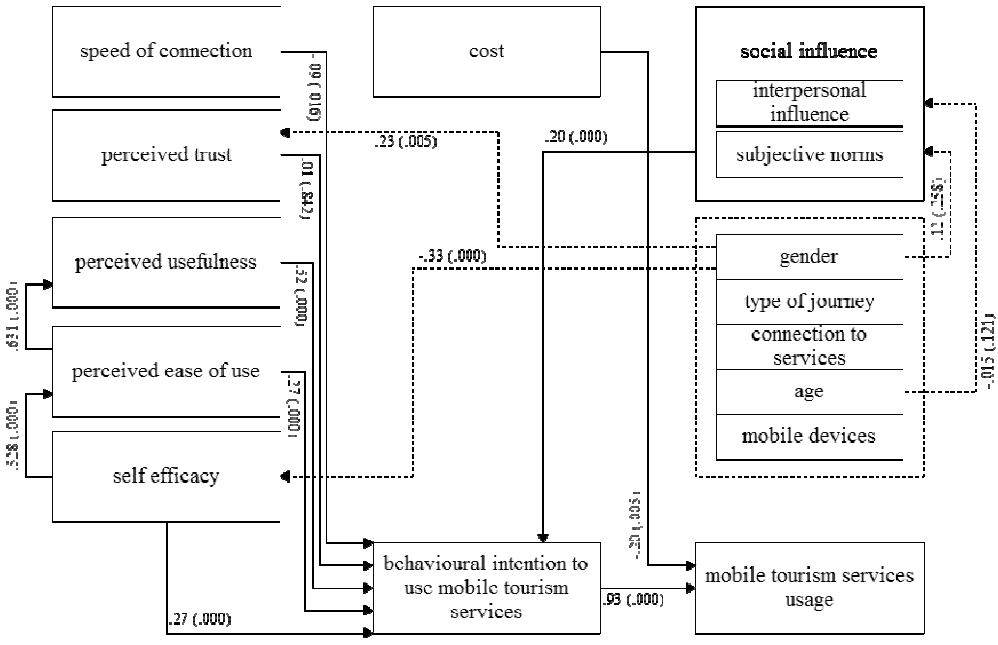
It can be confirmed, that perceived usefulness of a mobile tourism services have a positive direct effect on behavioural usage intentions toward mobile tourism services (H1). Thus, an increase in usefulness affects the intention to use services by 0.5 - ceteris paribus. The perceived amount of effort required to be able using the system

has a moderate impact on the intention to use mobile tourism services. Moreover, ease of use shows even a stronger significant impact of 0.631 on usefulness. These findings underline the stated relationship of Davis and Arbor (1989), that the ease of use might out weight the usefulness of the technology. Additionally, the statement of Schotter (2011) saying that a lack of application knowledge is a major barrier to the usage of mobile tourism services seems to be valid. Thus, perceived ease of use of mobile tourism services having a positive direct effect on the perceived usefulness (H2) and on behavioural usage intentions (H3) of mobile tourism services, can be confirmed.

Venkatesh & Davis (1996) suggested that an individual's perception of a particular system use is anchored to his or her level of computer self-efficacy. Accordingly, the study finds individuals with a high level of self-efficacy in terms of mobile devices form more positive perceptions of the 'ease of use' of mobile tourism services. Furthermore, people with a high self-efficacy judgment in the domain of mobile devices are more likely to use mobile tourism services. Hence, it can be confirmed that the degree of an individual's self-efficacy in the domain of mobile devices has a positive direct effect on the perceived ease of use of mobile tourism services (H4). Similarly, one cannot reject that self-efficacy has a positive direct effect on the intention to use mobile tourism services (H5).

Social influence as a direct determinant of the behavioural intention to use mobile services exerts a moderate significant influence. Thereby, the individual`s behaviour is influenced by the way in which they believe others will view them as a result of having used the technology (Venkatesh, et al., 2003). In contrast the study cannot conclude that the influence of social influence on behavioural intention is moderated by gender and age (H6).

Gender as well as age does not have a significant impact on social influence. A fast and reliable access to the services has neither significant effects on the intention to use nor the use itself. Controlling for different types of access as well as different providers the influence still is non-significant. Thus, the speed of connection has no positive direct effect on the behavioural intention to use mobile tourism services (H8). Therefore, the hypothesis has to be rejected. Hence it is assumed that the mobile tourism services available to the customers have taken current limitations into consideration and are developed according to the existing data transaction capacity. Surprisingly, trust does not show any effects on the use constructs. In particular, against the findings of Gefen, et al. (2003) and Wang, et al. (2003), claiming that trust is reducing individuals concerns and heightening their decision to use the system, the model estimation cannot depict a significant influence on the intention to use mobile tourism services. Therefore, the hypothesis, that perceived trust towards mobile tourism services have a positive direct effect on behavioural usage intentions towards mobile tourism services (H9) has to be rejected. A possible explanation is, that people trust their providers or the OS-platforms of their devices so much, that they do not care about the risks of the different mobile service applications itself. Monetary costs are essential in mobile tourism services settings. Higher prices for equipment, access and/or transactions are reducing the service use significantly. Like conjectured by expert interviews in advance and the stated hypothesis (H7), costs are one of the major factors influencing the technology acceptance of mobile tourism services.



Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.143	.131	.155	.000
Independence model	.200	.190	.209	.000

Fig. 1. Maximum Likelihood estimate of the conceptual model

Control variables. Gender differs in self-efficacy and perceived trust towards information technology respectively mobile tourism services. This finding is consistent with earlier studies (Broos, 2005). Even though females perceived themselves as more competent in using computers and have a significant higher trust level towards mobile tourism services, they experience a higher computer anxiety relative to males (Wong & Hanafi, 2007). According to Liaw's study (2002) it is not surprising because the study also indicated that males show a more positive perception towards computers and Internet technologies than females.

As the collected sample data does not vary enough in age, the model detects no effects. The sample consists of mostly young people aged between 20 and 30 years. To control for the different type of journeys, some separate models were estimated.

The technology acceptance, i.e. the behaviour towards mobile tourism services for tourists on private journey with and without overnight stays is representative for the whole sample (N=542 out of N=588). The estimated model for business travellers (N=46) shows a poor fit. In addition, most of the coefficients are insignificant. Integrating business travel dummies into the conceptual model confirms that the sample of business travellers is too small to gain robust information out of it.

Including the dummies in the conceptual model it can be concluded, that there is a significant higher intention to use and effective use the more functions a mobile device provides. In case of the effective use of mobile tourism services this result seems to be valid, as a person traveling with simple phones cannot access most of the services due to the limitations of the device. More interesting is the fact, that differences in the ownership of mobile devices have an impact on the intention to use the services. People with Internet or smart phones are more likely to use mobile tourism services. As there is a strong correlation between the number of functions of mobile devices and self-efficacy toward information technology, it is assumed that the effect is moderated by self-efficacy. In other words, only 34.4% of the simple phone users are concentrated in the three highest self-efficacy categories. In contrary 60.2% of the smart phone users are to be localized in those categories.

The model estimations differ in case of people accessing mobile tourism services through a foreign provider (N=30) and paid Wi-Fi (N=185/N=55 if dropped out persons with a Swiss provider). Most of the coefficients become insignificant if separate model estimation is conducted with that small sample size. Looking at free Wi-Fi, Swiss provider or offline access the people show behaviour according to the above analysed conceptual model. If the access is included in the conceptual model all predict a significant higher intention to use mobile tourism services expect for foreign provider. There is an insignificant negative effect on the intention to use mobile tourism services if the user has a contract with a foreign provider. The same accounts for the effective use of mobile tourism services.

The effect of foreign provider is insignificant. In particular, it is hard to win arguments to distinguish between different alternatives to access the services, because of the small sample sizes. Furthermore, there are no signs for a correlation between the different access options and the connection speed respectively the costs. The collected data is too broad to grasp these fine differences. As there are multiple responses allowed, when asking for their access, the database cannot allocate the use of mobile tourism services on specific access options, the interviewees mentioned connection speed and the costs among other factors.

5 Discussion and Conclusions

Channel acceptance - the analysis shows that different standards should become one for indirect and one for direct communication. The direct communication between devices also shows great potential, as it is an economic way of exchanging data. Near Field Communication (NFC) is an emerging technology for mobile devices making them a hotel key, a credit card and a bus ticket at the same time. While Wi-Fi hot spots are a cheap and useful solution mobile telephony can still be very costly. Wi-Fi might be seen as substitute to mobile data communication via telephony to save cost. When the LTE generation (4G - 4th generation mobile phone network standard) conquers the market, the mobile device can be considered as always online via the telephone network making Wi-Fi hotspots obsolete.

Mobile device usage - Like the humans five congenital natural senses the use of the sixth sense afforded by mobile devices should be as natural as possible in order to be seamlessly integrated in reality. Mobile devices should be intuitively understandable

and operable. Thus, human behaviour should determine the design of mobile devices. Not mobile devices should form human behaviour moreover, it is the information they are transmitting. *Service usage* - Services should interact with each other and finally bring the entire touristic world in one customer`s mobile device. Further the services itself have to be intuitively operable, i.e. easy to use. The study shows that the usage of the service categories booking services, support tools and social share is rather low. Thus, services have to move closer and must interact with each other. This interaction must be integrated and for the user almost invisible. But to fully reach this stadium of interaction the industry must start to develop not only for its own environment (e.g. Lufthansa mobile world) but also interact with other service providers (e.g. Accor cooperation with TripAdvisor for hotel rating). *Access* - As all of our five senses constantly attending us, the sixth sense should be too.

Thus, an area-wide net coverage at any time at most advantageous prices possible is a necessity. Given the fact that network systems with sunk costs evolve in a natural monopoly, there should be a single regulated network system constructor meeting the demand of its users and creating competition in the access provider market. Still policies have to establish a more competitive and dynamic telecommunication market for lower, more standardized costs within Switzerland and ensure coordination with the European Union. Although the market is facing steadily decreasing prices, competition is low. The entry barriers for access provider should be reduced by decoupling the network infrastructure ownership from access providers.

5.1 Recommendations for the Tourism Industry in the Age of Tourism 2.0.

The use of mobile tourism services is positively moderated by social influence. Therefore, make it social. The building of social communities as standalones or in existing social networks will result in higher usage of mobile tourism services. Furthermore overcome the negative impact of cost on the usage of mobile tourism services. Thereby, the service providers of touristic regions/destinations should form networks or strategic alliances to setup free public Wi-Fi networks. Moreover, service providers should minimize transaction costs of services and deliver information (content) for free. One should give people incentives not only to use the content but also to generate useful content on their own (Web 2.0).

Additionally, mobile tourism services which can be used offline could be developed. The relevant content should be downloadable and in interaction with the GPS generating useful information for the customer. Service providers should also consider the existing gender difference in information technology usage. As the operating systems and the mobile devices become increasingly diverse, the services should be programmed as browser applications and not as operating system specific apps. The most important needs of a customer demanding a mobile tourism service is to save time, being more efficient and effective. Therefore, the customer`s need should be the focus in developing mobile tourism services.

5.2 Recommendations for the Mobile Tourism Service Environment.

The technology provider should enforce innovative, continuous improvement of mobile devices - keep them easy to handle - especially smart phones due to the dramatic increase in their usage. Policies have to establish a more competitive and

dynamic telecommunication market for lower, more standardized and more transparent costs within Switzerland and ensure coordination with the European Union. Transparency in a market with just a few players alone can't evoke competition, but rather facilitating tacit collusions. As this study was a first step on analysing the technology acceptance of Swiss tourists, further research can be done about the implementation of given recommendations by tourism organisations and companies.

References

Accenture. (2010). *Mobile Web Watch - Studie 2010*.

Bandura, A. (1997). *Self-efficacy: the exercise of control*. W.H. Freeman.

Baumann, T. & Schiess, U. (2008). *Satellitenkonto Tourismus der Schweiz, 2001 und 2005, Grundlagen, Methodik und Ergebnisse*. Neuchâtel.

Broos, A. (2005). Gender and information and communication technologies (ICT) anxiety: Male self-assurance and female hesitation. *Cyberpsychology & Behavior* 8(1): 21-31.

Browne, M. W. & Cudeck, R. (1993). Alternative ways of assessing model fit. *Testing structural equation models* 154*:* 136–162.

Brözel, M. A. C. & Hiller, C. A. (2011). N*icht ohne mein Smartphone! Studienergebnisse zur Nutzung mobiler Technologien im Urlaub.* Paper presented at the ITB.

Constantinides, E. (2002). The 4S Web-marketing mix model. *Electronic Commerce Research and Applications* 1(1): 57-76.

Davis, F. D. & Arbor, A. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *Mis Quarterly* 13(3): 319-340.

Davis, F. D., Bagozzi, R. P. & Warshaw, P. R. (1989). User acceptance of computer technology: a comparison of two theoretical models. *Management Science* 35(8): 982-1003.

Fishbein, M. & Ajzen, I. (1975). *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research*. California: Addison-Wesley Pub. Co.

Gefen, D., Karahanna, E. & Straub, D. W. (2003). Trust and TAM in online shopping: An integrated model. [Review]. *Mis Quarterly* 27(1): 51-90.

He, D. & Lu, Y. (2007). An Integrated Framework for Mobile Business Acceptance. *Sixth Wuhan International Conference on eBusiness,* 14.

Jones, A. J. I. (2002). On the concept of trust. *Decision Support Systems* 33(3): 225-232.

Legris, P., Ingham, J. & Collerette, P. (2003). Why do people use information technology? A critical review of the technology acceptance model. *Information & management* 40(3): 191-204.

Leo, H. (2010). *Trends in Service Innovation enhanced by mobile services in the field of tourism in rural and mountain areas*. Paper presented at the How can Service Innovations support Sustainable Tourism in Rural Regions?

Liaw, S. S. (2002). An Internet survey for perceptions of computers and the World Wide Web: relationship, prediction, and difference. *Computers in Human Behavior* 18(1): 17-35.

Miller, J. B. (1986). *Toward a new psychology of women*. Beacon Pr.

O'Connell, S. (2011). *A round-up of the debate around the value of social networking.* Retrieved May 31, 2011, from http://www.global-blue.com/merchant-services/corporate/news/social-networking-should-you-take-notice/

Pääjärvi, M. (2004). *Future Mobile Data Services for Tourism*. Unpublished Examensarbhete. Luleå University of Technology.

Rhodes, S. R. (1983). Age-Related Differences in Work Attitudes and Behavior: A Review and Conceptual Analysis. *Psychological Bulletin* 93(2): 328-367.

Schegg, R., Liebrich, A., Scaglione, M. & Syed, A. (2008). An Exploratory Field Study of Web 2.0 in Tourism. Proceedings of the ENTER Conference 2008 - "*Information and Communication Technologies in Tourism 2008*". Innsbruck, Austria, Springer, Wien.

Schotter, C. (2011). *m-tourism - ein Leitfaden für Destinationen*. Paper presented at the ITB 2011.

Serenko, A. & Bontis, N. (2004). A model of user adoption of mobile portals. *Quarterly Journal of Electronic Commerce* 4(1): 69-98.

Stevens, C. K. (2011). Questions to Consoder when Selecting Student Samples. *Journal of Supply Chain Management* 47(3): 19-21.

Tan, E. M.-Y., Goh, D. H.-L., Theng, Y.-L. & Foo, S. (2007). *An analysis of services for the mobile tourist*. Paper presented at the Proceedings of the 4th international conference on mobile technology, applications, and systems and the 1st international symposium on Computer human interaction in mobile technology.

Thatcher, J. B. & Perrewe, P. L. (2002). An empirical examination of individual traits as antecedents to computer anxiety and computer self-efficacy. *Mis Quarterly* 26(4): 381-396.

Venkatesh, V. & Davis, F. D. (1996). A model of the antecedents of perceived ease of use: Development and test. *Decision Sciences* 27(3): 451-481.

Venkatesh, V. & Davis, F. D. (2000). A theoretical extension of the Technology Acceptance Model: Four longitudinal field studies. *Management Science* 46(2): 186-204.

Venkatesh, V., Morris, M. G. & Ackerman, P. L. (2000). A longitudinal field investigation of gender differences in individual technology adoption decision-making processes. *Organizational Behavior and Human Decision Processes* 83(1): 33-60.

Venkatesh, V., Morris, M. G., Davis, G. B. & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *Mis Quarterly* 27(3): 425-478.

Wang, Y. S., Wang, Y. M., Lin, H. H. & Tang, T. I. (2003). Determinants of user acceptance of Internet banking: an empirical study. *International Journal of Service Industry Management* 14(5): 501-519.

Wolfgang. (2010). *Tourism 2010: three web trends for destination management organizations.* Retrieved from http://www.web-cubed.com/2010/01/14/tourism-2010-three-web-trends-for-destination-management-organizations/

Wong, S. L. & Hanafi, A. (2007). Gender differences in attitudes towards information technology among Malaysian student teachers: A case study at Universiti Putra Malaysia. *Subscription Prices and Ordering Information* 158.

The New Landscape of Travel: A Comprehensive Analysis of Smartphone Apps

Dan Wang[a], and
Zheng Xiang[b]

[a] School of Tourism and Hospitality Management
Temple University, USA
dan.wang@temple.edu

[b] School of Merchandising and Hospitality Management
University of North Texas, USA
philxz@gmail.com

Abstract

Due to the increasing popularity of smartphones and their apps, mobile technologies have been effectively changed the nature of travel. This study employed a data mining approach to classify the types of information services and design features of travel-related iPhone apps as well as users' evaluation and reviews of these apps. The results identify the range of iPhone apps from both the developers' and users' perspectives. This study documents the state-of-the-art of mobile technologies and serves as an important foundation for understanding the emerging technological environment that will cause substantial change in travel and tourism.

Keywords: Mobile technology; smartphone; data mining; mobile applications.

1. Introduction

Imagine a traveller driving on an Interstate highway in America. With the assistance of a GPS-enabled iPhone, he/she accesses Google Map for driving directions. Rock music is played through the car's stereo system with the streaming music coming from the Spotify app on the phone. With a little help from the embedded search engine, he/she locates a restaurant of one of his favourite brands 20 miles away along the highway; and, fifteen minutes later arrives at the restaurant and instantly checks in through the Foursquare app equipped with the location-awareness capabilities. His/her status is immediately updated so that he/she can now interact with friends through various social networks, such as Facebook and Twitter. Although this scene by no means represents a complete picture of the "landscape" of mobile technologies supporting travel, this scenario clearly illustrates what is becoming increasingly common among today's travellers on the road. There are indications that mobile technologies, especially through the emergence of smartphones and software applications (apps), will become the next wave of innovation that drives travel and tourism (Wang, Park, & Fesenmaier, 2011). With power-efficient processors, modern operating systems, broadband Internet access, large viewing screens, user friendly interfaces, pervasive computing capabilities, as well as productivity-enhancing apps, smartphones are being adopted by the consumer market at a rapid speed (Chang, Chen, & Zhou, 2009; Roussos, Marsh, & Maglavera, 2005; Want, 2009). This collection of mobile technologies offers a wide range of possibilities to support travellers in their use of information and communications home and away. Particularly, smartphones offer increased connectivity, communication, content

consumption, and content creation. As such, mobile technologies have even been touted as the catalyst for the emergence of a new generation of modern tourists (Gretzel, 2010; Gretzel & Jamal, 2009). The goal of this study is to investigate the state-of-the-art of mobile software applications in order to address two questions: 1) what information services and design features do smartphone apps possess to support travel? and, 2) how do users evaluate them? The following sections provide a review of related literature, a description of methods, study results, and implications and conclusions based upon the findings.

2. Research Background

Recent studies indicate that mobile technologies have the capabilities to satisfy users' entertainment and spontaneous needs, help fulfil one's efficiency desires, and assist in making time-critical arrangements (Anckar & D'Incau, 2002; Brown & Charlmers, 2003; Gretzel, 2010; O'Brien & Burmeister, 2003; Rasinger, Fuchs, & Höpken, 2007). Importantly, Tussyadiah and Fesenmaier (2009) and Wang, Park, and Fesenmaier (2011) found that smartphones can mediate both the behavioural and psychological dimensions of the touristic experience by facilitating information search, processing, and sharing, by enabling a traveller to learn about new travel opportunities and get to know better a destination, and by sharing photos and other social activities at any time during the trip. All the changes and benefits brought out by mobile technologies to users are supported by diversified information services and innovative designs built into mobile phones. In the tourism context, mobile technologies have been developed primarily as electronic tourist guides to support information access and decision making during the travel process (Brown & Charlmers, 2003; Poslad, Laamanen, Malaka, Nick, Buckle, & Zipf, 2001). The main purpose of these systems is to create services that support efficient, intelligent, and accessible communication for travellers (Grün, Pröll, Werthner, Retschitzegger, & Schwinger, 2008; Kenteris, Gavalas, & Economou, 2009).

With enhanced multimedia capabilities and access to the Internet, mobile technologies can create "augmented realities" that enable rich contents to be viewed, documented, and shared among travellers (Lashkari, Parhizkar, & Mohamedali, 2010). Also, social computing combined with location-based technology can create messaging systems that deliver user-generated content to the social circle at the users' physical locations (Carlsson, Walden, & Yang, 2008). With the advent of the smartphone era and the emergence of open source software systems, information services available on the mobile platform have been substantially expanded with the support of continuously updated designs. This growth can lead to more changes in one's life (Katz, 2006). As such, it is important to develop a comprehensive and systematic understanding about the landscape of information services on the mobile platform. It appears that, however, only a handful of studies have been conducted to understand the range of travel-related capabilities offered through mobile technology. For example, Grün, Pröll, Werthner, Retschitzegger, & Schwinger (2008) analyze the system of mobile tourist services from the dimensions of information services and design features. Kenteris & Gavalas (2011) evaluate the design features of mobile tour guides using four dimensions: 1) information models such as personalized content, collaborative filtering, context awareness, and automatic updating of content; 2)

positioning/mapping; 3) network infrastructure such as cognitive capabilities and adaptability; and 4) input/output modalities. Unfortunately, these studies are limited by in terms of their depth and breadth of analysis. As such, there is a considerable gap in our understanding of the information services that are now provided by mobile systems, particularly represented by today's smartphones. Building upon this literature, this study attempts to systematically examine the range of information services for travellers on the smartphone platform for a more comprehensive understanding of emerging mobile technologies.

3. Methods

This study adopted a content analysis approach with the aim to extract and describe the information and design features of travel-related iPhone apps as well as users' evaluation and reviews of these apps.

3.1. Data Collection

The iPhone, as part of the first generation of smartphones, is one of the most influential mobile innovations (Want, 2009). To date, over 33 million iPhones have been sold worldwide (Dala-Ali, Lloyd, & Al-Abed, 2011). Correspondingly, the iTune app store, the online store used to sell mobile applications running on the iPhone platform, has grown to be one of the largest websites for downloading mobile applications. Thus, this study focuses attention on iPhone apps that are currently available through the iTunes store. Data on the use of iPhone apps was collected through a program designed to "scrape" descriptive information about each app (e.g., release date, version, developer, and function, etc), customer reviews, and associated rating scores (from 1 to 5) posted for the specific app from the iTune app store. By the end of July 2010, there were 19,072 apps included in the travel category. Only apps with ratings and customer reviews are included in this study; thus, data from 4,954 apps, which represented approximately 25 percent of all apps in the travel category, were collected. Preliminary analysis of the ratings indicates that among the 4,954 apps, 86 percent of the customer reviews focused on the first 300 apps (2,450,802 out of 2,835,061 ratings). Thus, the top 300 apps were included for further investigation. These 300 apps are associated with 69,437 reviews, which is 67% of the total number (106,105) of reviews in the travel section in the iTune app store.

3.2. Data Analysis

Content analysis was used to identify the app features and to summarize customer reviews of the respective smartphone apps. Wordstat, a computer text analysis program for text mining, was used to analyze two sets of text corpus (developers' descriptions about the apps functions and customer reviews). Data analysis followed a three step process: classification, summarization of functions and design features, and identification of consumers' preferences. First, the top 300 apps were classified according to the major information services they provide, which were then summarized into categories. The text corpus of developers' descriptions, which describe the main functions of apps, was used for this purpose. Specifically, the "automated text categorization" function in Wordstat was used to assist the classification of apps. Automated text categorization is a supervised machine-learning

task by which new documents are classified into one or several predefined category labels based on an inductive learning process (Sebastiani, 2002). The Wordstat program classifies the apps based on naive Bayesian classifiers whereby the classifiers were first obtained by using a set of discriminate functions and then estimated using the relevant probabilities derived from a set of previously classified documents (called "training dataset", including 50 apps in this study) (Domingos & Pazzani, 1997). The automated text categorization based on naive Bayes algorithm is appropriate for this study, because the terms describing the information services of apps (e.g. hotels, restaurants, and flights) are useful in differentiating the apps. In order to ensure the accuracy of classification, the results were reviewed by the authors. Twentytwo out of 250 apps (7%) were deemed misclassified and were manually assigned to appropriate categories.

Second, the information services and design features of each category of apps were summarized using Wordstat. Although one category of apps indicates a main kind of information service (e.g. flight searching), each app can provide a variety of information services. As such, in order to reflect the range of services provided, the phrase finder in the Wordstat was used to scan the textual corpus of 300 developers' descriptions and to identify the phrases and idioms illustrating the functions of each app. Most of developers' descriptions are relatively structured, which is appropriate for phrase analysis (Feldman & Sanger, 2007).

For this study, the length of the phrase was set to range between two to five words, and the short phrases that are part of larger ones were automatically removed from the phrase list. Also, the phrases that are literally the same (e.g. live flight status and live flight tracker) were combined by hand. Furthermore, the phrases describing information services and design features of the apps were differentiated by manual coding. The design features were coded according to the framework proposed by Kenteris and Gavalas (2011). Two authors coded the phrases list separately and inter coder reliability (Krippendorff, 2004) was checked. Third, customer reviews in each category were analyzed to identify consumer preferences. Specifically, the text corpus of customer reviews within each category was processed by Wordstat for cluster analysis. Thus, the customers' opinions towards apps were identified by the keywords in each cluster and indicated by the relationships between the words (the patterns of words co-occurrence). Furthermore, crosstab analysis was conducted for words-clusters and the rating scores (from 1(worst) to 5 (best)) and chi-square values were calculated. Thus, features that customers like (with ratings 4 or 5) can be differentiated from those customers do not like (with ratings 1 or 2).

4. Results

4.1. Classification of Apps and Information Services

The most popular travel-related smartphone apps in the iTune app store are grouped into eleven categories based on the major information services they provide (Table 1). Also, by applying the phrase finder function of Wordstat, the details of information service in each category are identified, and the unique information services provided in each category of apps are listed in Table 2. The following paragraphs illustrate each category and the information services they provide.

As can be seen, "Single City Destination Guide" is the biggest category in terms of number of apps (63 apps, 21%). Each app in this category is a mobile tour guide for a particular city (i.e. New York, Chicago) which usually provides a wide range of classified information about that city such as attractions, transportation, restaurants, and shows etc. All 63 apps in this category are about the metropolitan city, and 30 of them concentrate on seven cities in the United States (New York, Chicago, Boston, Seattle, Washington D.C., Las Vegas, and San Francisco) and the remaining 33 are about 21 cities in four continents (North America, Europe, Asia, and Oceania).

Table 1. Types of Travel-Related iPhone Apps

Category	ID	Definition	N	PCNT	Examples
Single city destination guide	F	Mobile tour guide for a particular city (i.e. New York, Chicago)	63	21.0%	NYC Way; DC Going Out Guide; London Tube.
Online travel agency	C	Apps for searching and reserving transportation tickets/hotels/cars	35	11.7%	Travelocity; Rental Car Sites; Hampton.
Language assistant	A	Apps for translation	34	11.3%	Language Translator; iTranslate it; Speak Spanish
Flight manager	B	Apps for searching, tracking and booking flights	34	11.3%	TripCase; Flight Sites; TripTracker.
Theme park & Resort guide	D	Apps for providing travel tips within theme parks or resorts such as Walt Disneyland theme park	31	10.3%	Disneyland Wait Times; Ride Hopper; Hersheypark
Facilitators	K	Apps for providing quick facts such as Wi-Fi spot, cheap gas stations, local times.	30	10.0%	Wi-Fi Finder; CheapGas; AccuTip.
Multiple city destination guide	G	Mobile tour guide for multiple cities	26	8.7%	Citysearch; GayCities; Goby.
Food finder	E	Apps for searching restaurants	24	8.0%	Urbanspoon; Happy Hours, LocalEats.
Entertainment	I	Apps for the purpose of fun	14	4.7%	Postman; Trip Journal; Virtual Tourist.
Live camera	H	Apps for viewing live situations in other places through webcams	5	1.7%	Live Cams; TraffiCam SG; WorldView.
Currency converter	J	Apps for calculating exchange rates	4	1.3%	Currency; Quid; CurCon

Note: Total N=300

Twenty-nine apps in this category are paid apps and their prices range from $0.99 to $5.99. Three categories, "Online Travel Agency", "Language Assistant", and "Flight Manager", account for about 35% apps included in this study. "Online Travel Agency" (35 apps, 11.7%) refers to apps that provide services to search and book air

tickets, hotel rooms, and rental cars. These apps are likely the extension of websites (29 out of 35 apps). That is, the developers are the owners of existing online travel agencies such as Travelocity and Tripadvisor or hotel groups such as Hilton. "Language Assistant" (34 apps, 11.3%) refers to apps that assist language translation. Some of these apps (8 apps) provide translation in multiple formats such as text- and voice-based. Also, users can search for local greetings and common phrases using the apps. Some apps in "Language Assistant" are specialized in the translation of one language (25 apps covering 18 kinds of language), while some apps serve for the translation of multiple languages (6 apps). Most of language assistant apps are free of charge (23 apps) and the others cost from $0.99 to $2.99. "Flight Manager" (34 apps, 11.3%) refers to apps providing live flights information. Users can use these apps to search and track flights, find out gates or terminal changes, check-in, and manage mileage reward accounts. Also, some of these apps provide some kind of services of online travel agencies such as flight and rental car booking. Among the apps in "Flight Manager", only four apps are developed by airline companies (Air Canada, British Airlines, Southwest Airlines, Alaska Airlines, and Lufthansa). Eleven of them cost from $0.99 to $9.99, while the rest are free.

"Theme Park & Resort Guide" and "Facilitators" account for about 20 percent of apps. This category (31 apps, 10.3%) refers to the app that provides travel tips within theme parks, resorts or cruises. Twenty apps are about Walt Disneyland and other eleven apps are about five theme parks, three cruise lines, and three resorts. The major function of theme park apps is to acquire rides waiting times usually submitted by the visitors who are present in the parks. Apps in the category of "Facilitator" (30 apps, 10%) provide a variety of information services for tourists including GPS-enabled theft tracker, cheap gas stations, bathroom locations, packing list, Wi-Fi spots, mobile battery life, times in different places, airport maps, and Catholic Mass times church directory.

The other two main categories are "Multiple City Destination Guide" and "Food Finder". "Multiple City Destination Guide" (26, 8.7%) is similar to "Single City Destination Guide" regarding the functions as a tour guide. However, each app in this category provides guide for multiple cities and strongly depends on location-based service for recognizing the positions of the user and providing context-relevant information. "Food Finder" (24 apps, 8.0%) refers to apps for finding restaurants, bars, and anything related to night life activities. These apps usually provide recommendations to users based on some user-specified criteria. Also, most of these apps provide access to customer reviews. Only one app requires payment in the amount of $0.99. Other three categories ("Entertainment", "Live Camera", and "Currency Converter") account for about 8% of apps. Apps in "Entertainment" (14, 4.67%) and "Live Camera" (5, 1.7%) are mainly for enriching touristic experience. For example, some apps can help to turn the pictures taken by tourists to postcards and share immediately. Some apps can pinpoint the places tourist visited within one trip on the map and link each point to a photo album related with that place. Last, "Currency Converter" (4, 1.3%) helps to convert currencies by applying the most updated exchange rates.

Table 2. Unique Information Services of Apps in Each Category

Category	Unique Information Services in each category of apps	
A. Language assistant (6 types)	basic_direction_in_the_local conversational_language exist_traveller_language_guide_podcast	language_translation offer_more_native_voice_translation phrase_book
B. Flight manager (15 types)	arrival_and_departure alternate_flight baggage_claim board_pass break_news car_rental_and_reservation contact_information flight_track	flight_alert flight_map flight_search flight_status reward_account trip_itinerary weather_report
C. Online travel agency (9 types)	avis_location book_a_stay_or_view choice_hotel_locator hhonors_reward_point_balance hotel_location	hotel_photo modify_upcoming_reservation priority_club_reward secure_book
D. Theme park & Resort guide (7 types)	crowd_level events_schedule hidden_treasures park_map	parade_time ride_wait_time rides_height_requirement
E. Food finder (5 types)	chain_restaurants_contacts food_recommendations local_eats	local_specialty restaurants_reviews
F. Single city destination guide (10 types)	car_park_and_underground_station favourite_guidebook include_restaurant information_from_wikipedia_and _openstreetmap leisure_poi	london_tube museum_tour night_life post_office ticket_price
G. Multiple city destination guide (4 types)	local_expert mom_map	rocket_taxi taxi_company
H. Live camera(1 type)	cam_live_view	
I. Entertainment (3 items)	npr_station_finder photo_and_video	track_trip
J. Currency converter(1 item)	currency_converter	
K. Facilitators (7 items)	aaa_roadside cheap_gas classify_ad dog_friendly	pack_list postal_code roadside_assistance

4.2. Design Features of Apps

Table 3 presents the design features of each category of apps, which can be organized into four main categories (information models, positioning, input/output modalities, and network infrastructure) and twelve sub-categories by applying Kenteris and Gavalas' (2011) framework.

Table 3. Design Features of Apps

	Design Features	**Descriptive Phrases**	
Information Models	Personal Profiling	bookmark_manager favorite_list	personalised_and_fully_searchable save_your_station_and_route
	Collaborative Filtering	similar_recommendations	
	Context Aware	application_indicate_the_station_close popup_show_you_the_distance	locate_all_nearby_poi push_notification
	Updating Content	add_it_by_hand itinerary_update	constantly_update recent_time_submitted
	Sharing	itinerary_share editor_and_reporter_share	recent_time_submitted family_and_friend
Input/ Output Modalities	Voice Input/Output	translate_with_voice	strike_image
	3D Modeling Output	augment_reality zoomable_map	
	Interface Design	easy_and_so_much_fun intuitive_user_interface landscape_and_portrait_mode perfectly_formatted_and_optimize	easy_navigation information_interactive quickly_access simple_interface
	User-Interaction	enter_flight sort_and_filter easily_and_quickly search_by_city read_and_submit double_tap_to_call	street_search_functionality search_and_book instant_access list_view single_click category_to_choose
Network	Offline Mode	perfectly_while_in_the_subway no_roam_charge	offline_map connect_to_the_internet
Positioning	Mapping Tech	application_locate pinpoint_routes	locate_on_the_map interactive_map

Generally, apps were developed to leverage the benefits from mobility and network connectivity of smartphones. First, the results demonstrate that the designers catered to a variety of information models to enable users to acquire personalized recommendations efficiently and to assist users in intelligent and care-taking ways by taking into account contextual factors and through various auto-synchronization

features. For example, some apps help users create and save lists of favourite items for future use (personal profiling). Some apps also provide recommendations according to the user's locations (context aware) and search history (collaborative filtering). Particularly, with the always-on feature of smartphones, some apps utilize "push notification" to remind users for important things. Second, the results demonstrate that most of the apps are equipped with multiple modalities of input according to the principle of simplicity and convenience. For example, users can input words for translation by speaking to the app. Users can search for restaurants by defining in the filters according to cuisines, distance, opening hours etc., and call for reservation by just a single tap on the text of phone number rather than inputting phone number in a traditional way. Third, some apps make maps more intuitive and interactive (positioning). The target locations that users look for can usually be positioned on the map and the route to reach a specific target can be highlighted on the map. Thus, the touch screen enables a "conversation" between user and map. Last, some apps can provide information without requiring Internet connection (network infrastructure/offline mode). This design is based on the consideration that tourists are more likely to need assistance in the area without Internet connection. Therefore, some apps can store information in the smartphone for back-up services.

4.3. Analysis of Customer Reviews

All the customer reviews (33,867 records) related with the top 300 apps were analysed. The top ten keywords clusters (ranked in the number of reviews that mention these keywords) were identified for each category of apps in the cluster analysis of the customer reviews. This resulted in a total of 110 clusters for the 11 categories of apps. In order to identify the customers' preferences toward the features inferred by the clusters, a series of steps were followed to differentiate the features that customers preferred and the ones that are complained about. First, for each cluster, the number of customer reviews with different rating scores (i.e. 1 to 5) were counted and chi-square analysis was applied to examine the capability of this cluster (the bunch of keywords) in differentiating the rating scores (i.e. low (1 or 2) or high (4 or 5)). Second, the differences of the number of customer reviews between high ratings (total percentage of 4 and 5 rating scores) and low ratings (total percentage of 1 and 2 rating scores) were calculated. A positive result indicates more users report high rating scores on the feature inferred by the cluster than the ones report low rating scores. That is, customers are more likely to be fond of the feature inferred by the keywords of clusters. Conversely, a negative result indicates that customers dislike the feature. The analysis identified a total of 74 clusters associated with higher ratings and 36 clusters with lower ratings.

Table 4 presents some examples of the information services and design features that customers like and dislike. For information services, it seems that the most favourable information service (with most high rating scores) is the service that is only available on the mobile platform (Cluster 40 in Category D) including "live chat", "real wait time", "watch park from webcam anytime". For example, customers comment that they "feel fun to get waiting times that reply on consumers data input" and they enjoy "the live chat with the people in the parks". Also, customers like to "reserve dinner tables by a few tabs in the app" (Cluster 10). Besides, customers also show their

preferences on the accurateness of the information and quick responses provided by responsive developers.

Table 4. Examples of User Preferred Information Services and Design Features

	Category	Cluster ID	Keywords	Difference between Positive and Negative ratings
Information services	D	40	live chat; real wait time; watch park from webcam anytime	81.7%
	D	10	easy for dinner reservation; res menus and prices	70.9%
	D	41	easy to find history; excellent resources	70.9%
	E	8	favourite app; help to find local food	68.8%
	J	23	fast and accurate converter	66.0%
	K	38	handy tool, save time	64.3%
	C	85	bidding price is cool, easy and effective	63.8%
	D	14	help to avoid lines; check inside and outside (parks)	62.9%
	E	50	club, bars info; happy hour special drink	61.3%
	I	12	contact lists; address; email; easy find	60.9%
Design features	E	5	help decide where to eat; easy, fun and useful; live in big city; nice idea; handy tool; perfect for indecisive people	76.7%
	G	7	simple, easy, quick interface	74.6%
	E	2	shake function is pretty cool; pick up randomly	68.4%
	K	103	accurate, fast, simple and straightforward interface	66.0%
	J	61	static keypad and multiple ways to input	64.4%
	B	89	clean interface; intuitive design; easy, simple, and fast; user friendly	59.8%
	E	1	filter ability;browse; search by category; choose cuisine; price range; current location; randomly select; nearby location	57.4%
	F	16	easy navigation, wish list	53.3%
	J	26	flawless and simplicity job	52.6%
	D	12	browse attraction list, save favorites	45.2%

Note: C= Online travel agency; D= Theme park & Resort guide; E= Food finder; F= Single city destination guide; G= Multiple city destination guide; I= Entertainment; J=Currency converter; K= Facilitators

For the design features, generally, customers like the design of multiple filters which can enable them to personalize their search. They also like the listed information (i.e. lists of hotels nearby) when they have difficulties to set up search criteria. Particularly, customers like the innovative design such as the "shake" function provided by the app "Urbanspoon" which can help them randomly select restaurants (Category 5 and 2). In addition, customers likes "simple, easy and intuitive interface" (Cluster 103 and 89). Besides, customers like the designs for personal profiling such as the functions to make wish or favourite lists (Cluster 26 and 12). In sum, customers

showed great interests in innovative services offered by the smartphone. The results also identify several "problems" with the respective apps. The low rating scores are most likely to be associated with comments such as "waste money", "piece of crap", and "not recommend to buy." From these analyses, one can see that customers are easier to be dissatisfied if they paid, but the apps did not work properly. Also, the customers do not like the design in which the apps simply directed them to the websites while not providing services independently. Last, it seems that customers care deeply about the quality of support provided by the developer; that is, low ratings are more likely to be associates with situations whereby the developer does not respond to customers' inquiries and/or fix bugs.

5. Implications and Conclusions

This study focuses attention on the emergent mobile technologies represented by iPhone apps that support travel. The results clearly show that smartphones are truly handheld computers with the all-in-on design. They have comprehensively extended the Internet-enabled services such as information search, reservation and e-commerce, multimedia content consumption and creation as well as social communications to the mobile platform. Besides, smartphones leverage a series of innovations through the applications to provide innovative services such as location-based services, context-aware recommendations, and improvisational selections and decisions. As such, smartphones and their apps reflect the technical potentials and capabilities of today's mobile technologies in transforming tourism experiences. They are, indeed, the new landscape of travel for today's tourists. The results on customers' preferences over the information services indicate that customers are the fans for the kind of "smartphone-enabled" services, which are only available on the mobile platform such as real-time trip route recording and the "shake" function to randomly select restaurants. These services provide new experiences to customers and enable them to make decisions or consume destinations in innovative ways. Also, the analysis on the design features suggests that customers are usually fond of intuitive interface and considerate user-interaction design. On the opposite, customers tend to frown upon apps that only serve as a gateway to the regular websites and those with unreliable programming. Although exploratory in nature, this study establishes an important foundation to support the study of mobile technologies for travel, including system evaluation, use behaviour, as well as mobile business models. Technology constantly changes, and future research should focus on the evolution of the information services available on a range of mobile platforms. Following Fulk (1993), the mutual determinative relationships between technology and the user suggest that the user and mobile technological innovations will continue to shape and reshape each other through a process of interaction. As such, future research can focus on the various processes which affect the depth and breadth of changes, their impact of travel behaviour, as well as business opportunities for the travel and tourism industry.

References

Anckar, B. & D'Incau, D. (2002). Value creation in mobile commerce: findings from a consumer survey. *Journal of Information Technology Theory and Application* 4(1): 43-64.

Brown, B. & Charlmers, M. (2003). *Tourism and mobile technology.* Paper presented at the Eighth European Conference on Computer-Supported Cooperative Work, Helsinki, Finland.

Carlsson, C., Walden, P. & Yang, F. (2008). *Travel MoCo - a mobile community service for tourists.* Paper presented at the 7th International Conference on Mobile Business.

Chang, Y. F., Chen, C. S. & Zhou, H. (2009). Smart phone for mobile commerce. *Computer Standards & Interfaces* 31(4): 740-747.

Dala-Ali, B. M., Lloyd, M. A. & Al-Abed, Y. (2011). The use of the iPhone for surgeons. *The Surgeon* 9*:* 44-48.

Domingos, P. & Pazzani, M. (1997). On the optimality of the simple Bayesian classifier under zero-one loss. *Machine Learning* 29(2): 103-130.

Feldman, R. & Sanger, J. (2007). *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*. New York: Cambridge University Press.

Fulk, J. (1993). Social construction of communication technology. *Academy of Management Journal* 36(5): 921-950.

Gretzel, U. (2010). Travel in the network: redirected gazes, ubiquitous connections and new frontiers. In M. Levina & G. Kien (Eds.), *Post-global Network and Everyday Life* (pp. 41-58). New York: Peter Lang.

Gretzel, U. & Jamal, T. (2009). Conceptualizing the creative tourist class: technology, mobility and tourism experiences. *Tourism Analysis* 14(4): 471-482.

Grün, C., Pröll, B., Werthner, H., Retschitzegger, W. & Schwinger, W. (2008). *Assisting tourists on the move - an evaluation of mobile tourist guides.* Paper presented at the 7th International Conference on Mobile Business.

Katz, J. E. (2006). *Magic in the Air: Mobile Communication and the Transformation of Social Life*. New Brunswick, NJ: Transaction Publishers.

Kenteris, M. & Gavalas, D. (2011). Electronic mobile guide: a survey. *Personal and Ubiquitous Computing* 15(1): 97-111.

Kenteris, M., Gavalas, D. & Economou, D. (2009). An innovative mobile electronic tourist guide application. *Personal and Ubiquitous Computing* 13(2): 103-118.

Krippendorff, K. (2004). *Content Analysis: An Introduction to Its Methodology* (2nd ed.). Thousand Oaks: SAGE.

Lashkari, A. H., Parhizkar, B. & Mohamedali, M. A. (2010). *Augmented reality tourist catalogue using mobile technology.* Paper presented at the International Conference on Computer Research and Development.

O'Brien, P. & Burmeister, J. (2003). Ubiquitous travel service delivery. *Information Technology & Tourism* 5(4): 221-233.

Poslad, S., Laamanen, H., Malaka, R., Nick, A., Buckle, P. & Zipf, A. (2001). *CRUMPET: creation of user-friendly mobile services personalized for tourism.* Paper presented at the 3G 2001, London.

Rasinger, J., Fuchs, M. & Hopken, W. (2007). Information search with mobile tourist guides: A survey of usage intention. *Information Technology & Tourism* 9(3/4): 177-194.

Roussos, G., Marsh, A. J. & Maglavera, S. (2005). Enabling pervasive computing with smart phones. *IEEE Pervasive Computing* 4(2): 20-27.

Sebastiani, F. (2002). Machine learning in automated text categorization. *ACM Computing Surveys* 34(1): 1-47.

Tussyadiah, I. & Fesenmaier, D. R. (2009). Mediating tourist experiences: Access to places via shared videos. *Annals of Tourism Research* 36(1): 24-40.

Wang, D., Park, S. & Fesenmaier, D. R. (2011). The role of smartphones in mediating the touristic experience. *Journal of Travel Research,* Accepted.

Want, R. (2009). When cell phones become computers. *IEEE Pervasive Computing 8*(2): 2-5.

Recommending Hotels based on Multi-Dimensional Customer Ratings

Dietmar Jannach[a],
Fatih Gedikli[a],
Zeynep Karakaya[a], and
Oliver Juwig[b]

[a] Fakultät für Informatik
TU Dortmund, Germany
{firstname.lastname}@tu-dortmund.de

[b] HRS Hotel Reservation Service
50676 Cologne, Germany
oliver.juwig@hrs.de

Abstract

Recommender Systems (RS) have shown to be a valuable means to support the traveller or tourist in his pre-trip information search and decision making processes. These systems often rely on rating information provided by the user community to make recommendations for individual users. In classical application domains such as movie or book recommendation, users provide one overall rating for each item. Customers in the travel and tourism domain however are often allowed to evaluate their hotel or holiday packages along several dimensions after the trip. In this work, we show through an empirical evaluation based on a real-world data set from the tourism domain that the predictive accuracy of an RS can be significantly improved when the multi-dimensional rating information is taken into account. In particular, we demonstrate that regression-based methods and in particular the novel combination of user- and item-based models leads to more accurate recommendations than previous approaches. In addition, we show that not all dimension (criteria) ratings are equally valuable for the prediction process and that a careful selection of rating dimensions can help to further increase the quality of the recommendations.

Keywords: recommender systems, collaborative filtering, multi-criteria ratings.

1. Introduction

Recommender Systems (RS) are information filtering and decision support tools, which have been successfully applied to support the online customer in various domains. In the travel and tourism domain, such systems are for example developed to help the customer in the pre-trip information search and decision making process. Examples of recent research include knowledge-based and conversational approaches to filter destinations and select travel packages (Jannach et al., 2009; Jannach et al., 2007; Zanker et al., 2008), context-aware recommendation of places of interest (Baltrunas et al., 2011), mobile recommenders (Ricci, 2011), or the development of more intelligent user interaction strategies (Mahmood et al., 2009).

Recommendation in the tourism domain has some specific particularities and challenges, which are not present in more classical RS application domains, in which especially collaborative filtering (CF) techniques have been successfully applied in the past. Customers in the tourism domain, for example, do not purchase items as

frequently as customers of an online book store or movie rental system. Thus, the amount of user feedback and the buying history available for building systems based on collaborative filtering techniques may be limited, which is why conversational approaches are often chosen. Furthermore, the context of the traveller or tourist is particularly relevant. Think, e.g., of making recommendations for a group of people traveling together. Also, the type of the trip (business or private) or seasonal aspects can be important when recommending tourism products.

With regard to the sparsity of the data note that while the number of past transactions and ratings per user is relatively low when compared to other domains, one particular aspect in the tourism domain is that customers are often not only allowed to assign an overall rating to a hotel or travel package, but can rate the product or service along several dimensions. In the hotel domain, the typical dimensions or criteria include "room quality", "service quality", or "value-for-money", see Figure 1.

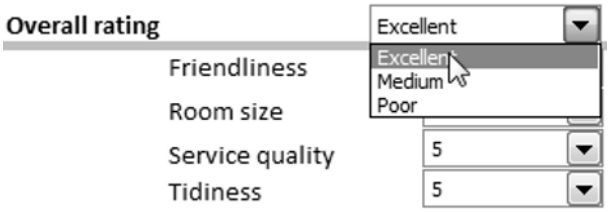


Figure 1. Example rating screen.

The value of such multi-criteria ratings for improving the accuracy of CF recommender systems has been analyzed in the past, e.g., by Adomavicius and Kwon (2007), who also proposed different algorithms that take these detailed ratings into account in the prediction process, see also (Adomavicius et al., 2011).

In this work, we build on previous work of Adomavicius and Kwon and propose a CF recommendation approach based on Support Vector regression (Drucker et al., 1997), which combines user- and item-models in a weighted approach to further improve the accuracy of the recommendations. We evaluate our method on a real-world data set provided by a major European tourism platform and compare it with state-of-the-art baseline algorithms based on matrix factorization. In addition, we aim to explore in this work whether or not all available criteria ratings are equally valuable to increase the predictive accuracy or if only a subset of the ratings should be exploited.

2. Improving RS Accuracy Based on Multi-Criteria Ratings

2.1. Single-Rating and Multi-Criteria Rating Algorithms

In the classical setting for collaborative filtering, the input to the RS is a user-item rating matrix and a recommender system can be seen as a function that returns a rating prediction (or relevance score) for a given user-item combination. In Figure 2, the problem for example consists of predicting the rating of User1 for Item4.

Over the last fifteen years, a variety of algorithms have been proposed to calculate these relevance scores from the known ratings, starting with from nearest-neighbour methods over probabilistic approaches to recent matrix factorization algorithms[1].

	Item1	Item2	Item3	Item4
User1	5	5	5	?
User2	4	3	5	5
User3	2	1	-	2
User4	4	-	4	5

Figure 2. CF based on single ratings.

In multi-criteria rating settings, however, we assume that we know more about the user's preferences than just his overall evaluation for the item. In particular, we assume that the user has additionally expressed his opinion on several different aspects of the item, e.g., the friendliness of the hotel staff or the tidiness of the rooms. The general idea of multi-criteria is thus to take that additional information into account in the recommendation process. If we, for example, know that the room tidiness is particularly important for a customer (by analysing the relation between the overall rating and the rating for tidiness), we could try to recommend more hotels which have received a high rating for tidiness also by other customers.

2.2. Aggregation Function Based Multi-Criteria RS

The most precise approaches to generate recommendations based in multi-criteria ratings presented in the study by Adomavicius and Kwon are called "aggregation function based". The main goal of such approaches is to determine how the individual criteria ratings r_1 to r_k are related to the overall rating r_0 of an item and use that information as a prediction function for the overall rating as shown in Equation (1).

$$r_0 = f(r_1,...,r_k) \tag{1}$$

In the movie domain, this for example would mean to learn a function that describes in which way the ratings for the "story" and other aspects are related to the overall rating. In (Adomavicius and Kwon, 2007), the authors propose to determine the aggregation function by learning a regression function of the following form for each item, where the w_i's are the learned weights for the dimensions r_i and c is a constant.

$$r_0 = w_1 * r_1 + w_2 * r_2 + ... + w_k * r_k + c \tag{2}$$

Once the (estimated) values of the function are known, the overall rating for item i and user u can be predicted by parameterizing the regression function for item i with u's ratings for the different criteria. The problem, of course, is that u's criteria ratings for item i are also not known. In Adomavicius and Kwon (2007), it is therefore proposed to predict those dimension ratings for a given user-item combination using, for example, standard CF methods in a first phase. For the Yahoo!Movies

[1] See (Jannach et al., 2010) for a recent overview.

(http://movies.yahoo.com) data set used in their experiments, they for example use a nearest-neighbour approach to predict *u*'s ratings for the dimensions "action", "story" and so forth. We will follow this approach also in our experimental study.

Figure 3 illustrates the approach. The entry 5(4,3,4,5) in the table, for example, denotes that the user gave an overall rating of 5 and the criteria ratings 4, 3, 4 and 5 respectively. In the first phase, the "neighbours" for User1 with respect to the last rating dimension are determined (let us assume User2 and User4). Their predictions are then used to predict the rating for the last dimension of Item4, for which we also seek an overall recommendation. Once all criteria ratings are determined, the overall rating is estimated based on the learned regression function from Equation (2).

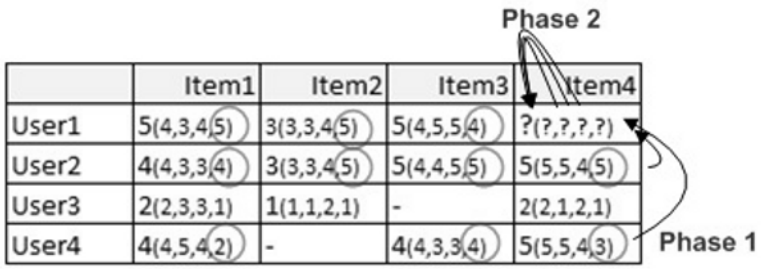


Figure 3. Aggregation function based approach.

2.3. Proposed Enhancements

We propose the following improvements and extensions compared with previous works.

- *Use a weighted combination of user and item models.* In Adomavicius and Kwon (2007), aggregation functions as shown in Equation (2) are learned per movie. The goal is therefore to determine the importance of the individual dimensions for the overall rating for each movie across all users. However, it is also possible to learn such a function also per user, that is, we could try to learn how important the "action" aspect is for a particular user across all movies. Since we assume that both regression functions carry valuable information, we propose in this work to combine the two predictions obtained from the different models in a weighted approach.
- *Use support vector regression.* Instead of using least squares regression as done in previous works, we propose to use Support Vector (SV) regression to calculate the regression functions, because this method has shown to be more accurate than other regression methods also in other recommender applications (Sen et al., 2009).
- *Selection of rating dimensions.* In our real-world data set from the tourism domain used in this study, customer ratings are available for more than a dozen different aspects for each hotel. Our hypothesis is that not all rating information is equally valuable to predict the overall rating and that some of the ratings might be better predictors than others. We will therefore evaluate different settings in which we intentionally omit parts of the available rating information.

3. Empirical Evaluation

We conducted several offline experiments on a data set provided by HRS.com, a major European platform for hotel reservations, and compared the predictive accuracy of our newly proposed techniques with top-performing, recent recommendation algorithms. In order to demonstrate the general usability of the weighted regression approach also for other domains, we ran further experiments on a data set obtained from Yahoo!Movies, which also contains multi-dimensional ratings and which was also used in the studies by Adomavicius and Kwon.

3.1. Data Sets

Hotel ratings. In order to analyse how the density of the rating information affects the recommendation accuracy, we created variations of the original data set. To that purpose, we varied the constraints *minimum number of ratings per user* and *item* (see the columns *R/U* and *R/I* respectively in Table 1) on the data set and removed users or items for which not sufficient data was available, thus increasing the density of the data set. Table 1 lists the different quality levels used for evaluation. In the `hotel-high` data set, for example, each user has rated at least 5 hotels and each hotel has at least 5 ratings attached. Note that the hotel rating data set is extremely sparse when compared to other data sets from literature as most users have rated only very few hotels. In the well-known MovieLens data set, for example, each user has rated at least 20 movies, which we consider to be an unrealistic constraint in the tourism domain.

Beside the variations with respect to the rating sparsity, we also varied the set of dimension ratings to be taken into account for the prediction task. Table 2 shows an overview of the characteristics of the used data sets, which differ with respect to the considered subset of dimension ratings.

Table 1. Hotel data set quality levels.

Dataset/Constraint	*R/U*	*R/I*
`hotel-high`	5	5
`hotel-medium`	3	3
`hotel-low`	1	1

Overall, the described density variations listed in Table 1 and the variation with respect to the criteria ratings listed in Table 2 lead to a total number of 12 data set variations which we analysed in our evaluation.

In order to determine the most relevant dimensions for the `dim-chi-squared` data set, we computed the value of the chi-squared statistic with respect to the overall rating for each dimension. We retained the 14 dimensions with the highest weights since the weight of the next dimensions was significantly lower. In the `dim-price-performance` data set, we based the prediction of the overall rating only on the customer's rating on the price performance ratio aspect of the hotel. Note that with the emergence of Web-portals, the hotel market became very transparent, open, and competitive. Our hypothesis is thus that the rating on the price performance ratio

strongly correlates with the overall rating. For the `dim-domain-knowledge` data set, finally, the 7 dimensions taken into account were selected based on domain expertise. Note that four of them were also in the top-7 list determined based on chi-squared statistics.

Table 2. Different subset combinations of the rating dimensions in the hotel data set.

Name	Description
`dim-all`	All dimensions are taken into account.
`dim-chi-squared`	The most relevant dimensions according to a chi-squared statistic test are used.
`dim-price-performance`	Only the price-performance dimension is used.
`dim-domain-knowledge`	A set of seven dimensions determined by a domain expert are used.

Yahoo!Movies data set. Based on the movie rating data set provided by Yahoo! Research, we collected additional criteria rating information from the Yahoo!Movie web site, where users can rate the available movies along four different dimensions (Acting, Story, Visuals, Directing). In order to make our results comparable to previous research, we applied the constraint that each user and item must have at least 10 ratings. We refer to the resulting data set as `YM-10-10`.

3.2. Algorithms Compared in the Evaluation

Single-rating baseline algorithms.

- `slope-one`: Slope One is a family of prediction schemes as proposed by Lemire and Maclachlan (2005). It was designed to be an algorithm which is comparably easy to implement, supports the incorporation of new ratings, has a reasonable accuracy and is efficient at run-time. Slope One predictors have the form $f(x) = x + b$ and are based on rating differences between items. In our experiments we used the (basic) Slope One scheme because the results are on a par with slower memory-based schemes such as nearest-neighbour approaches often used in comparisons.
- `funk-svd`: Over the last years, matrix factorization techniques have shown to be a good basis to develop highly accurate recommender systems. In these approaches, the goal is to automatically identify a set of latent semantic features (aspects or factors) which characterize the available items using for example Singular Value Decomposition (SVD) or probabilistic approaches. Predictions can then be made by determining and combining the information about the position of each item and user interest profile in this reduced feature space. In our evaluation, we used an algorithm recently proposed in the context of the Netflix Prize by Simon Funk (pen name)[1]. For the SVD-based recommender, we used 30 latent features, a number which we determined empirically.

[1] See: http://sifter.org/~simon/journal/20061211.html

Multi-criteria algorithms.

- `item-based SV regression`: This algorithm corresponds to the idea of item-based regression models as proposed in (Adomavicius and Kwon, 2007). In our work, however, we use Support Vector (SV) regression to learn the regression function.
- `user-based SV regression`: The regression model in which we learn the coefficients for each user. Again we use SV regression.
- `weighted SV regression`: This method combines the estimates of item- and user-based SV regression using the harmonic mean of the two estimates which yielded the best results in our evaluation.
- `item-based lin-regression`: This algorithm is similar to the original regression approach proposed in (Adomavicius and Kwon, 2007) and relies on Ordinary Least Squares regression instead on SV regression.

Note that customers in some cases only gave ratings for some but not all dimensions. For the multi-criteria algorithms, we approximated these missing ratings by taking the average of the user's other dimension ratings for an item.

For the SV regression methods, we furthermore used the empirically determined penalty parameter $c = 0.15$ and $p = 0.1$ for epsilon in the loss function. Note that we also made experiments with the Ordinary Least Squares regression algorithm `item-based lin-regression` on the hotel data set. This method is however not applicable for sparse data sets with many rating dimensions as the method requires that the number of data points is higher than the number of regression coefficients. Note that in the hotel data set we have more than a dozen criteria but most hotels and users have much less than 12 ratings. Experiments on the Yahoo!Movies data set, however, showed that SV regression outperforms the Least Squares regression approach.

All algorithms were implemented in our Java-based *Recommender Suite* framework, which also includes components to run offline experiments, do cross-validation and measure various metrics such as accuracy or coverage. For the SV regression calculations we used the `libsvm` library[1].

Evaluation Metrics. We rely on standard metrics to evaluate the accuracy of our method. For determining the *Root Mean Squared Error* (RMSE) the available rating information was split into training data (95%) and test data (5%). In order to factor out effects of randomness, we used random subsampling, repeated the experiments appropriately and report the average value of all runs.

Beside the RMSE, we also report the standard Information Retrieval (IR) metrics precision, recall, and F1 and rely on the procedure described by Nakagawa and Mobasher (2003). In this approach, the numerical rating predictions are transformed into binary "like" and "dislike" statements. In order to measure precision and recall, we compare the number of existing like statements (ELS) in the test set with the

[1] See: http://www.csie.ntu.edu.tw/~cjlin/libsvm

number of predicted like statements (PLS) returned by the algorithms. Precision and recall are measured as follows: $\text{Precision} = \frac{|PLS \cap ELS|}{|PLS|}$; $\text{Recall} = \frac{|PLS \cap ELS|}{|ELS|}$.

F1 is calculated as the harmonic mean of the two. When determining these IR metrics, we used five-fold cross-validation.

3.3. Results and Observations

Accuracy results. Tables 3 to 6 show the RMSE numbers for the different hotel data sets. Note that the overall rating (indicated by the user-specified general "feel good factor") in the hotel data sets is given on a 1-to-3 rating scale.

Table 3. RMSE results for the hotel data set (`dim-all`).

Algorithm	hotel-high	hotel-medium	hotel-low
`weighted SV regression`	0.561	0.575	0.640
`item-based SV regression`	0.613	0.614	0.722
`funk-svd`	0.633	0.640	0.668
`user-based SV regression`	0.642	0.676	0.683
`slope-one`	0.725	0.718	0.786

The following observations were made for all hotel data set variations. Firstly, the results confirm the general value of exploiting multi-criteria ratings for recommender algorithms. Moreover, the herein proposed `weighted SV regression` method significantly outperforms both single-rating methods `funk-svd` and `slope-one` as well as the individual regression approaches throughout all hotel data set variations.

Table 4. RMSE results for the Hotel data set (`dim-chi-squared`).

Algorithm	hotel-high	hotel-medium	hotel-low
`weighted SV regression`	0.520	0.528	0.677
`item-based SV regression`	0.581	0.577	0.758
`user-based SV regression`	0.594	0.629	0.757
`funk-svd`	0.632	0.638	0.668
`slope-one`	0.749	0.717	0.789

Table 5. RMSE results for the Hotel data set (`dim-price-performance`).

Algorithm	hotel-high	hotel-medium	hotel-low
`weighted SV regression`	0.521	0.533	0.640
`item-based SV regression`	0.576	0.578	0.768
`user-based SV regression`	0.581	0.614	0.694
`funk-svd`	0.634	0.638	0.653
`slope-one`	0.725	0.723	0.788

Table 6. RMSE results for the Hotel data set (`dim-domain-knowledge`).

Algorithm	hotel-high	hotel-medium	hotel-low
`weighted SV regression`	0.515	0.528	0.533
`item-based SV regression`	0.565	0.582	0.758
`user-based SV regression`	0.587	0.610	0.757
`funk-svd`	0.633	0.638	0.668
`slope-one`	0.721	0.718	0.789

The individual regression-based approaches alone are also able to outperform the single-rating approaches except for the setting in which all dimensions are taken into account (Table 3) and where user-based regression works slightly worse than `funk-svd`. Item-based regression on the other hand is consistently better than the single-rating approaches in all settings and dataset variations.

Note that even though the weighted approach is already the best performing one, we see potential for further gains with respect to accuracy, e.g., by optimizing the parameter settings for the regression learner or by combining the user- and item-based predictions in a different way.

We can see from the evaluation data that increasing the quality level by adding constraints on the minimum number of ratings per user and item the predictive accuracy increases. Intuitively, the RMSE values thus decrease when more rating information per user and item is available. Note however, that the different data sets are not fully comparable with respect to the number of hotels and items. The "hotel-high"-data sets are for example derived from the same raw data set as the "hotel-medium"-ones, which means that they comprise less hotels, items and ratings. Still, we can observe that the RMSE values are significantly better in the "hotel-high" data set even though less training data was available.

Table 7. F1, precision, and recall results for the hotel data set `dim-chi-squared`.

Algorithm	hotel-high			hotel-low		
	F1	Prec.	Recall	F1	Prec.	Recall
`weighted SV regression`	93.47	93.47	93.47	70.97	73.24	68.84
`item-based SV regression`	92.97	92.97	92.97	69.51	71.79	68.74
`user-based SV regression`	92.68	92.68	92.68	70.87	73.14	67.39
`funk-svd`	92.53	92.53	92.53	69.23	71.51	67.10
`slope-one`	46.13	49.29	43.40	08.31	10.17	07.03

Table 7 shows the results for the standard IR metrics precision, recall, and F1 on the hotel data sets `hotel-high` and `hotel-low` (`dim-chi-squared`). We can see that the results for the IR metrics are comparable to the RMSE results reported above and that our weighted approach outperforms the other algorithms also on this measure. In

that context, note that the single-rating method slope-one performs extremely poor on the highly sparse data set.

Impact of feature selection. The tables showing the RMSE values (Table 3 to Table 6) also highlight the importance of choosing an adequate subset of dimensions as this selection can significantly affect the performance of a multi-criteria recommender algorithm. However, the results we obtained so far are not conclusive and will be analysed in more detail as a part of our future work.

Our results indicate that feature selection based on statistics or machine learning paid off for the high quality data set (hotel-high). On the other hand, for the low-quality data set (hotel-low), which is more realistic in real world scenarios, feature selection by a human expert with domain knowledge was the most successful approach. Based on the observation that the price-performance ratio is one of the most important factors in the domain and overpriced hotels are quickly penalized with low reviews by customers, we can see from the results in Table 5 that relying only on this detailed rating information can lead to comparably good results.

Finally, as mentioned above, the user-based regression approach does not work better than funk-svd when all available dimension ratings are taken into account (Table 3). This again stresses the importance of selecting the most important features for the prediction process.

Accuracy results for the movie domain. We conclude this section with a discussion of the results obtained on the Yahoo!Movies data set. These experiments have been conducted to determine whether or not the weighted regression approach also works in domains other than hotel recommendation. Note that in this data set, the overall ratings were originally given on an A+ to F score and have been transformed to the usual 1-to-5-scale. Therefore, the RMSE values have to be interpreted differently than the values for the hotel data set.

Table 8. Accuracy results for the YM-10-10 data set.

Algorithm	RMSE	F1	Precision	Recall
weighted SV regression	0.625	90.69	90.69	90.69
user-based SV regression	0.652	88.18	88.18	88.18
item-based SV regression	0.694	89.93	89.93	89.93
item-based lin-regression	0.779	88.14	88.14	88.14
funk-svd	0.871	83.29	83.29	83.29
slope-one	0.888	82.64	82.64	82.64

Table 8 shows the accuracy results (RMSE, Precision, Recall) for the different algorithms. Similar to the hotel domain, the weighted regression approach performs best on all metrics. In contrast to the hotel data sets, however, all regression-based approaches are consistently better than the best single-rating based approach funk-svd. A comparison with the Least Squares regression method item-based lin-regression also indicates that SV regression is favourable for multi-criteria recommendation also for this domain.

4. Conclusion and Outlook

In contrast to other traditional application domains of recommender systems, multidimensional ratings play an important role on today's online tourism platforms. In this work, we have proposed to use Support Vector regression and the combination of user-based and item-based model to implement highly accurate multi-criteria recommenders. We analysed the predictive accuracy of our method in the domain of hotel recommendation based on a real-world data set. Our experiments showed that in particular the weighted combination of user- and item-based prediction models leads to the best performance measured in terms of standard accuracy metrics and that the results are better than those that can be achieved with state-of-the-art matrix factorization methods.

Beside the analysis of the accuracy for different density levels, we also ran experiments in which we varied the number of rating dimensions to be taken into account for the recommendation process. Although we could not develop a final guidance of how to select the best subset of dimensions, our initial results show that the selection measurably influences the recommendation accuracy. Our future work therefore includes the investigation of additional strategies of selecting the best subset of dimensions for the recommendation task. In addition, we plan to develop other schemes of combining the predictions returned by the user- and the item-based approach, e.g., based on the aspect of model quality. Currently, we are investigating an approach in which we try to learn individual weights for the item- and user-based component for each item and user as was done in (Gedikli et al., 2011).

An analysis of whether or not the existing information about the context of the customer (e.g., is it a business trip or a private travel) can be exploited to further increase the accuracy of the recommendations is part of our on-going work. In addition, future work in the area could also include the design of new approaches to extract additional customer opinions (on different aspects) from the free-text reviews. A first review of the data from a real-world platform in that context however revealed that most probably specific new algorithms are required as standard sentiment analysis methods might have difficulties in analysing text reviews that in many cases only consist of fragmentary sentences and individual keywords.

References

Adomavicius, G. & Kwon, Y. (2007). New recommendation techniques for multicriteria rating systems. *IEEE Intelligent Systems* 22: 48–55.

Adomavicius, G., Manouselis, N. & Kwon, Y. (2011). Multi-criteria recommender systems. In F. Ricci, L. Rokach, B. Shapira and P. B. Kantor (Eds.), *Recommender Systems Handbook*. Springer US.

Baltrunas, L., Ludwig, B., Peer, S. & Ricci, F. (2011). Context-aware places of interest recommendations for mobile users. In Proc. *HCI 2011*, Orlando, FL.

Drucker, H., Chris, Kaufman, B. L., Smola, A. & Vapnik, V. (1997). Support vector regression machines. *Adv. in Neural Inf. Proc. Systems* 9: 155–161.

Gedikli, F., Bagdat, F., Ge, M., Jannach, D. (2011). Fast and Accurate Computation of Recommendations based on Rating Frequencies, In Proc. *IEEE CEC 2011*, Luxembourg.

Jannach, D., Zanker, M., Felfernig, A. & Friedrich, G. (2010). *Recommender Systems – An Introduction*. Cambridge University Press.

Jannach, D., Zanker, M. & Fuchs, M. (2009). Constraint-based recommendation in tourism: A multiperspective case study. *International Journal of Information Technology and Tourism* 11(2): 139-155.

Jannach, D., Zanker, M., Jessenitschnig, M. & Seidler, O. (2007). Developing a conversational travel advisor with ADVISOR SUITE. In Proceedings *ENTER 2007*, Ljubljana, Slovenia.

Lemire, D. & Maclachlan, A. (2005). Slope one predictors for online rating-based collaborative filtering. In Proc. *SIAM Intl. Conf. on Data Mining*, Newport Beach, CA.

Mahmood, T., Ricci, F. & Venturini, A. (2009). Improving recommendation effectiveness: Adapting a dialogue strategy in online travel planning. *International Journal of Information Technology and Tourism* 11(4): 285-302.

Nakagawa, M. & Mobasher, B. (2003). A hybrid web personalization model based on site connectivity. In Proceedings of the *Workshop on Web Mining and Web Usage Analysis (WebKDD'03)*,Washington, DC, USA.

Ricci, F. (2011). Mobile recommender systems. *International Journal of Information Technology and Tourism* 12(3): 205-231.

Sen, S., Vig, J. & Riedl, J. (2009). Tagommenders: Connecting users to items through tags. In Proc. *WWW'09*, Madrid, Spain.

Zanker, M., Fuchs, M., Höpken, W., Tuta, M. & Müller, N. (2008). Evaluating recommender systems in tourism - a case study from Austria. In *Proceedings ENTER 2008*, Amsterdam, Netherlands.

Information Extraction for a Tourist Recommender System

Ago Luberg[a,b],
Priit Järv[a,b], and
Tanel Tammet[a,b]

[a] Eliko Competence Centre
Tallinn, Estonia
ago.luberg@eliko.ee

[b] Tallinn University of Technology
Tallinn, Estonia
tammet@staff.ttu.ee
priit@cc.ttu.ee

Abstract

We will present the information extraction algorithms for a semantic personalised tourist recommender system Sightsplanner. The main challenges: information is spread across various information sources, it is usually stored in proprietary formats and is available in different languages in varying degrees of accuracy. We will address the mentioned challenges and describe our realization and ideas how to deal with each of them: scraping and extracting keywords from different web portals with different languages, dealing with missing multi-lingual data and identifying the same objects from different sources.

Keywords: recommender system; information retrieval; entity disambiguation

1. Introduction

We will describe the information retrieval techniques used in the tourist recommender system Sightsplanner for Tallinn (http://tallinn.sightsplanner.com, [2011]). The first iteration of the system was called Smart City and is presented in Luberg et al. (2011). We use semi-automatic web scraping of different web portals to gather information about tourism objects. The data is transformed into our own custom ontology and cleaned up for later use by the recommender.

Information about the objects on the Internet is spread across various sources and is usually stored in different languages. The outcome shown to a tourist should be complete and comprise precise facts about the sights. The paper focuses mainly on how the data about the tourism objects is gathered, how it is stored and how it is cleaned up. We will present both the problems we have overcome and the problems we still have to tackle to improve the results of our data gathering module.

In the following two sections we will provide a brief overview of the system. In section 4 we will describe data retrieval process in more detail. We will end our paper with related work and a conclusion.

2. Tourism Recommender System

Sightsplanner is a semantic recommender and route composer system for tourists. A tourist can specify her interests on different categories, her location and time of the visit. Provided characteristics create a user profile which is sent to the recommender. Based on the profile, the system calculates a ranking score to every object in the database. The objects that the user likes more will get a higher score and vice versa. A schedule is created out of the objects which will give the best overall score. Several processes are involved when calculating the final list of objects: object verification process, matching process, planning process, result representation process, and feedback process.

The recommendation process starts when the user opens the web page and provides her preferences. A slider can be used to define the strength of interest on every object category. In our system we are using two levels of categories. One general category (like "*Music*") has subcategories (like "*Rock*", "*Jazz*"). Every object in the database has also defined types from the same set of categories. Besides interests the user can also define her start time, duration of the visit and preferred travelling method.

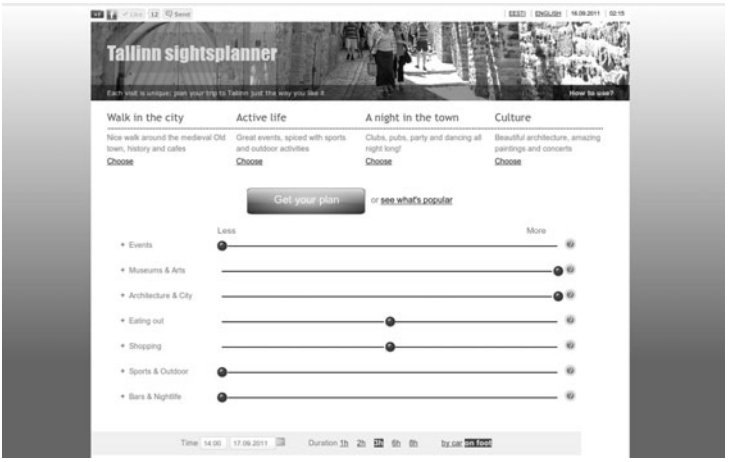


Fig. 1. Screenshot of the Sightsplanner.com frontpage

After the user has submitted the form, the created profile is sent to the recommender. The planner filters out objects by location, opening times and categories. Found objects are combined into a timetable iteratively. After a certain time limit the system returns the result with the highest total score which is presented to the user as a timetable and a map. The user can modify the result by giving feedback for objects: objects can be removed and the visiting duration of objects can be changed. A new planning process is started after the submission of the modifications.

3. Technical Overview

The main components of the system are presented in Figure 2. In this section we will describe data discovery and recommendation process with details about each of the components.

Information about the tourism objects stored in our database is gathered from various web sources in proprietary formats and is available in different languages in varying degrees of accuracy. The more detailed and coherent we can gather this data, the better recommendations we can provide.

In the current implementation we scrape information about objects in a semi-automatic fashion. First we crawl different web portals and normalise the gathered data into our custom ontology. In some cases the information available is not complete, therefore we have added an option to modify or add data manually. All the gathered information is stored in PostgreSQL (www.postgresql.org, [2011]) database. In section 4 we will describe the data discovery process in more detail.

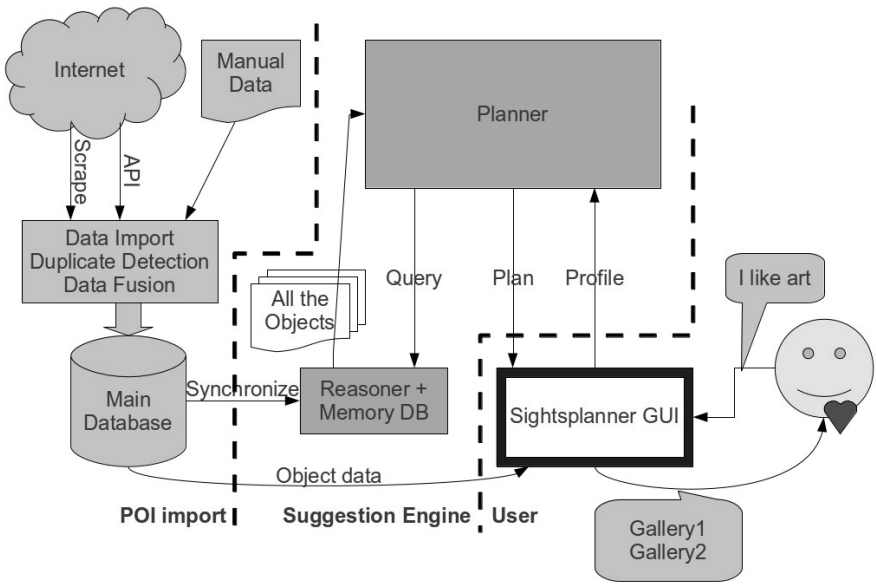


Fig. 2. System architecture

Object information is saved into the database using extended RDF triple structure (http://www.w3.org/TR/rdf-concepts/, [2004]). Every fact about each object is conceptually represented as an object, property, value triple (in standard RDF terminology subject, predicate, object). We have extended this structure by adding new fields for every fact. The main additional fields are the unique triple identifier, connection score, source, and timestamp. The unique identifier is automatically assigned by the database engine. The connection score indicates the probability of correctness of the given fact. The score is within the range [0; 1] and indicates the intuitive likelihood that the fact is true. The source field stores the name of the source

from where the fact is gathered (web portal name, manual insertion, generated etc.). The timestamp stores the date and the time when the fact was created.

The recommendation engine has to quickly handle a large amount of data: all properties of all objects and a relatively large ontology. Fetching all this data from the conventional relational database takes too much time. In addition to fetching the data we also would like to deduct new facts by applying rule-based reasoning. We have built a new reasoner-equipped fast in-memory database to serve our needs. This in-memory database is implemented in shared memory as a library, without any continuously running processes.

Data imported from the web pages is synchronised from PostgreSQL to the memory database once a day. Currently the synchronisation works only one way: all the data from persistent database is copied to the memory database. We employ our specialized reasoner to derive new types, opening times etc.

The recommendation process verifies suitable objects by loading the tourism objects from the in-memory database and applying an object filter: objects that do not have the required properties (coordinates, opening/closing times), belong to a different locale or are not accessible during the given visit duration are ignored. For each loaded object the recommender assigns a ranking score based on a matching between object types and preferences taken from the user profile.

The recommender runs the planning algorithm that produces a number of candidate solutions. These are ranked using an aggregate score which is based on the average score of the objects included, as well as how the time is divided between different categories of objects, when compared to the relative importance of the categories in the user's profile. The top ranking solution is returned to the user. The output of the planner is a timetable of objects, containing a Unified Resource Identifier (URI), arrival and departure time and method of travel from the previous location for each object.

Using the URI-s from the planner output, the recommender retrieves additional data, including visualisation components, descriptions in the user's language and contact information from the PostgreSQL database and creates the timetable representation.

4. Data Discovery

For our recommender system we are currently scraping six different web portals with completely different structures. For every portal we have written mapping rules to normalise raw data for our system. Some of the objects in the recommender are based on dynamic, rapid changing information such as one-time events (concerts, performances etc.) which are scraped from the web daily. Another category of objects is formed by places of interest (POIs) with mostly static information - nothing changes in months or even years. The POI objects are updated typically once a month.

The importer can connect to several web pages using different manually created scraper algorithms. Each scraper downloads the content of the web pages, finds the necessary information (for example using Xpath; http://www.w3.org/TR/xpath/, [1999]) to extract title, description etc.), normalises data for our system and sends it to

the Clusterer. The Clusterer compares the new object with the existing clusters by computing the similarity between the object and every cluster. From certain similarity value, the object is considered to belong to the cluster. In case of a match, the object and the cluster data is merged. When the match is not found, the object creates a new cluster. More details about clustering and merging is presented in the next sections.

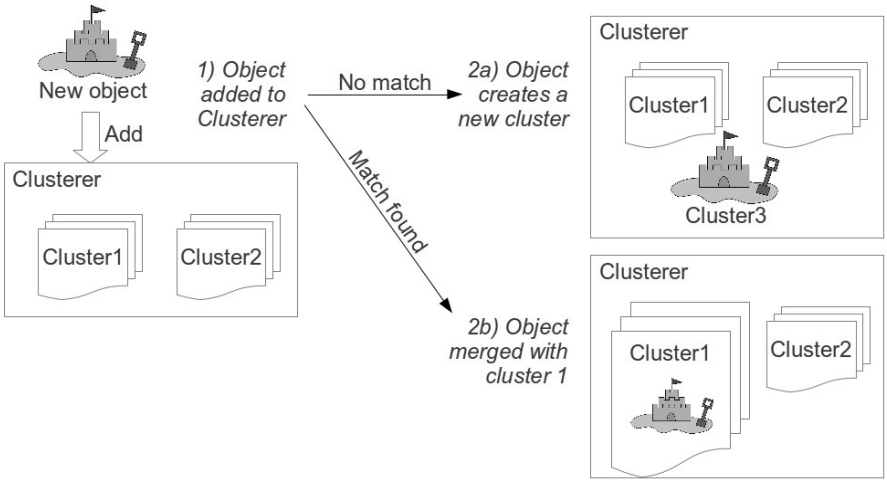


Fig. 3. Clustering of objects. An object is added to Clusterer. Case 2a: result when no matching cluster is found. Case 2b: result when a matching cluster is found and the object is added to that cluster

The recommendation process uses properties *latitude*, *longitude*, *opening_time*, *type*, *popularity_tourist*, *visit_time* to calculate the recommendation for the user profile. The coordinates are mandatory and cannot be empty. If the opening times are missing, we use the default values based on the object type (restaurants and museums have different default opening times). We extract types from the *title* and the *description* values. *Tourist popularity* indicates the average amount of tourists visiting the object in one day. If the value is missing, it is defaulted to 1. *Visit time* represents the recommended visit duration for the given object. The default value is 30 minutes, which can be modified by the tourist.

Our system stores information in various languages. Properties like *name* and *description* have language code along with the textual value. With multilingual data we often have a problem of missing translations. Usually the web pages we are scraping only have data in one language which makes it harder to get all the information in all the languages. To overcome the problem of missing multi-lingual information, we use the Clusterer. An object is scraped from one web page which only has data in Estonian. We add the object into the database. If the system finds an object from another web site, which is in English, the found object is compared to the existing objects. Based on the location, types and other properties, the Clusterer knows that the Estonian object and the new English object describe the same object. The objects are added into the same cluster. Data from the objects is merged and the original Estonian object gets an additional English name and description. In addition

to the multilingual textual information, the new scraped object can have additional metadata (phone number, homepage etc.) which can be merged to the existing cluster. If we cannot scrape information in other languages for an object, we run into the problem that some objects in our database do not have translations.

Table 1. An example of multilingual description

Object	Property	Value @ language	Score	Source
#church	#name	Oleviste Church @ en	1.0	#scraper1
#church	#name	Estonian name @ et	1.0	#scraper1
#church	#description	Translated English description @ en	0.8	#translator
#church	#description	Scraped English description @ en	1.0	#scraper2

We have implemented automatic translation functionality, which currently uses Google Translate API (http://code.google.com/apis/language/, 2011). We are translating title and description properties. The translated value is saved in the database with a lower probability score than for originally scraped values. In Table 1 we have presented a simple example with the language for every value and also the source for every fact. The scenario itself is as follows. We get the description of a church from one Estonian web page. The fact (description in Estonian) will get a score of 1.0. To be able to give some information about the object for foreign tourist, we translate the description automatically into English with a score of 0.8. At some point we will add a new web page scraper into our system which gets an English description for the same church. The scraped description would get a score of 1.0 and is therefore preferred over the automatically translated description (the fact with the source "#scraper2" is used to show English description). The scores for translation and different scrapers can be modified - instead of 1.0 and 0.8 other numbers can be used. For the web pages with low text quality, even the automatic translation can be preferred. In the following subsections we will present more details about different data discovery topics. First we describe keyword extraction from object multi-lingual description. After that we will present our data clustering and merging technique with some examples.

4.1 Keyword Extraction

The web pages usually do not provide enough meta-information about the object types or categories. Those are, however, crucial for suggesting tourism objects which match the user selected interests. Some web pages only offer a very general topic like "music". In our system we would also like to know the genre of the music, for example rock or classical. Therefore we have introduced a keyword extraction methodology in our system. We realize conceptual indexing in an automatic fashion. Conceptual indexing usually involves firstly, the detection of relevant concepts, and secondly, the calculation of a relevance weight for each detected concept. We follow the approach of projecting the ontology onto the document by extracting all the concepts the following way: as mentioned above, we store object information in different languages. To get object types, we analyse all the description properties in all the languages. Every language has its own keywords which match with possible

object types. The descriptions are parsed and all the matched keywords are found. The keywords are grouped by corresponding object types. The object will get all the found types with certain scores. The type which corresponds to more extracted keywords, gets a higher score and vice versa.

4.2 Data Clustering

We have already given a brief overview of clustering. In this subsection we will present more details about the process. First, we have to define the cluster. When data is scraped from the web into our system, everything is stored as an imported object (referred as "object" or "imported object"). Cluster itself is also a tourism object (referred as "cluster", but sometimes also as "object"). It represents one physical object or event. If the same church is defined in different web pages resulting, for example, three imported objects, the Clusterer should put all the imported objects into one cluster. Cluster can consist of one or more imported objects, it has all the information (facts) imported object has. In fact, the only main difference from an imported object is that cluster has an additional fact which just states that the given object is cluster. During Clusterer work, imported objects are left unchanged. Instead, clusters will be modified or added when needed. The recommendation process uses only cluster objects. After the import and clustering every imported objects should be in a cluster. If no similar objects are found, then the number of clusters will be the same as the number of imported objects.

Figure 3 describes how the Clusterer works. A new object (imported object) is sent to the Clusterer. The Clusterer compares the given object to every cluster. If no matching cluster is found, a new cluster is created and the object is added to the cluster. If a matching cluster is found, the object will be added to the matching cluster. Facts about the object and the found cluster are merged. We will describe data merging in subsection 4.4.

Comparing an object and a cluster is done by finding a similarity value between the two. Similarity in our system is a number in range [0.0; 1.0]. If the similarity is above a certain threshold, two objects are considered to be the same. More about the similarity calculation will be presented in the next subsection.

Comparing objects for similarity is expensive in terms of time consumption. To avoid comparing an object to every cluster, we try to find a smaller number of candidates for a new object. We have implemented location based candidate search: there is no point in comparing objects which are too far away from each other. Currently, with 3000 clusters and 300 new objects, the total running time of comparison process is around 10 minutes on a quad-core laptop. The number depends on the concrete data and the number of similar objects. If we had 5000 objects, it would probably take hours. The clustering process is usually run once a day. As mentioned, we regularly update only event data. More static POI data is updated only once a month. To improve the speed of clustering, we are looking into better candidate selection to get fewer candidates, different clustering algorithms and simpler comparison functions.

4.3 Similarity

Different data sources can describe the same physical tourism object or event. Usually those sources do not link data together. When importing the information, we end up

with several different objects in our system. Now it is up to us to decide, which of those objects are actually the same. We are using the *similarity* to decide that. The system calculates *similarity value* of two objects and if the value is above a certain threshold, the objects are considered to be the same and are put into the same cluster.

We have implemented several functions named *comparators* which deal only with certain characteristics. For example, we have a separate function for getting location similarity (whether two objects are similar based on the coordinates or address) or for titles (whether the titles of the objects are similar) and so on. Every function gives a result in range [0.0; 1.0]. All the *comparators* have *importance value*, which states how important the result of this function is in the total similarity. The calculation of *total similarity* can be viewed as a weighted mean. In our implementation, the formula is:

$$similarity = \frac{2 * location + title + type + properties + eventstart\ time}{6}$$

All five elements of the sum are different similarity functions. As can be seen, the *importance* for *location* is twice as high as for other elements. Importance values are currently set manually. We plan to use machine learning algorithms to learn better values. If the *similarity* is above 0.9, the system puts the objects into the same cluster.

In the next subsections we will describe the main *comparators*. Every *comparator* provides a *check function* which is used to test whether the corresponding comparator can be applied. For example, if one of the objects is not an event, *start time comparator* cannot be used as the similarity result would be 0. When the *comparator* is not used, it is not included in *total similarity* calculation. In the case of non-events, the *similarity* formula will have 5 as divisor (instead of 6).

Location comparator

One of the most important comparisons in our system uses object distances. Objects in database are described by *latitude* and *longitude*. Using coordinates, we calculate direct distance between the objects. We have defined three ranges: first is allowed distance without penalty, second one is allowed distance with penalty, and third one is distance not allowed. Currently, if the distance is smaller than 10 meters, the *location similarity* will be 1.0. If the distance is between 10 and 100 meters, the *similarity* will fall linearly: 10 meters would still yield in *similarity* of 1.0, where 100 meters would give 0.0. Distances over 100 meters would give 0.0 as a result.

Comparing only the coordinates might be misleading. As we scrape information from various sources, they can apply different coordinate systems (Google versus local Estonian system) or simply provide different location for the same object. For example, if the building is 200 meters long, one source can point to one side of the building as the other can point to the opposite side. If we compare the locations, we would get similarity 0.0. To overcome the problem, we also compare addresses.

Address comparison is done in two different ways. The first one is string comparison: we remove common words like “street”, “road” etc., extract the street name and the building number. If both the street part and the building part match, the similarity will

be 1.0. We also use Google maps API (http://code.google.com/apis/maps/, 2011) to get coordinates for the address. Google accepts address in various formats and returns one certain location for different address strings. We use coordinates provided by Google and compare locations as we did with our original coordinates.

Event start time comparator

Comparing only the location does not help us in the case when some events take place at the same building. Therefore we compare starting time of the events. Similarly to location comparison we have defined three ranges. If the time difference is below 15, the *similarity* will be 1.0. The difference between 15 and 60 minutes would decrease the *similarity* from 1.0 to 0.0 linearly.

Text comparator

With only *location* and *start time comparators* we would still experience problems, for example a cinema with three movies starting at the same time. The only way to make difference is to compare titles. Another example would be small shops or cafés inside a huge supermarket. Again, we need to compare titles.

As with location, we have implemented several different comparison functions. First is to compare all the words from objects' titles. We try to remove common words like “café”, “shop” to find out that “Cafè X” and “Cafè Y” are not similar although “cafè” matches. With different languages, it is often a problem to know the common words.

Another example is “McDonalds” and “McDonald's”. Depending on how the splitting into separate words is done, the problem still occurs when there are some symbols different (due to misspelling or special symbols). For the given case, we have introduced two string comparison algorithms: Levenshtein distance (also known as edit distance) and trigram comparison. They are both good in the case of misspelled symbols or when only small number of symbols are different. For example “Mc Donald's” and “McDonalds” are similar when using those methods.

Described algorithms do not solve all the cases. For example, in a shopping centre X there are two eating places: “McDonald's X” and “Pizzeria X”. The address is the same, they both are eating places and X part of the word is matching. The *title similarity* will not be 1.0, but it is high enough to push the *total similarity* above the threshold. Here we have an idea to use the knowledge, that X is a shopping centre (if it is present in our database) and consider this part of the titles as common word. Then we would compare “McDonald's” and “Pizzeria”, which would yield in similarity 0.

There are several cases where we still have to find a better solution. Often the title comparison is decisive whether the objects are the same or not. This is one priority in our current work.

4.4 Data Merging

After two objects have been put into one cluster, the cluster has to have representative information. An easy solution would be to just add all the facts from every imported object to the cluster. This may end up with objects which have two different names. To simplify the task of presenting data with user interface, we have implemented a merging process, which tries to remove obsolete information about the object.

During the clustering process all the facts from imported objects are added to the clusters. The merger component is run separately after clustering. Current implementation keeps the value with highest connection score for every property. Certain properties, for example all the imported types, are always kept.

Often values have the same probability score. For that case we have introduced a source priority list: value from the source which has a higher priority is preferred. Current approach has weaknesses and more sophisticated selection based on the context should be implemented.

5. Experiments and Results

On Sightsplanner.com we currently have about 2,200 scraped objects about Tallinn: 1200 POIs and 1000 events. The number of different events is about 150, but we keep every occurrence as a separate object (a play which takes place three times will be stored as three objects in our database). We scrape data from six different portals both in English and in Estonian. Every night we automatically update event information.

We have gathered a dataset of Tallinn eating places to measure the clustering precision. We have a total of 1,700 eating places (including duplicates). Total number of different objects when duplicates are merged is 800. Eating places are gathered from 5 different sources, which results in high number of comparison space (the number of possible object pairs to compare is above 10,000,000). Some objects have to be merged from 5 different data sources, where others only from 2 sources. Given the size of comparison space, we have gathered a different dataset to measure the clustering accuracy.

We have manually found all the duplicates in a dataset of about 3,500 scraped objects for another city Riga. Created dataset along with all the similar objects is used as a control dataset. We have used the same configuration of the Clusterer as with Tallinn data to get the results. In the given dataset, there are 77 duplicate objects, e.g. 154 objects will create 77 clusters, 2 objects in every cluster. The results with our current implementation are 87% precision (67 objects found from 77) and 87% recall (10 false negatives). 90% precision should be achievable. However, the main problem is title comparison, where we would take into account additional context information, like object type, location, common keywords and existing shopping centre names, etc.

Our next goal is to apply the same accuracy measure on Tallinn data. We have to create a subset of our data to limit the search space in order to be able to manually indicate all the duplicates. The selected subset can then be used as a training data for machine learning, which could optimize the similarity function parameters to improve clustering and merging of the objects.

6. Related Work

Geographical information extraction task is covered by Alves et al. (2009), where they describe KUSCO system which searches internet to enrich a Place of Interest (POI) data. Words from the results are used to discover the *Semantics of Place*. In the work they use natural language processing to extract nouns which will be used as concepts for entities. We are currently extracting keywords only from specified

sources and are looking into extending the lookup to the whole web, which would give us more information and also help to improve deduplication process.

In (Tré, 2010) a formal theory about POI deduplication and data merging is given. The article focuses mainly on *coreferent* (*clustered* in our terminology) POI data merging. The approach described in the paper is somewhat similar to our methodology. We do not calculate true and false probabilities. Instead we make use of the probabilities of our facts. In some cases (for example for descriptions) proposed techniques could be used to combine descriptions. As mentioned in the article, they do not deal with multilingual data. In one example, a translated title for a POI is lost due to merging. Our system has a requirement to merge data separately for every language.

Bleiholder and Naumann (2008) have given a good overview of data merging techniques. They described different approaches for different requirements. Along with every approach, they provided SQL examples. When using terminology from the article, we are using mostly *Take the information* and *Trust your friends* strategies. In addition we use fact probabilities to choose the one with the highest score.

A very similar approach is used in (Zheng, 2010) where the authors present a problem of finding near duplicate records in location datasets. They compare POI titles, addresses and types. As they have proposed a structure for the address field, they probably have rather high distances between objects. In our case, most of the objects are located in one part of the city. Our lowest hierarchy should be street, which is actually included in our address similarity calculation. Our current similarity measures are also near 0.9, which probably improve after we apply machine learning.

7. Conclusions and Future Work

In this paper we have described the tourism recommender Sightsplanner. A brief overview of the system architecture is given. The main focus of the paper is on the data discovery topics, where data importing, entity disambiguation and data fusion processes are described. We have presented the main functionality about different data processes along with some problems we have to overcome in order to improve the quality of the system and data itself.

In our current implementation we are using six different data sources. The goal is to be able to use an unlimited number of portals from the web. To achieve this, we have to improve the clustering process which would enable gathering information which for example, does not have coordinates. If we have an object in our database, we could just look for a more detailed description. Currently, the disambiguation depends strongly on location. We could also add information from user generated content by the title of the object: a tourist has visited one restaurant and writes about that in a blog. If we can identify the object in our database, we can add the new data into our cluster and merge all the facts.

We are already working on adding open linked data into our importer. There are several web portals which could be used to retrieve linked data (e.g. DBpedia, OpenStreetMap, etc.). Another future improvement will be user feedback on both data

and recommendations. Based on that we could evaluate the quality of our planner and information retrieval.

References

Alves, A., Pereira, F., Biderman, A. & Ratti, C. (2009). Place Enrichment by Mining the Web. In M. Tscheligi, B. de Ruyter et al. (Eds.), *Ambient Intelligence,* 5859: 66-77. Berlin: Springer.

Bleiholder, J. & Naumann, F. (2008). Data Fusion. *ACM Computer Surveys* 41(1): 1-41.

Luberg, A., Tammet, T. & Järv, P. (2011). Smart City: A Rule-based Tourist Recommendation. In *Information and Communication Technologies in Tourism 2011*. New York: Springer.

Tré, G. D. & Bronselaer, A. (2010). Consistently handling geographical user data: Merging of coreferent POIs. *Fuzzy Information Processing Society (NAFIPS), 2010 Annual Meeting of the North American* 1(1): 117-122. New York: IEEE.

Zheng, Y., Fen, X., Xie, X. et al. (2010). Detecting nearly duplicated records in location datasets. *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems* 1(1): 135-143. New York: ACM.

Acknowledgements

This work has been funded by the European Commission as part of the TEAM IAPP project (grant no. 251514) within the FP7 People Programme (Marie Curie).

Towards a Typology of Retail Cybermediation in Tourism Markets

Benjamin Rensmann[a]

[a] Department of Information Systems
University of Münster, Germany
bere@wi.uni-muenster.de

Abstract

Cybermediaries are organisations that bring customers and suppliers of tourism and travel products together on the Internet. They bridge the structural gap between highly complex tourism submarkets and customers facing various types of decision problems. Customers find a large variety of different cybermediary websites assisting them in composing a holiday or booking a flight ticket or a hotel room. The aim of this paper is the identification of different types of business models of retail cybermediaries in tourism markets and their value propositions for customers. A conceptual framework is developed which describes the position of cybermediaries in the cloud between complex supply-side market structures and a segmented demand-side. Furthermore, an analysis of 30 cybermediary websites has been conducted, the results of which point to three types of cybermediary business models.

Keywords: intermediation, cybermediaries, value proposition, typology, tourism markets.

1 Introduction

Intermediaries in markets are phenomena that have been researched from different angles. With the introduction and diffusion of electronic linkages among market participants (suppliers, intermediaries, customers), the traditional ways of doing business changed: The added value of intermediaries was questioned as direct interactions between suppliers of products and services and the customers or buyers became more and more feasible with increased electronic linkage. Broad predictions for disintermediation were made and the role of connecting suppliers and buyers was attributed to electronic markets (Bakos, 1991; Benjamin & Wigand, 1995; Malone, Yates, & Benjamin, 1987).

Sarkar et al. (1995) rebut these predictions by showing with a simple model that an intermediated transaction can incur lower transaction costs than a direct one (between supplier and buyer). This is enabled by new kinds of intermediaries which implement innovative, ICT-based business models. Sarkar et al. call them *cybermediaries* (Sarkar, Butler, & Steinfield, 1995). Subsequent studies have focused on the question what type of value cybermediaries create for different market participants by fulfilling certain roles with regard to market transactions and exchange relationships (cf. Bailey & Bakos, 1997; Giaglis, Klein, & O'Keefe, 2002; Rensmann & Klein, 2011).

The markets for tourism products were among the first markets that were heavily affected by these changes. Cybermediaries in particular discovered the potential of e-marketing and online sales as innovative means of bringing together suppliers and customers of tourism services. The main drivers for these shifts were changing customer expectations (more individualistic style of travelling, penetration of the Internet into homes) and an increasing competition between the different market

players. The changes resulted in a highly dynamic market environment with many new players employing innovative business models. At the same time, a market consolidation took place, pushing traditional travel agents out of the market or forcing them to adapt their way of doing business (E-business W@tch, 2006). After years of on-going change, what we see today is a large variety of cybermediary websites that address various travel needs of tourists and business travellers, next to traditional travel agents, which cater less internet-savvy customers and provide high service levels (Buhalis, 2003; Mayr & Zins, 2009).

The variety of cybermediary business models makes it difficult to develop an understanding of the way cybermediaries create value for customers in tourism markets. Obviously there are various types of retail cybermediaries which address different customer travel needs. Up to now, little work exists that systematically describes different types of cybermediary business models in tourism markets and their respective value propositions towards customers. Existing work about cybermediaries in the tourism industry rather focuses on cases of disintermediation and successful reintermediation (Garkavenko, Bremner, & Milne, 2003; Novak & Schwabe, 2009). In order to be able to evaluate different value propositions, a thorough and systematic understanding of different types of existing retail cybermediary business models and their value propositions to customers is a prerequisite. This leads to two research questions which are addressed in this paper, namely (1) what are different types of cybermediaries in tourism markets? And (2) how do the different types create value for customers, i.e., what is their value proposition?

The goal of this paper therefore is to identify types of retail cybermediaries in the tourism market and to describe their unique value propositions. This contributes to a better understanding of the value creation of retail cybermediaries that operate in the diverse markets for tourism and travel products. A conceptual framework is developed which describes the position of cybermediaries in-between the supply-side and the demand-side of the tourism market. The analysis of 30 different cybermediary websites points to distinct ways of value creation and thus different value propositions that are provided by different types of business models.

Section 2 describes the research approach and methodological background for this study in more detail. Section 3 introduces the conceptual framework. In section 4 the empirically derived types are described and illustrated with examples. Section 5 discusses the types and their value propositions. In section 5 a conclusion is given and the limitations of the paper are briefly explained along with some ideas for future research.

2 Research Approach

For the purpose of identifying types of retail cybermediaries and their distinct value propositions to customers, the business model concept is used as a theoretical lens. The value proposition is an inherent part of every business model, as it “gives an overall view of a company’s bundle of products and services” (Osterwalder, Pigneur, & Tucci, 2005). The value proposition describes exactly how the business or company creates increased economic utility for its customers and generates profits

from that (Gordijn & Akkermans, 2003). Types of cybermediary business models can be compared to categories, each of which contains exemplary business models that create customer value in a common way. The identification of types is similar to typological approaches in strategic management and organisational sciences. Typologies are top-down, theory-driven approaches that seek to conceptually derive types or classes of organizations. However, for a typology to be useful it needs to be conceptually well grounded in existing literature, but at the same time supported by empirical evidence. A well-developed typology can serve as a theory in itself (Doty & Glick, 1994).

For the identification of types of cybermediary business models and their value proposition, a two-fold approach is followed: Top-down as well as bottom-up. A conceptual framework is developed, which describes the market environment in which retail cybermediaries are operating. The framework (section 3) provides a conceptual background for the evaluation of cybermediary value propositions. For the empirical, bottom-up part of the study, 30 cybermediary websites were qualitatively analysed. Publicly available information on the respective websites and tests of the offered functionality provided the necessary data (actual bookings were not conducted). This research is to be understood as a pre-study, which aims at collecting initial evidence for different types of cybermediaries.

Although this evidence has been collected and analysed in a systematic way (see below), the types are deducted and described in a qualitative way, using the empirical evidence and the conceptual framework to argue for different types. The identified types and criteria can serve as a basis for future research regarding a typology of cybermediaries in tourism markets. Existing research on multi-channel strategies of retailers took a similar approach (Müller-Lankenau, 2007).

The sample was composed of retail cybermediaries focussing on different tourism markets and -products, as well as of cybermediaries that integrate two or more markets and their products on their websites. Only websites of cybermediaries are included which cover some aspect of a market transaction between primary suppliers and customers. The analysis of functionality was conducted along the line of the different phases of a market transaction, i.e. the information phase, the agreement phase and the settlement phase (Lindemann & Schmid, 1999). From the 30 cybermediary websites, 6 focussed exclusively on flight tickets, 10 on hotel rooms, 6 on rental cars, 4 were travel portals focussing on several tourism markets and 4 were specialized on tourism niche markets (accommodation in the Alps, hostels, organic hotels). Destination sites were not included, as they aim at providing general information about a specific region or destination and do not focus on specific tourism products.

For each of the websites the general market focus and the product specialisation were evaluated by checking which markets were addressed (e.g. flight tickets), what kind of products were specifically offered or promoted (e.g. nofrills flights, lastminute deals) and how many markets were combined (one versus several). The search functionality and the search scope were extensively tested and described (*information phase*). It was checked to what degree the different search functions that are provided by the cybermediaries help customers to state and flexibly vary their search criteria in

order to extend the search scope. Booking functionality was tested for the integration of third party services (*agreement and settlement phases*). Furthermore the websites were screened for information on destinations, travel recommendations and customer generated content. In addition to testing the provided functionality, company descriptions, FAQs and terms and conditions provided information on search strategies, sources of income for the cybermediary and company strategies[1].

3 A Framework for Analysing Cybermediary Business Models

Cybermediaries create value for consumers by making use of the middle position they occupy between the supply- and the demand-side of the market. This middle position creates specific challenges that have to be addressed. On the one hand, one of the main tasks of cybermediaries is the screening of the supply side and an aggregation of the available supply (i.e., tourism and travel products).

Cybermediaries then create value by matching the various offers of the supply-side of the market with customer's demand, supporting the actual booking and offering complementary services to customers. In doing so, they often rely on the services of others, e.g. third-party services, and sometimes collaborate with selected suppliers. These factors have to be taken into account. We call the gap the "cybermediation cloud", hereby allowing a wider focus for the analysis of cybermediary value creation. Figure 1 delineates this framework, the elements of which will subsequently be described.

3.1 Supply-Side Market Structure

The tourism industry consists of widely disparate groups of firms from different industries, offering a large variety of different tourism products (Eadington, 1991). For each of these products a specific market exists, e.g. for accommodation, flight tickets, rental cars, package holidays, luxury holidays and others. These markets can be further subdivided into more homogenous submarkets for specific branches of the respective market. The market for accommodation consists of submarkets for hotel rooms, hostel accommodation, holiday apartments, camping sites and others.

The market for flight tickets mainly consists of flights catered by big national and international airlines, low cost flights by nofrills carriers and charter flights. The different submarkets are characterized by a varying degree of concentration and supplier size. It goes beyond saying that the market segmentation is primarily an analytical tool to deal with the complexity of the market and that boundaries between market segments cannot always been drawn sharply. Services from submarkets might also be substitutable by services from other submarkets (e.g. short flight by train trip).

[1] A list of the different criteria that were evaluated for every website, as well as a list of the websites that were analysed can be provided by the author upon request.

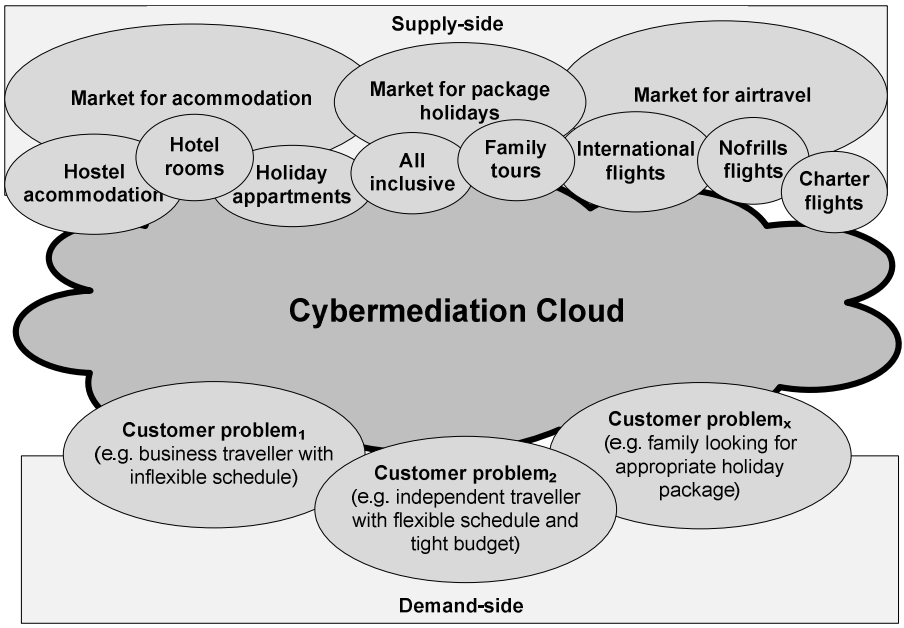


Fig. 1. Framework used for the analysis of cybermediary business models showing exemplary tourism markets, tourism and travel products, customer problems and the "cybermediation cloud"

3.2 Demand-Side

The complex supply-side market structure reflects different classes or types of customer problems, which result in customers looking for individual tourism or travel products or services (e.g. a rental car) vs. customers seeking a service configuration (e.g. transport + accommodation + entertainment). The preferences and needs of a customer are influenced by a range of factors like travel motivation or purpose, time preferences, family status, the available budget / income or other demographic factors (Bieger & Laesser, 2002; Dolničar, 2007). The high complexity of the supply-side market structures that customers are confronted with makes it difficult to find the right offer(s) which fit these needs. In a fragmented market with many small suppliers like the market for budget accommodation, the consumer has to evaluate many submarkets and offers. The complexity is further increased if a customer needs to book several single tourism services in order to compose a trip. The specifications of the different components need to fit together for the trip to make sense.

The expertise the customer brings along determines how well he can handle these problems. First of all, customers need to be familiar with booking and buying on the Internet. The identification of appropriate supplier and cybermediary websites requires knowledge, time and perseverance, especially if a customer is looking for the lowest price (Anckar & Walden, 2002). Furthermore, the more customers know about the structure of the tourism market, distribution strategies of suppliers and value propositions of intermediaries, the better they are able to find the product that fits

their needs. Another point is trust: Consumers might not feel comfortable booking a holiday apartment, as they do not know the supplier, or they do not trust a hotel's own description about rooms and services.

Altogether, this creates different types of problems for the customer. Independent travellers, who need to compose many single tourism products in a flexible way, are faced with a different type of problem than business travellers who need to book a flight plus hotel room for a certain destination. Customers seeking the lowest possible price either have to screen all potential supplier websites they know, or visit a cybermediary website of which they know it offers a complete market overview. Other customers might have the problem of not knowing where they want to go, so they need advice and inspiration, a task that travel agents usually fulfill. A business traveller, on the contrary, might only have restrictions regarding location, availability and class, but does not care much about price or touristic details. Families that want to go on holidays have special requirements regarding accommodation and activities.

3.3 The Cybermediation Cloud

Werthner and Klein (1999) describe the space between suppliers and consumers in the tourism sector as a "communication cloud" that has to be crossed by both sides, suppliers and consumers of tourism and travel services (Werthner & Klein, 1999). This domain is an ideal playground for cybermediaries of all different kinds. Customers need to identify the specific suppliers on the other side that provide the kind of services that together fulfil their needs. Cybermediaries support or fulfil the necessary communication and information exchange between the supply and demand sides. Customers can choose between many different cybermediaries offering similar services. The intense competition between cybermediaries leads to a lot of variety of cybermediary websites.

Cybermediaries also integrate services of 3rd party providers into their own service portfolio, e.g. of payment providers and insurances. Furthermore, cybermediary services often build upon the services of other cybermediaries in the cloud. In these cases, cybermediaries search the databases of other cybermediaries and integrate their offers into their own search results. Customers are then referred to that intermediary for the actual booking process. Cybermediaries in the market for tourism services are thus embedded in a net of relationships with suppliers of tourism services, third party providers and other cybermediaries. The business models of cybermediaries not only address customers' problems, they are also aligned to specific issues of the supply-side market structure, such as the various distribution strategies of suppliers. Cybermediaries can be either biased towards the supply-side, towards the demand-side or they can act as neutral market agents (Klein & Teubner, 2000).

4 Types of Cybermediary Business Models

In this section, the results of the analysis of cybermediary websites will be presented. At first glance, two major trends can be observed: On the one hand there are cybermediary websites (CM) that are specialised and have a focus on a single submarket (SM). Some of them conduct the actual booking and include services of service providers (SP) like travel insurances, some merely forward the customer to the

respective supplier website or to other cybermediary websites. On the other hand, there are websites that integrate two or more submarkets and provide a lot of information about destinations and tourism products. The analysis of how cybermediary websites relate to customers' problems (CP) revealed a more detailed picture and resulted in the identification of three distinct types of cybermediary business models. The types will subsequently be described. Figure 2 provides a schematic overview of the different types. Table 1, at the end of the section, gives an overview of the distinctive characteristics.

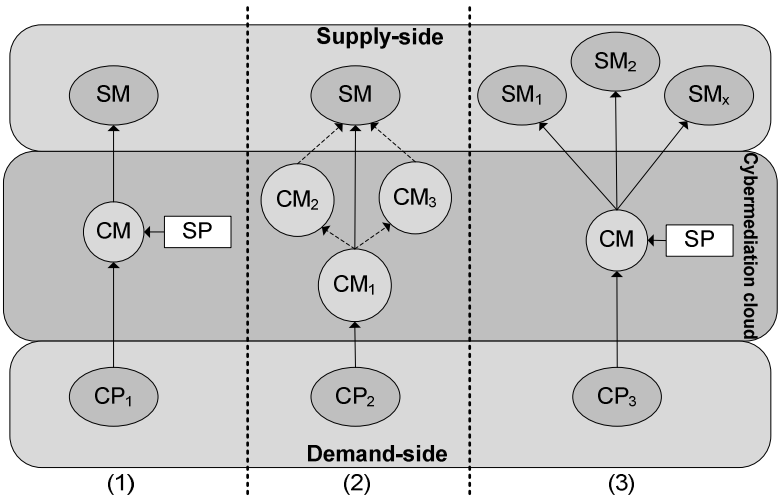


Fig. 2. Schematic overview of the three different types of cybermediary business models: (1) Booking Sites (2) Comparison Sites and (3) Online Travel Agencies

4.1 Booking Sites

This type of cybermediary focuses on one single submarket and enables the booking of tourism products on that specific submarket. The offered search functionality includes all important search criteria like dates, locations and number of people travelling. This creates a good market overview for the customer, however, not all suppliers and special deals might be searched. Only suppliers that either provide an interface to their own customer reservation system or are connected to the interface of the cybermediary are included. The actual booking process, i.e. payment and settlement, is always supported by this type of cybermediary. During the booking process, additional third party services like travel insurance or activities can often be added. The income of this type is either based on a commission which has to be paid by the supplier or on a booking fee which has to be paid by the customer.

Tbooker is an example for this type of cybermediary[1]. Tbooker maintains a search engine for flight tickets and also enables the actual booking of the ticket. When searching for a flight, only airlines are searched which enable the booking of tickets by cybermediaries (this excludes e.g. Ryanair). During the booking process, several

[1] See: http://www.tbooker.com/

kinds of insurances as well as a rental car can be added to the actual flight. Tbooker charges a ticket fee which has to be paid by the customer. Booking.com is another example operating in the hotel submarket. Booking.com enables customers to search for hotels in specific destinations and offers binding reservations or bookings (to be paid in advance).

4.2 Comparison Sites

Cybermediaries implementing this type of business model aim at creating the highest possible market transparency for their customers. Extensive search functionality is offered, which enables the customer to precisely specify what he is searching for. In addition, search criteria are often flexible in a way that helps the customer finding deals which might e.g. not perfectly fit his schedule, but which are also acceptable because of a much lower price. This type of intermediary focuses on one single submarket and often searches on other cybermediary websites, too. The actual booking process is not carried out by the cybermediary. Instead customers are forwarded to the respective supplier- or intermediary-website, where the booking can be done. This type of cybermediary creates an income by referring customers and by placing ads on their websites.

Qfly[1] is an example for this type. Qfly is a new and innovative cybermediary website which covers the European market for nofrills flights. Qfly claims to offer a complete market overview, including all nofrills airlines which are operating routes in Europe. It offers sophisticated search functions which enable the customer to flexibly describe his travel needs (e.g. to include several departure and destination airports, time periods, budget, etc.). A special feature on Qfly is the possibility to specify a budget and look for possible routes which fit into that budget. Qfly simply forwards customers to the website of the respective airline, where the actual booking takes place. Trivago[2] is a comparison site which enables searches for hotel accommodation. Trivago thereby conducts searches on more than one hundred hotel booking websites, i.e. performs a meta-search on booking sites for hotels. Here again the consumer is forwarded to the respective website, where the hotel can be booked.

4.3 Online Travel Agencies

The online travel agency type of cybermediary business model integrates several submarkets for providing one-stop shopping to customers. Customers can search for just a single tourism product (e.g. a flight) or execute a combined search for several products, e.g. flight, hotel and rental car, or any combination of these. In addition, online travel-supermarkets offer a large variety of other tourism products on their websites like package holidays, cruises, last-minute deals, holiday apartments and others. Furthermore, online travel-supermarkets provide information on activities at the destination and give advice for booking the trip that best fits the customer's needs (Expedia e.g. provides a "beach-inspirator"). The actual booking is also conducted by the cybermediary.

[1] See: http://www.qfly.com
[2] See: http://www.trivago.com

There are plenty of well-known examples for this type. One is Expedia, which claims to be the world's largest travel business. On the website of Expedia, customers can book flights, hotels, rental cars (or combinations of these), package holidays, cruises, last-minute deals, family holiday trips, holiday apartments and other tourism products. In each case, the different products can be booked on Expedia, which offers additional travel insurances during the booking process. The search screens are adapted to the specific products, i.e. for package holidays other search criteria can be specified than for flights. In addition, Expedia helps customers in composing their trip by providing information on destinations, activities and travel arrangements. The beach-inspirator and the city trip-inspirator are specialized search screens on which customers can "play around" in order to find interesting potential destinations.

Table 1. Main characteristics of the three cybermediary types

	Booking sites	**Comparison sites**	**Online Travel Agencies**
Market focus	Focus on 1 submarket	Focus on 1 submarket	Multiple submarkets included
Booking support	Booking conducted by cybermediary	No booking, forwarding to supplier or cybermediary	Booking conducted by cybermediary
Search scope	Search limited to focussed submarket and side market(s)	Search exclusively in focussed submarket	Integrated search in several submarkets
Search depth	Only direct suppliers of tourism products are included	Other cybermediary websites are also searched	Only direct suppliers of tourism products are included
Search functionality	Standard search functionality	Extensive search and filtering functionality	Special search masks for integrated search, standard functionality
Source of income	Booking fees, commission paid by suppliers	Mostly ads and paid referrals	Partly booking fees, partly commission paid by suppliers
Examples	Tbooker, fluege.de, booking.com, hotel.de, carrentals.com	Qfly, Trivago, Swoodoo, Skycanner, Kayak	Expedia, Travelocity, Opodo, lastminute.de

5 Discussion of the Findings

The types that were identified in the previous section bridge structural gaps that originate from the complex market segmentation on the supply-side and the various customer problems on the demand-side. This is reflected in the functionality and scope of different cybermediary websites and results in different value propositions towards customers.

The first type, *booking sites,* are restricted to a specific submarket and thus provide a limited set of products, e.g. hotel accommodation. These sites help customers to get a quick overview of a certain submarket and to save time by enabling an immediate

availability check and booking. This is especially useful in highly fragmented submarkets, as customers would have to spend a lot of time on searching and evaluating different suppliers directly. Because the booking sites type provides a platform for the actual payment and settlement, it acts as an independent third party that is trusted by the customer. For availability check and the booking, however, booking sites rely on the offers they receive from suppliers or GDS. This means that customers might not always get an exhaustive market overview in terms of available suppliers, routes or prices. Booking sites thus support all three phases of a transaction, with a focus on the agreement and settlement phases which describe the actual booking.

Comparison sites strive to provide a complete market overview by including as many suppliers and booking sites as possible. While focussing on a single submarket, comparison sites want to make sure that customers find the lowest possible price for a certain tourism product. The extensive set of search functions gives customers a high flexibility concerning schedule, route and destination preferences. This enables a variation in travel preferences in order to find an even lower price. The disadvantage is the lack of booking facilities. While price and availability is usually checked by comparison sites, customers are forwarded to supplier or booking websites for the actual booking. This increases the time needed to conduct a booking, as customers have to deal with the specifics of yet another website. Comparison sites thus focus exclusively on the information phase of a transaction.

Online travel agencies, by combining several submarkets and enabling the combined booking of different products, give their customers the possibility to compose a complete holiday trip. The market coverage for a specific submarket is comparable to that of a booking site. By creating the convenience of one-stop shopping, customers do not have to match the offers of several websites so that they fit their travel preferences. Online travel agencies concentrate a lot of information about different submarkets and products in a very customer-accessible way. A lot of information on destinations, tourism suppliers (e.g. tour operators) and activities assists the customer to find what he wants. This serves customers lacking inspiration on where to go and what to do, but also helps larger groups of travellers, in particular families, to find appropriate accommodation. Its value proposition thus very much resembles the traditional travel agent from around the corner. Online travel agencies address all three phases of a transaction. Compared to booking sites they partly put more emphasis on the settlement phase, by providing extensive hotline support and own reviews of suppliers (e.g. hotels), which provides a kind of guarantee that these suppliers really provide the kind of service that they advertise.

The different types thus solve several structural problems. They map the supply-side submarkets and market segments with a variety of customer segments and problems. The schematic overview in Figure 2 only indicates this structural bridging. Not shown in the Figure are the different customer segments: Customer from different segments may have the same kind of decision problem, e.g. the booking of a hotel for a specific date / city or the composition of a trip. For customers the question is not only which type of cybermediary website best fits their needs, but also which website to pick out of many. There is an intense competition going on between cybermediaries, with cybermediaries maintaining different brand names which serve the same market. Next

to visiting a cybermediary website, customers also have the choice to book directly on the websites of tourism suppliers. Cybermediaries thus not only compete against each other, but are also facing an increasing competition from direct distribution strategies of suppliers.

6 Conclusion, Limitations and Future Research

The goal of this paper was to identify different types of retail cybermediaries in different tourism markets and to describe their value proposition to their customers. After having introduced the theoretical background on intermediation and the tourism sector in the Introduction, section 2 presented the research approach that was followed. The framework that was used for analyzing different business models was presented in section 3. In section 4, the three different types of cybermediary business models that were identified are described. Section 5 discusses the findings with regard to different value propositions.

This study has several limitations. First of all, the website analysis that has been conducted might not encompass all existing types of intermediary business models that exist. However, this research approach should be regarded as a pre-study which points out existing *types* of business models. What further needs to be noted is that there are differences within the types, too. There are booking sites for various purposes, ranging from general hotel booking sites via websites offering accommodation in the Alps to websites for booking rental cars. The point of this paper was however to indicate different types of cybermediary websites based on structural characteristics of the tourism market. Lastly, in order to fully analyse the business model of a cybermediary, more data is needed which gives an insight into sources of income, connectivity to GDS and suppliers and the size and usage of different distribution channels.

References

Anckar, B. & Walden, P. (2002). Self-Booking of High-and Low-Complexity Travel Products: Exploratory Findings. *Information Technology & Tourism* 4(3/4): 151–165.

Bailey, J. P. & Bakos, J. Y. (1997). An Exploratory Study of the Emerging Role of Electronic Intermediaries. *International Journal of Electronic Commerce* 1(3): 7-20.

Benjamin, R. & Wigand, R. (1995). Electronic Markets and Virtual Value Chains on the Information Superhighway. *Sloan Management Review* 36(2 (winter)): 62-72.

Bieger, T. & Laesser, C. (2002). Market Segmentation by Motivation: The Case of Switzerland. *Journal of Travel Research* 41(1): 68-76.

Buhalis, D. (2003). *eTourism. Information technology for strategic tourism management*. Harlow, UK: Pearson Education.

Dolničar, S. (2007). Market Segmentation in Tourism. In A. G. Woodside and D. Martin (Eds.), *Tourism Management: Analysis, Behaviour and Strategy* (pp. 129-150). Oxfordshire, UK: CAB International.

Doty, D. H. & Glick, W. H. (1994). Typologies as a Unique Form of Theory Building: Toward Improved Understanding and Modeling. *Academy of Management Review* 19(2): 230-251.

E-business W@tch. (2006). *ICT and e-Business in the Tourism Industry* (Sector Report No. 8/2006). Brussels, Belgium: European Commission.

Eadington, W. (1991). Economics and tourism. *Annals of Tourism Research* 18(1): 41-56.

Garkavenko, V., Bremner, H. & Milne, S. (2003). Travel Agents in the "Information Age": New Zealand Experiences of Disintermediation. In A. J. Frew, M. Hitz and P. O'Connor (Eds.), *Information & Communication Technologies in Tourism 2003: Proceedings of the International Conference in Helsinki, Finland, 2003*. Vienna, Austria: Springer-Verlag.

Giaglis, G. M., Klein, S. & O'Keefe, R. M. (2002). The Role of Intermediaries in Electronic Marketplaces: Developing a Contingency Model. *Information Systems Journal* 12(3): 231-246.

Gordijn, J. & Akkermans, H. (2003). Value Based Requirements Engineering: Exploring Innovative e-Commerce Ideas. *Requirements Engineering Journal* 8(2): 114-134.

Klein, S. & Teubner, R. A. (2000). Web-based Procurement: New Roles for Intermediaries. *Information Systems Frontiers* 2(1): 19-30.

Lindemann, M. A. & Schmid, B. F. (1999). Framework for Specifying, Building, and Operating Electronic Markets. *International Journal of Electronic Commerce* 3(2): 7-21.

Malone, T. W., Yates, J. & Benjamin, R. (1987). Electronic Markets and Electronic Hierarchies. *Communications of the ACM* 30(6): 484-497.

Mayr, T. & Zins, A. H. (2009). Acceptance of Online vs. Traditional Travel Agencies. *Anatolia: An International Journal of Tourism and Hospitality Research* 20(1): 164–177.

Müller-Lankenau, C. (2007). *Multikanalstrategien im stationären Einzelhandel. Eine empirische Untersuchung in der Kommunikationsbranche*. PhD-thesis, University of Münster, Germany.

Novak, J. & Schwabe, G. (2009). Designing for Reintermediation in the Brick-and-Mortar World: Towards the Travel Agency of the Future. *Electronic Markets* 19: 15-29.

Osterwalder, A., Pigneur, Y. & Tucci, C. L. (2005). Clarifying Business Models: Origins, Present, and Future of the Concept. *Communications of the Association for Information Systems* 16(1): Online Publication.

Rensmann, B. & Klein, S. (2011). Sarkar, Butler & Steinfield (1995) "Intermediaries and Cybermediaries" Revisited: A Review and Identification of Future Research Directions for Intermediaries in Electronic Markets. *Proceedings of the 24th Bled eConference, Bled*, Slovenia.

Sarkar, M. B., Butler, B. & Steinfield, C. (1995). Intermediaries and Cybermediaries: A Continuing Role for Mediating Players in the Electronic Marketplace. *Journal of Computer-Mediated Communication* 1(3): Online Publication.

Werthner, H. & Klein, S. (1999). *Information technology and tourism: A challenging relationship*. Vienna, Austria: Springer-Verlag.

Network Analysis of the Austrian eTourism Web

Roland Piazzi[a],
Rodolfo Baggio[b],
Julia Neidhardt[c], and
Hannes Werthner[c]

[a] Graz University of Technology,
Graz, Austria
roland.piazzi@iaik.tugraz.at

[b] Master in Economics and Tourism
Bocconi University, Milan, Italy
rodolfo.baggio@unibocconi.it

[c] E-commerce Group – ISIS
Vienna University of Technology
{julia.neidhardt, hannes.werthner}@ec.tuwien.ac.at

Abstract

Network analysis methods have gained much attention in the last years and have provided a wealth of insights into the structural and dynamic properties of many systems. Here these methods are applied to the study of a destination's tourism webspace. This exploratory analysis aims at showing how these techniques can be used and what outcomes can be obtained. After a short introduction to network analysis and a brief review of the literature, the network analysis of the Austrian eTourism Web is described in detail. The data collection methods are explained and the characteristic network parameters are calculated. Then the results are discussed and interpreted.

Keywords: Network analysis, eTourism, Austria, World Wide Web.

1 Introduction

The Web has dramatically transformed the tourism industry and its landscape, and this even more as foreseen in the optimistic early days of eTourism in the 1990s (Poon, 1993; Schertler et al., 1994; Sheldon, 1997; Werthner & Klein, 1999). It is common knowledge that nearly all tourism players operate a web site, and that a very large proportion (approx. between 30 and 50%, dependent on region and / or reference, e.g., PhoCusWright, 2011) of the turnover generated by the sector originates from online activities. It is clear then how strong the impact of Internet applications has been on tourism and why this industry is one of the prominent application fields in the Web, also due to its specific features. This reciprocal relationship was also already predicted in the early days of eTourism (Werthner & Klein, 1999).

The Web induced transformation of tourism happened on all levels, i.e., individual, enterprise related and market level. Consumers have changed their search and booking behavior, and benefit from a supposedly more "transparent" market, which allows them to compare different online, offers and services in a convenient way. They become *actively interactive* players in the field. Tourism companies and

organizations use the Web to directly access the electronic market, to promote themselves and to sell their products and services. A further effect of the virtualization of tourism activities is that the distinction between (types of) players blurs. Suppliers become intermediaries and vice versa, and consumers became part of the supply chain, for example using Web 2.0 applications and generating contents, information and services. These phenomena point at the third level of change, i.e. the structural one. One can observe permanently new services as well as players in the field, and at the same time a somehow opposed *winners-take-it-all* (or most) trend with a high level of concentration. At this level the Web shows a dialectic relationship of order and un-order.

But one should note that information technologies not only facilitate or force these changes, but also offer a historically unmatched possibility to observe the world and its changes. Since nearly all human activities are moving to the web and the web itself is a transparent medium, we have the opportunity to observe both activities and the infrastructure used, the Web becomes a mirror of the world. Data are easily available, in a cheap and fast manner. This will also lead to a change in the way economic and social studies will be performed (Hendler et al., 2008).

This second consideration is also of importance for the tourism sector. Up to now there has only been little quantitative research focusing on structural issues, using for example network analysis. Network theoretical methods allow a structural analysis of the relationships and connections between the respective market participants, and may provide good insights, as already done in numerous other fields (da Fontoura Costa et al., 2011). In this paper such a structural analysis for the country of Austria, which is based on Piazzi (2011), is presented.

The paper is organized as follows. In section 2 the network analysis concepts that are used in the paper are briefly described. Furthermore, a short overview of the state of the art of network analysis in tourism and eTourism is given. In section 3 the data collection methods and the analyzing tools are described. The results of the analysis of the Austrian eTourism web are presented and discussed in section 4. Finally, section 5 is dedicated to further discussion and concluding remarks.

2 Theory

In this section a short introduction to network analysis is given. For a thorough review of all the concepts and metrics described see, e.g., da Fontoura Costa et al. (2007).

A network is usually modeled as a graph $G=(V,E)$. The set V is called *nodes*; the set E consists of *edges* that link pairs of nodes. If two nodes are joined by an edge, they are called *neighbours*. In *an undirected graph* the edges do not have an orientation. In *a directed graph* each edge has an origin and a destination. The *degree* $d(v)$ of a node v in an undirected graph is the number of its neighbours. In a directed graph, *in-* and *out-degree* of a node are separately considered: in-links are edges that end in v, out-link are edges with origin v. The *average degree* of graph G is the arithmetic mean of all degrees. A *path* is a sequence of nodes such that two consecutive nodes are joined by an edge; and its *length* is its numbers of edges. The *distance* $d(v,w)$ between two nodes v and w is the length of the shortest path between them. If there is a path from a

node *v* to a node *w*, they are called *connected*. These quantities, when considering directed graphs, take into account the edges' orientations. The graph's *density* is defined as the number of edges divided by the maximum possible number of edges (those present if there were an edge between each pair of nodes). The *clustering coefficient* of a node *v* reflects the probability that two arbitrary neighbours of *v* are connected by an edge. The clustering coefficient of the entire graph is defined as the arithmetic mean of the clustering coefficients of all nodes.

Since network analysis methods should provide a better understanding of the "real world" structure, some concepts that facilitate a richer interpretation have been proposed. In this respect, a very important category is the class of so-called *centrality measures*. These try to formalize the idea that in many instances some nodes or edges play a more important role than others, hence they should be considered as more central. For the study of the eTourism networks three commonly used centrality measures are: degree centrality, closeness centrality and betweenness centrality. The *degree centrality* $C_D(v)$ for node *v* is defined as the number of edges it is connected to: $C_D(v)=d(v)$. In a directed graph two kinds of degree centrality are usually distinguished, namely *in-* and *out-degree centrality*. The *closeness centrality* $C_C(v)$ for a node v is defined as the reciprocal value of the sum of all distances between *v* and each other vertex *w*: $C_C(v)=1/(\sum d(v,w))$. In a directed graph *in- and out-closeness* can be considered separately. The *betweenness centrality* $C_B(v)$ for node *v* is defined as the number of shortest path between all other pairs of nodes that run through node *v*: $C_B(v)= \sum(\sigma_{uw}(v)/ \sigma_{uw})$. Here σ_{uw} denotes the number of shortest paths between nodes *u* and node *w* and $\sigma_{uw}(v)$ the number of shortest path between those nodes that run through *v*. Usually the values calculated for these metrics are normalized.

One property that many "real world" networks have in common is the so-called *power-law degree distribution*; i.e., the degree distribution of the network's nodes can be approximated by a function of the form $p(k)=ck^{-x}$, where k denotes a node's degree and c and x are real positive constants. Hence, in such a network the majority of nodes have a very low degree while a very few have a remarkable high number of neighbours. Such networks are also called *scale-free*. Another common property is the so-called *small-world property*. The term expresses that the average distance within such a network is relatively short.

It is well known that the World Wide Web exhibits these properties (Newman, 2010). It can be modeled by a directed graph whose nodes are Web pages or websites and whose edges are the hyperlinks between them. According to the literature, the average degree of a Web page is 7.2; and there is a power-law degree distribution for both the Web pages' in-degrees and out-degrees (with an exponent x=2.1 for the in-degrees and x=2.7 for the out-degrees). In addition, the World Wide Web seems to be a small world network with an average path length of 16.18.

As mentioned above, by now network analysis has been little applied in the field of tourism. The few, existing studies are rather different in nature and hardly comparable. Roughly, the existing work can divided into three categories: a. papers and studies dealing with the analysis of web network structures, where also comparisons between electronic and real world networks are made; b. network analysis applied to physical infrastructure networks such as roads or transportation

networks; and c. works dealing with the relationships between individuals or organizations engaged in tourism; e.g. actors developing tourism policy strategies.

The first category covers studies that treat web networks and the reflected "real world" at the same time and compare them. Real world networks consist here of tourism businesses like hotels or other accommodations that are connected through business relations, or just through information flows. At the other side, technological networks can be identified as networks that consist of the websites of tourism businesses that are connected through hyperlinks. Here, the most comprehensive studies were conducted by Baggio et al. with the analysis of Elba tourism destination (Baggio, 2007; da Fontoura et al., 2009; Baggio et al., 2010). A comparison between different destinations was performed by studying the Web spaces of the Fiji Island and of Elba (Baggio et al., 2007). Both islands show similar features. Furthermore, a study of the Web network of Italian travel agencies (Baggio, 2006) showed again similar connectivity characteristics. Finally Scott et al. (2008) studied the networks of four differently developed Australian tourism regions. The most interesting result of this series of works is the topological similarity between the real and the virtual networks, which may allow, for example, to consider the Web network of a destination as a significant sample of the real relationships between the destination actors.

In the second category, Shih (2006) pursued a different approach. Also network analysis was used, but here the network's nodes are the towns and villages of the region Nantou (Taiwan) and the links are the transport facilities. The overall aim of the study was to find out where new tourism infrastructure should be built and of which type it should be. Other studies have dealt with airlines, train and navigation networks. The main aim was to find possible reconfigurations of the infrastructures to optimize travel times or to reduce congestions and bottlenecks (da Fontoura Costa et al., 2011). In the last category, Pforr (2005) focused on the network of actors developing tourism policy strategies (thus being only slightly related with this paper's line of work).

This paper can be classified into the first category, but it considers also nodes' centralities, which were not mentioned in the other works belonging to this category

3 Methods

The idea of this study was to find the big tourism players (like HRS, Booking.com, Expedia etc.) that have the most impact on the business. They were expected being superior compared to the destination management organizations. However the study showed something different, as we will see later on.

The data for analyzing the network structure of the Austrian eTourism Web was gathered by a simple Web Crawler. This "search engine" was developed in Java by one of the authors and used a MySQL Database to store the relational information of the websites. The Web Crawler worked as follows: a. read a website's homepage, search through this page for links, save the links pointing to the same website in a list; save the links pointing outside this website in another list; b. read the next page of this website and save its links; c. do this up to a predefined pages-per-site limit, then go to

the next website. This procedure was done until a significant amount of websites was collected. A pages-per-site limit of 5000 web pages was chosen to reduce the crawl time. This limitation does not reduce the quality of the results, because on the one hand it turned out that most of the outgoing links were found at the beginning of the crawl - probably due to the websites' structures; and on the other hand, many of the websites did not even reach this amount of pages.

When modelling the eTourism web it had to be decided whether: a. the websites or the web pages should represent the network's nodes; b. a weight should be assigned to the number of links between two websites; c. dynamic web pages, i.e. pages where the content depends on parameters within their URL, should be taken into consideration. Regarding point a. it was decided to define the entire websites as the network's nodes rather than the single web pages. Since the actors of the network are tourism enterprises, they are better identified by the whole website than by just one web page. Links between two websites were not weighted assuming that a high number of links between two sites does not necessarily reflect a stronger relation, e.g. if there are many Facebook like-buttons or links to the website's developer on each page. Finally, regarding point c., also dynamic web pages were taken into account, because of the assumption that in this way more outgoing links are found.

Table 1. Details about the four crawls executed: starting website(s), number of crawled websites, number of tourism related sites and indirectly tourism related sites within the network

	Crawl 1	**Crawl 2**	**Crawl 3**	**Crawl 4**
Starting website(s)	austria.info	tiscover.com	austria.info tiscover.com touris-net.co.at touristinformation.at alpenstadt.com urlaubinoesterreich.net urlauboesterreich.com thermen.travel cusoon.at wienfuehrungen.com	austria.info blog.austria.info tirol.at genussland.at oberoesterreich.at vorarlberg.travel neusidlersee.com oetztal.com innsbruck.info burgenland.info
Number of websites crawled	2116	1385	8276	2395
Tourism related websites	58%	43.5%	37.2%	69.2%
Indirectly tourism related websites	7%	6.7%	7.3%	15.8%

As the Web Crawler is not able to distinguish whether a website belongs to the tourism sector or not, the starting points for the program had to be selected carefully, in order that these starting points mainly link to tourism enterprises' websites. So, four crawls were executed to improve the quality of the result. In Table 1 some details about the crawls are provided. Using the results of the previous crawls, as starting

points for the fourth crawl, pages were selected that linked many tourism related sites (according to Werthner & Klein, 1999). Furthermore, to include a site in the Austrian eTourism network, the enterprise behind the associate website had to have its business mainly in Austria.

For further analysis, the result of the fourth crawl was chosen, because it contained the highest fraction of tourism enterprises. The network uncovered by this crawl consists of 2395 nodes of which 69.2% can be considered as tourism related websites and 15.8% as indirectly tourism related websites. The latter are sites of enterprises that are indirectly depending on tourism, e.g. that offer products, which typically are consumed by tourists.

So, the resulting network can be regarded as a significant sample of the Austrian tourism web space also taking into account that the Austrian tourism landscape consists of 91600 enterprises (www.austriatourism.com/xxl/_site/intde/_area/465222/_subArea/465274/_subArea2/476161/index.html [Aug. 27, 2011]). For analyzing the data, UCINET by Borgatti, Everett & Freeman (2002) was used; NetDraw, which is integrated in UCINET, was used for the network's visualization.

4 Results

In this section we present the results of the analysis of the Austrian tourism web network. First, a graphical representation of the network is given. Afterwards, the characteristic network parameters are discussed.

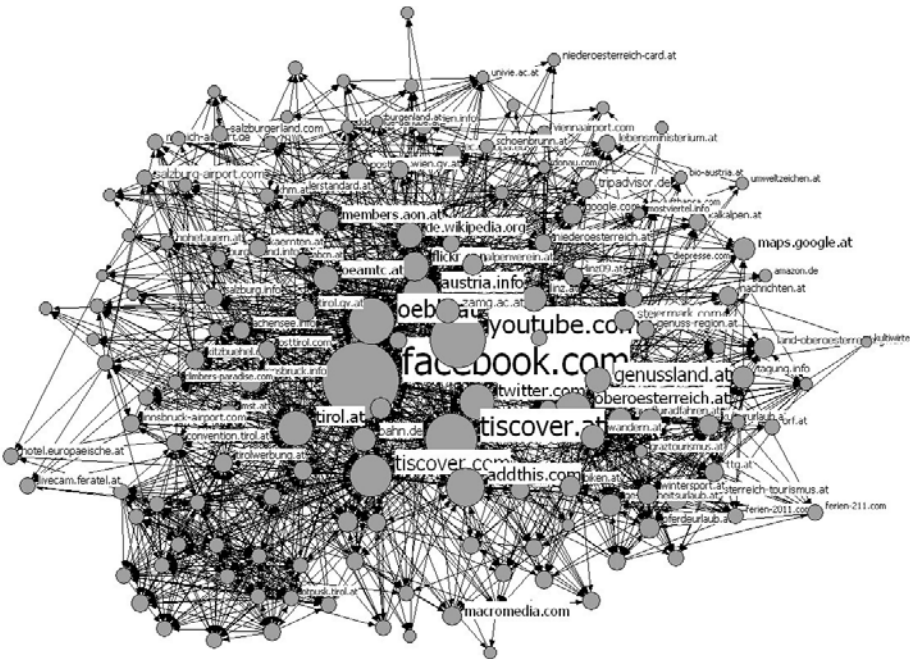


Fig. 1. Network of the Austrian tourism web, nodes with in-degree>=25

4.1 Graphical Representation of the Network

Figure 1 shows the Austrian eTourism network with nodes that have an in-degree of at least 25 (for better visualization). The size of the labels and circles is scaled by the corresponding in-degree value. As we will discuss in the next section, Facebook.com and Tiscover (.at and .com together) are the biggest nodes.

Also other social media sites like Youtube.com, Twitter.com and Addthis.com are linked quite often, as well as destination management platforms like Austria.info, Tirol.at, Genussland.at and Oberoesterreich.at, and also the site of the Austrian rail company Oebb.at.

4.2 Centrality Results

Now the network's centrality indices and other typical measures, like the network's density or the average degree, are discussed in detail.

Table 2. Websites with the highest in- and out-degree and in-closeness

Website	In-Degree	Website	Out-Degree	Website	In-Closeness
facebook.com	18.4	austria.info	47.2	facebook.com	39.5
tiscover.at	12.0	tirol.at	28.0	tiscover.at	35.1
youtube.com	11.7	genussland.at	25.9	youtube.com	34.9
oebb.at	10.5	mobile.austria.info	12.8	oebb.at	34.2
tiscover.com	8.2	austriatourism.com	10.8	tiscover.com	31.1
genussland.at	6.9	oberoesterreich.at	6.6	twitter.com	30.6
twitter.com	6.1	vorarlberg.travel	5.6	austria.info	30.5
addthis.com	5.9	blog.austria.info	4.7	addthis.com	30.4
oberoesterreich.at	5.6	ausflugstipps.at	4.4	oberoesterreich.at	30.2
tirol.at	5.5	linz.at	4.1	tirol.at	29.6

Table 2 shows the results for the in- and out-degree as well as the in-closeness. It can be seen that Facebook.com and Tiscover.at/Tiscover.com are by far the most prominent actors according to the in-degree centrality measure. It is remarkable that four out of the top ten are social media sites. This underlines their important role in providing a platform for tourism enterprises. Here these enterprises can reach many customers without high promotion costs. The out-degree column in Table 2 shows first of all destination management platforms on the top of the ranking. These semi-public platforms are crucial in supporting the mainly small and medium tourism enterprises (see Egger et al., 2007). These platforms are central points for attracting customers and presenting the tourism enterprises in a consistent and integrative way, while underlining the special character of a tourism region. Here the countrywide site for Austria Austria.info reaches by far the highest value. Looking at the in-closeness, clear similarities to the in-degrees can be seen. This seems to be evident because a high in-degree also makes a node closer to other nodes within the network. Looking at Table 3, also a high similarity between the values of the out-closeness and those of the out-degree can be noticed.

Table 3. Websites with the highest out-closeness and betweenness

Website	Out-Closeness	Website	Betweenness
austria.info	71.1	austria.info	24.1
tirol.at	60.3	facebook.com	14.9
genussland.at	53.3	genussland.at	13.7
mobile.austria.info	53.0	tirol.at	13.5
austriatourism.com	52.0	tiscover.at	9.1
scnem.com	51.2	oberoesterreich.at	5.5
region.austria.info	50.9	youtube.com	2.6
oberoesterreich.at	49.8	twitter.com	2.2
vorarlberg.travel	49.4	vorarlberg.travel	2.2
blog.austria.info	49.0	tiscover.com	2.1

When comparing the values of in- and out-closeness, it can be seen that the latter ones are almost twice as high as the first ones. So, it can be concluded that a site with a high closeness is more likely to link others than it is itself linked by them. Furthermore, the high closeness values (both in and out) reflect a very compact network, which may also reflect a high cooperation in the real world. This is underlined by the fact, that in the network many tourism enterprises and destination management platforms can be found. Typically, those actors cooperate strongly. Columns 3 and 4 in Table 3 contain the results for the betweenness centrality. It is remarkable that except for Austria.info all values are quite low. This means that almost all nodes are either connected directly or are not connected at all. So, here it is reflected that there is only a limited number of intermediaries in the network. Furthermore, the low betweenness values suggest a low density of the network.

4.3 Other Results

In Table 4 the characteristic network parameters, are shown. As mentioned above, the density of 0.003 seems to be quite low. However, this value is quite similar to that of the World Wide Web in general. Furthermore, the average degree is also similar to that of the World Wide Web (for more information on the values for the World Wide Web see Baggio et al., 2010). The average path length and the clustering coefficient of the network of the Austrian eTourism market are respectively lower (path length) and higher (clustering coefficient) with respect to what would be expected for a World Wide Web's subnetwork of the density recorded here. Examples (see Table 4) are the studies of Adamic and Adar (2003) that compare the technological networks (consisting of homepages) of students and staff of the Stanford University and the Massachusetts Institute of Technology (MIT), or the network analysis of the subdomain nd.edu (also cited in Piazzi, 2011).

Table 4. Comparison of small world properties (see text for references)

	Average path length	Clustering coefficient
Stanford	9.2	0.22
MIT	6.4	0.21
Nd.edu	n.a.	0.29
Austrian eTourism market	3.3	0.34

Both investigations show a lower average path length and a higher clustering coefficient than the whole World Wide Web. But these values for the average path length and the clustering coefficient of the Austrian eTourism market and also of the World Wide Web are - despite being something different - similar to those of small world networks (including real world social networks), whose main characteristic is the connectedness of the nodes through short paths (average path length is around the logarithm of the number of nodes) and the formation of clusters (Table 5). That is to say that the connectedness within the local neighborhood is higher than in the network as a whole.

Table 5. Important network parameters for the Austrian eTourism web network

Parameter	**Value**
Number of nodes	2395
Number of edges	16.893
Network density	0.003
Average degree	7.065
Average path length	3.328
Clustering coefficient	0.341
Key parameter of the power law distribution of the in-degrees (for in-degrees >12)	2.47
Key parameter of the power law distribution of the out-degrees (for out-degrees>29)	2.63

The distribution of the degrees of the Austrian eTourism network follows a power-law, at least for in-degrees > 12 and for out-degrees > 29. The corresponding exponents, which are reported in Table 4, were calculated following the procedure of Clauset, et al. (2009). For this, MATLAB (The Mathworks Inc., 2010) was used.

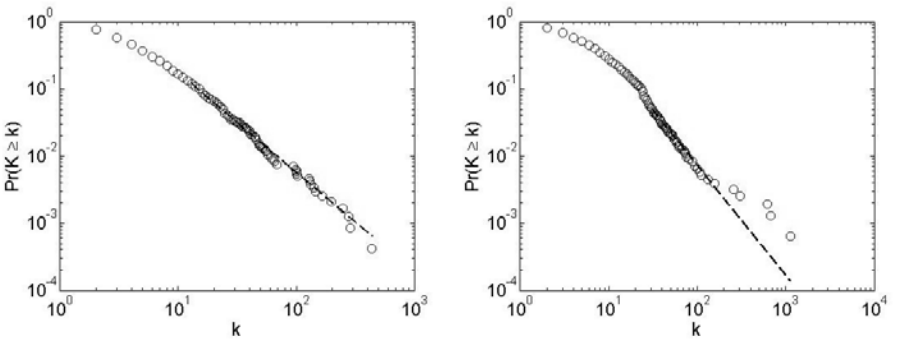


Fig. 2. Cumulative power law degree distribution of in- (left) and out-degrees (right)

Figure 2 shows the empirical cumulative distribution Pr(K>=k) of the probability of occurrence of degree k. The dotted lines indicate the approximated theoretical power law distribution with the key parameters stated in Table 4. Also in this case there is according to the literature a high conformance with the World Wide Web's values: the exponent of in-degree distribution is slightly higher than that of the World Wide

Web while the exponent of the out-degree distribution is more or less equal. Apart from the obvious consideration of the difference in size, a steeper in-degree distribution, which indicates a greater difference between the number of nodes with a high degree and those with a low degree, can be a symptom of some form of higher modularization of the network, a clear sign of a higher cooperative propensity of the Austrian tourism system (see Baggio et al., 2010).

5 Conclusion

The aim of this paper was to show how network analysis techniques can be applied in the study of tourism destinations' webspaces, and to discuss what insights they can provide in this context. After a short introduction to the methods and a brief review of the literature, the results of the analysis were presented and interpreted. The network under consideration exhibits well-known characteristic properties of complex networks. Only a few studies on network analysis and tourism exist so far, but the results provided in this paper confirm their findings. Furthermore, in the area of the Austrian eTourism, centrality measures have not been applied before. As it turned out, social media sites and destinations management platforms are the most central and most prominent actors within the Austrian tourism webspace. However, some big players of the "real world" tourism market are missing in the network. New intermediaries like HRS.com and Booking.com are not linked by tourism enterprises and vice versa. The reason for that might be that these intermediaries want customers to book via their own sites. The same might apply for travel agencies and tour operators.

Another finding of this paper is that in Austria there are many destination management systems that all have their own platforms – from countrywide ones down to systems for local tourism territories. This might be due to the fact that in Austria tourism is a very important economic sector generating 11.8% of the national GDP (www.wttc.org/eng/Tourism_Research/Economic_Research/Country_Reports/Austria / [Aug. 31, 2011]). Furthermore, since many tourism enterprises are small- or medium-sized (Egger et al., 2007), they have to cooperate and have to operate destination management systems jointly to remain competitive. Also the network's high clustering coefficient shows a relatively strong cooperation, especially between small enterprises. Moreover, it is interesting that three local destination management platforms (Genussland.at, Oberösterreich.at, Tirol.at) are higher in the in-degree ranking than the countrywide platform Austria.info. This might be due to a higher identification of tourism actors and guests with a smaller region.

The high in-degree centrality respectively in-closeness of the website of the Austrian rail company Oebb.at shows the importance of this means of transportation for tourism. This is underlined by the fact that also the website of the German rail company Bahn.de is in the top twenty in-degree ranking. Also two providers that can be associated with car traffic can be found in the top twenty. On the other side, websites of airports and airlines do not have a central position (Piazzi, 2011). But since 60% of all tourists in Austria are either from Austria or Germany (http://www.austriatourism.com/xxl/_site/int-de/_area/465219/_subArea/465248/ marktdaten.html [Sep. 6, 2011]) this is evident; due to the short distances, they rather travel by train or car than by plane.

Although Twitter.com is linked very often in the Austrian eTourism network, it is different from other social media sites. It can be seen when taking a closer look, that it is barely used. Except the twitter service of Tirol.at, which had at that time around 1200 followers, the twitter services of all other big players together only had around 200 followers. So, these findings do not affirm the high economic effects of Twitter usage proposed for example by Kaplan and Haenlein (2010).

Future research in this field should try to strengthen this paper's findings by analysing larger networks. Other future work might comprise a comparison with market analysis to find out the impact of the inter-linking, or inquiries of the tourism enterprises to learn why they link others.

References

Adamic, A. & Adar, E. (2003). Friends and neighbors on the web. *Social Networks 25*(3).

Baggio, R. (2006). Complex systems, information technologies, and tourism: a network point of view. *Information Technology & Tourism* 8(1): 15-29.

Baggio, R. (2007). The web graph of a tourism system. *Physica A* 379(2): 727-734.

Baggio, R., Scott, N. & Cooper, C. (2008). *Network science and socio-economic systems. A review focused on a tourism destination* (Dondena Working Paper No. 7): "Carlo F. Dondena" Centre for Research on Social Dynamics, Bocconi University. Retrieved September, 2011, from http://www.dondena.unibocconi.it/wp7.

Baggio, R., Scott, N. & Cooper, C. (2010). Network science: A Review Focused on Tourism. *Annals of Tourism Research* 37(3): 802-827.

Baggio, R., Scott, N. & Wang, Z. (2007). What network analysis of the WWW can tell us about the organization of tourism destinations. In *Proceedings of the CAUTHE 2007,* Sydney, Australia, 11-14 February.

Batagelj, V. & Mrvar, A. (1996). Pajek - Program for Large Network Analysis. Retrieved August, 2011, from http://pajek.imfm.si/

Borgatti, S.P., Everett, M.P. & Freeman, L.C. (2002). *Ucinet for Windows: Software for Social Network Analysis*. Harvard, MA: Analytic Technologies.

Clauset, A., Shalizi, C. R. & Newman, M. E. J (2009). Power law distributions in empirical data. *SIAM Review* 51(4): 661-703.

da Fontoura Costa, L. & Baggio, R. (2009). The web of connections between tourism companies: Structure and dynamics. *Physica A* 388(19): 4286-4296.

da Fontoura Costa, L., Oliveira, O. N., Travieso, G., Rodrigues, F. A., Villas Boas, P. R., Antiqueira, L., Viana, M. P. & Correa Rocha, L. E. (2011). Analyzing and modeling real-world phenomena with complex networks: a survey of applications. *Advances in Physics* 60(3): 329-412.

Egger, R., Hörl, J., Jellinek, B. & Joos, M. (2007). Virtual Tourism Content Network TANDEM – A Prototype for the Austrian Tourism Industry. In *Information and Communication Technologies in Tourism*. Vienna: Springer.

Hendler, J., Shadbolt, N., Hall, W., Berners-Lee, T. & Weitzner, D. (2008): Web Science: An interdisciplinary approach to understanding the World Wide Web. *CACM* 51/7, July 2008.

Kaplan, A. M. & Haenlein M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons* 53(1): 59-68.

Newman, M. E. J. (2010). *Networks - An introduction*. Oxford: Oxford University Press.

Österreich Werbung (2011). *Wirtschaftskammer Österreich*. Retrieved August 2011, from http://www.austriatourism.com/xxl/site/int-de/area/465222/_subArea/465274/subArea2/476161/index.html.

Österreich Werbung (2011). *Der österreichische Tourismus in Zahlen*. Retrieved September 2011, from http://www.austriatourism.com/xxl/_site/int-de/_area/465219/_subArea/465248/ marktdaten.html

Pennock, D., Flake, G., Lawrence, S., Glover, E. & Giles, C. (2002). Winners don't take all: Characterizing the competition for links on the web. *Proceedings of theNational Academy of the Sciences of the USA* 99: 5207–5211.

Pforr, C. (2005). Tourism Policy in the Making: An Australian Network Study. *Annals of Tourism Research* 33(1): 87-108.

PhoCusWright. (2011). *Global Online Travel Overview* (2nd ed.). Sherman, CT: PhoCusWright.

Piazzi, R. (2011). *Analysis of the network structure of the Austrian eTourism market and identification of the central points*. Master Thesis at the Vienna University of Technology, 2011.

Poon, A. (1993). *Tourism, Technology and Competitive Strategies*. Wallingford, UK: CAB International.

Schertler, W., Schmid, B., Tjoa, A M. & Werthner, H. (1994). *Information and Communication Technologies in Tourism.* Wien, New York: Springer.

Scott, N., Cooper, C. & Baggio R. (2008). Destination Networks. Four Australian Cases. *Annals of Tourism Research* 35(1): 169-188.

Sheldon, P. J. (1997). *Tourism Information Technology*. Oxon, UK: CAB International.

Shih, Hsin-Yu (2006). Network characteristics of drive tourism destinations: An application of network analysis in tourism. *Tourism Management* 27(5): 1029-1039.

The MathWorks Inc. (2010). *MATLAB Version R2010a*. Natick, Massachusetts.

Werthner H. & Klein, S. (1999). *Information Technology and Tourism – A Challenging Relationship*. Springer Vienna.

World Travel & Tourism Council (2011). *Economic report for Austria*. Retrieved from www.wttc.org/eng/Tourism_Research/Economic_Research/Country_Reports/Austria/.

Electronic Marketing for China Outbound Tourism: Implications from Chinese Online Travellers

Edmond H.C. Wu[a,1],
Rob Law[b], and
Chaozhi Zhang[a,2]

[a] School of Tourism Management
Sun Yat-sen University, P.R.China.
[1]wuhaocun@mail.sysu.edu.cn
[2]zhchzhi@mail.sysu.edu.cn

[b] School of Hotel and Tourism Management
Hong Kong Polytechnic University, Hong Kong.
hmroblaw@polyu.edu.hk

Abstract

Online travellers may exhibit distinct traveller behaviour in different travelling decisions. Their requirements are the primary considerations for E-marketing strategies and services provided to these online users. An investigation was recently conducted to understand the outbound tourism preferences of Chinese online travellers. By comparing the survey results with another similar study from Chinese general population, some differentiation aspects of online travellers are identified and then tailor-made E-marketing strategies are proposed for the fast growing population of Chinese outbound tourists through online channels.

Keywords: Electronic Marketing; Information Search; Outbound Tourism; Recommender Systems.

1 Introduction

With the largest potential outbound tourist population, China outbound tourism has drawn much research attention in the academic community (Cai, Li, & Knutson, 2007; Zhang, Chong, & Ap, 1999). Although China had a population of nearly 1.4 billion, the opportunity for Chinese to travel outside for leisure travel purpose is fairly recent (Sparks, 2009). Many observers dated the beginning of Chinese outbound tourism to 1990, when Mainland Chinese citizens were firstly allowed to take leisure trips to Singapore, Malaysia, and Thailand at their own expense (WTTC, 2003).

In the past ten years, the Chinese outbound tourism market has witnessed unprecedented growth due, in part, to relaxed government policies and increased individual disposable income. In 2010, a record 56 million mainland tourists travelled overseas, representing an increase of 8.3 million over 2009 (CNTA, 2011). In other words, China's outbound tourism has entered a new round of golden development.

At the same time, with the promotion and development of Internet technology, network information has infiltrated into all walks of modern life, the provision and acquisition of travel information, organization of travel activities, and even the experience of travel begin to be characterized by networking. Increased from 59.1 million in 2002, there were almost 457.3 million Internet users in 2010 in China (CNNIC, 2011). The real figure should be higher than the official statistics which

does not include users who share the usage of public computers. In China's marketplace, over 36 million Chinese Internet users search travel information and purchase tourism products online (CNNIC, 2011). As independent travelling and self-drive trips become more popular in China, the demand of Internet information searching and bookings online is expected to increase (Xiao, 2003). Undoubtedly, the importance of Chinese online travellers seeking for outbound tourism should draw our special attentions. Innovative wireless devices, such as iPhone and iPad, which are widely used among Chinese tourists, may provide excellent opportunities for the development of future mobile tourism products and services. Therefore, analyzing the new trends and preferences of Chinese Internet users is crucial for the success of relevant e-Tourism applications.

This paper reports the findings of a large scale Internet survey that was recently conducted in public in 2010 on a Chinese leading Internet portal website NetEase (www.163.com). The survey is to interview the outbound travelling behaviour of Chinese online users. The primary objective of this survey is to update the new trends of China's outbound tourism and try to identify the future services and applications that are suitable for this largest and fast growing outbound tourist population.

2 Literature Review

Internationally, outbound tourism, which is also called travel to foreign countries, refers to citizens' travel activities to another country or area (Jafar, 2000). Nowadays, the Mainland Chinese outbound tourism market is growing rapidly. Wang (2007) figured out that the expansion of the Chinese outbound tourism is an inevitable result of the economic development. The appreciation of the Chinese RMB would further enhance the purchasing power of Chinese visitors.

In a previous study, with regard to consumer decision, 31% of tourists made their outbound travelling decision due to the attractiveness of the destination, and 25% were curious about the destination (Li, 2006). Inevitably, shopping became the main segment of the outbound tourism for most of the Chinese tourists, accounting for nearly 50% of the overall expenditure. With the maturation of the outbound consumption psychology, the expansion of the outbound scale and the generalization of the outbound tourism, the high expenditure on shopping may get changed (Li, 2006).

Owing to the tourism market development, well organized outbound tourism, especially travelling abroad through travel agents, will still be the main stream of this market share. However, from the concrete organizing methods, more self-guided tours will be involved in the short-distanced outbound tourism. With the introduction of the Individual Visit Schemes (IVS) to many destinations and the free trade policy of the ASEAN, more Mainland Chinese tourists will choose Hong Kong and Southeast Asian countries as their first self-guided outbound travelling destinations. The trend of self-guided tour will appear in these destination countries and areas for Mainland Chinese tourists. After that, and the destinations will be strengthened in other global destinations as the maturation of the Chinese outbound tourism market (Li, 2006).

Sparks (2009) used the theory of planned behaviour (TPB) to investigate potential Chinese outbound tourists' values in terms of destination attributes. In the survey, five destination attributes were rated as most important by the group of potential tourists and included natural beauty and icons of a destination, quality infrastructure, autonomy, inspirational motives, and social self-enhancement.

With the development of e-commerce, online searching of travel information and hotel booking become popular in China. Li and Buhalis (2006) found that 90.2% of surveyed online users are in the age group of 31 to 40, and 95.3% of online users who live in the Pearl River Delta of China will make reservations of travel trips online. Moreover, age, area of residence, type of travel websites most visited, and length of time using the Internet, self efficacy, domain-specific innovativeness, and perception of the Internet are influential factors of Chinese online browsers for becoming online buyers.

Although previous empirical studies provided some conclusions about China outbound tourism, the Chinese outbound tourist population has experienced significant changes in the past few years. Examples of these changes include the fast development of middle and wealthy classes and the relaxation of visa restrictions for Chinese to travel abroad. In order to update some new findings, this study focuses on investigating Chinese online travellers' preference and behaviour with the major aspects of: trip intention and motivation, choices of outbound destinations, arrangement of travel, travel season preferences, and consuming behaviour.

3 Research Findings

In the study, Internet users from different areas of Mainland China completed an online questionnaire on the China NetEase website during a two weeks period from September 6th to September 20th, 2010, using a convenience sampling procedure. The survey is conducted in Chinese language and all the participants are Chinese residents.

In our research, we define the target Chinese Internet users as Chinese residents who are at least 16 years old and who have used the Internet in the past six months. Internet travellers refer to the Internet users who travel domestic or abroad in the last three years. As stated in the start of the online questionnaire, Internet users who never travel in the last three years will be asked to stop the survey and thus only Internet travellers are included in the effective sample. The final effective sample size is 2,797.

The demographic profile of the respondents is listed in Table 1. From the table, one can find that most respondents were 21-40 years old and had a relatively high level of education background. The personal incomes varied from person to person but generally below RMB 10,000 per month (USD1=RMB6.8). Nearly half of the respondents had previous outbound travelling experience in the past three years.

Table 1. Demographic characteristics of the respondents (N=2,797)

Variable	%	Variable	%
Gender		**Monthly personal income (RMB)**	
Male	54.5	2,000 or below	27
Female	45.5	2,001-5,000	51.6
Age (years)		5,001-10,000	16.0
<21	4.8	>10,000	5.4
21-30	58.6	*Note: around 6.4 RMB=1 USD in 2011	
31-40	25.6		
41-50	8.2	**Number of times for travelling abroad in the past 3 years**	
51-60	2.3	Never	53.2
>60	0.5	Once	25.1
Educational level		2 times	7.5
Secondary school or below	2.4	3 times	5.0
High school or polytechnic school	12.2	>3 times	9.2
Undergraduate or colleges	75.4		
Post graduate or above	10.0		

Another recent and relevant study about China outbound tourism was conducted by Ivy Alliance & Pacific Asia Travel Association (IA & PATA, 2010). This study conducted a survey between August and November 2009, by the China Association of Travel Services and Ivy Alliance to collect first-hand data by sending questionnaires to thousands of outbound tourists and around 100 travel agencies. Finally, they obtained 2,800 returned questionnaires from 4,200 survey forms given out. Among the respondents, 46.2% are male and the other 53.8% are female. Since IA & PATA's target is general Chinese population, we compare our research findings with this report to find the differentiation aspects of Chinese online travellers.

3.1 Analyzing the Characteristics of Online Travellers

Occupation. Diversity of respondents' occupations is observed in the samples. Based on the findings, professionals, civil servants/government employees, corporation staff, students, and freelancers comprised a large portion to the Internet travellers. In contrast, factory workers, farmers and retirees, which also account a large population in China, are not a dominant group in terms of Chinese Internet travellers. As to the levels of outbound travelling experience, corporation staff and civil servants/ government employees were the main groups in outbound travelling, since outbound travelling generally requires relatively higher spending ability and more leisure holidays. Moreover, these groups of Chinese have more chances of outbound travelling for business or visiting purposes. Their travelling expenses may be sponsored by their employers.

Education. The Internet user groups are characterized with high education, and middle or high incomes. The percentage of respondents with college/university or higher education is 85.4%, which is much higher than the current average education level of the China population.

Income. More than 50% Internet users have income between RMB 2000~5000 and over 20% earn more than RMB 5000 per month. It shows that Internet users who travel for leisure have higher education, and have enough money to participate in travel and leisure activities. Additionally, outbound travelling Internet user groups appear to have high education and middle level of incomes. Outbound travelling requires more of the quality of travellers, and generally charges higher. Thus Internet travellers who travel abroad have higher education and slightly higher incomes than the general Chinese population.

Region. The regional distribution of the interviewed Internet travellers is not even and it is relevant with the relative regional economic development in China. Divided by provincial regions, the Internet users who travel abroad and the Internet users who travel domestic most concentrated in provinces with strong economic development, such as Guangdong which accounts for around 25% of Internet travellers. From east, central, west and north-east mega regions of China, the percentage of Internet travellers decreases, which are also consistent with the fact that China outbound tourists are mostly from well developed provinces or cities. The source regions of China Internet travellers are classified below:

1. **East Mega Region**: Beijing, Tianjin, Hebei, Shanghai, Jiangsu, Zhejiang, Fujian, Shandong, Guangdong, and Hainan.
2. **Central Mega Region**: Shanxi, Anhui, Jiangxi, Henan, Hubei, and Hunan.
3. **West Mega Region**: Inner Mongolia, Guangxi, Chongqing, Sichuan, Guizhou, Yunnan, Tibet, Shanxi, Gansu, Qinghai, Ningxia, and Xinjiang
4. **North-east Mega Region**: Liaoning, Jilin, and Heilongjiang

Based on the above observation and analysis, we can assume that the collected samples are adequate and well represent the current and potential population of Chinese outbound online travellers. In the following sections, we further analyze the key characteristics and differentiations of the target Internet travellers.

3.2 Outbound Intensions and Motivations of Online Travellers

Respondents generally expressed their preferences for outbound travelling for various purposes, such as experience exotic culture (70.8%), recreation and vacation (65.3%), and explore the world and enrich experiences (50.5%). As for the travel budget, 55% of the respondents thought that it was very worthy and 32% thought it was worthy of spending a lot on outbound travelling. Meanwhile, only 1.3% of them held the opposite view. Therefore, respondents showed a positive attitude towards outbound travelling and they also had a high motivation to take an outbound travelling in the future.

As indicated in Table 1, about 53% Chinese online travellers did not take an outbound trip in the past three years. Comparing with their high motivation and willingness to

travel abroad, the potential outbound market of China is very huge. Thus, promoting outbound travelling through e-Tourism may be a very effective way to explore this tourist source market in the future. Based on the population of Internet users in China in 2010, one can estimate that at least 200 millions Chinese Internet users who did not travel abroad in the past three years may intent to travel abroad in the future. Since the China government is encouraging their residents to consume more, including tourism spending, providing outbound tourism guideline and advertisement to these tourists with less outbound travel experience through the Internet would be a promising e-Business opportunity.

It seems that the internal conditions of outbound tourism in China are becoming mature. With the strong interests of travelling outside, once with more leisure time and stable income growth, the motivation will be greatly stimulated. Also, with consistent improvement of income levels and currency appreciation of the RMB, the outbound tourism market of Chinese Internet users is expected to keep expanding.

3.3 Outbound Travelling Destinations of Online Travellers

Based on previous outbound travelling studies (IA & PATA, 2010), the choices of destinations had a negative relevance with the distance between the tourism-generating region and a destination (Table 2).

Table 2. "Have visited" and "Plan to visit" destinations of respondents

Outbound Destinations	Destinations have been in past 3 years (%)	Destinations plan to visit (%)
Hong Kong, Macau, Taiwan	47.5	9.4
Surrounding countries of China	35.7	15.3
Other Asian countries	17.2	9.5
European countries	14.5	35.0
Australia, New Zealand	6.7	15.2
United States	4.4	4.8
Canada	2.7	2.7
African countries	2.9	3.5
South American countries	1.1	2.3
Others	4.2	2.2

The further a destination is, the fewer people have ever been there, as travel distance has something to do with time and money. Therefore, Hong Kong, Macau, Taiwan (47.5%) and other surrounding areas of mainland China (35.7%) are still the most popular outbound travelling destinations.

Also, 14.5% of Chinese online travellers chose European countries as their destinations in the past three years. This percentage is significantly higher than other similar studies on China outbound tourism. According to IA & PATA's estimation in 2010, less than 5% of China outbound tourists chose Europe as their destinations. The higher percentage preferences of European tourism in our study suggested that Chinese online travellers showed much stronger interests to travel in Europe.

Also, when asked about which countries they wish to go for outbound tourism next time, about 35% of the respondents chose European countries, which was the highest ratio among all alternatives. Australia and New Zealand are also desirable destinations for China′s Internet users. These results are interesting and may have very constructive implications for European tourism officials and industries. The Chinese outbound tourism market to Europe from online travellers is expected to have large growth in the next few years. In other words, European tourism destinations are likely to increase their market share in China outbound tourist market, especially from online channels.

On the contrary, Internet users′ yearning for those neighbouring countries or regions like Hong Kong, Taiwan and surrounding Asian countries declined compared with their previous preferences. Even though it is just a forecast, it implies that Chinese online travellers curiously tend to explore more destinations they never visited before and those nearby popular destinations thus need to improve their tourism facilities and promotion to keep their competitiveness in the future.

This is also an interesting finding to our understanding of China′s outbound tourism. Longer distance of a destination does not necessarily imply less attractiveness, especially for the young and energetic generation of Chinese Internet travellers. They desire many aspects of European tourism, including history, architecture, festivals and cultures. Among the respondents used to travel to Europe before, a high percentage shows a revisit intension, suggesting that European tourism will be a popular area for Chinese outbound tourism in the next few years. Many e-Tourism services and applications that can meet these customers′ needs, such as route planning, hotel and transportation reservations, dining recommendation, and location based services (LBS) will be popular considering the huge potential customer base.

3.4 Outbound Trip Arrangement of Online Travellers

Currently, package tour is the major way of outbound travelling for Chinese travellers. Through our survey, 44.8% of the respondents choose package travel because outbound travelling generally has a higher request for safety and language communication. Thus, travel agencies still have the advantages in organizing outbound tours. However, with more abroad travel experience, Chinese Internet users show a strong desire to have a freely self-guided outbound travelling. When asked about what form of travel to choose in next tour, 57% choose self-guided travel, thus we can see that even in outbound travelling, self-guided tour has been a new trend in the future.

According to IA & PATA′s estimate in 2010, only around 1/3 of Chinese travellers prefer self-guided tour (independent or semi-independent travel). Our studies suggested that Chinese online travellers have a significant distinction with the general Chinese residents as online travellers desire freedom in arrangement for their outbound travelling. With a high education level, Chinese online travellers may have better language skills and thus they are more confident in planning and making reservations online from tourism related websites outside China. Therefore, demand for online travel information sharing platforms, and demand for self-guided travel guides will become more popular.

In the survey, about 55.7% of the respondents chose the Internet as their primary source of outbound travelling information, which they consider as the most convenient and reliable way to get updated information. However, online travel information search still has much room to improve. Some tourism destination websites are already in Chinese language. However, the majority is not. In order to attract more Chinese online travellers, it is suggested that more tourism related e-Commerce websites provide more convenience for Chinese travellers to online purchase tourism products and services, such as Chinese language web pages and more online payment methods.

Another finding is that Internet users showed a preference for companion during outbound travelling. Most of them (47%) chose lovers or family members as companion rather than friends or colleagues. Therefore, it will be necessary to take privacy into consideration when choosing a form of outbound travelling. More concerns of privacy and personality may be a characteristic of Chinese online travellers.

Although package travel is still popular among Chinese outbound tourists, it seems that self-guided travel is more likely to meet the demands of China′s Internet users. Therefore, more personalized online tourism services and business applications should be developed in the future to meet such demands.

3.5 Seasonality of Online Travellers

Another interesting discovery through our survey is that 55.7% of Chinese Internet travellers prefer to travel abroad during non-public holiday periods in order to avoid peak seasons. They planned to avoid travelling in crowded long public holidays, such as the first week of May of the labour holidays, and the first week of October of the National Day holidays. They choose other non-peak seasons to travel as these peak seasons are popular tourism periods among Chinese outbound package tours.

In contrast, general Chinese outbound tourists still prefer long public holidays as their primary outbound travelling time (IA & PATA, 2010). This differentiation suggests that online travellers pay more attention to the quality and comfort during travelling and try to make a schedule avoiding crowded population of outbound tourists.

Spring and autumn are Internet travellers′ most favourite seasons for outbound travel, and summer is the favoured season for all types of tourists. Actually there is not much gap between best months for outbound travelling except January to March, which is usually the spring festival period for Chinese but still more and more Chinese tourists chose to travel during this period. Since the Chinese outbound population is so large, every single month still has many Chinese to travel abroad across the world. Some people will choose to travel in off-peak season to enjoy a big discount in travelling expense. Therefore, further market segmentation can be done on the Chinese online travellers in the future.

3.6 Travel Consumption of Online Travellers

Our survey also validated that Chinese online travellers spend much on outbound travelling comparing to their income level. More than 50% of the respondents spend more than RMB5,000 and 30% of Internet users spend over RMB10,000 in outbound

travelling. For those package travellers, over 66% of them had extra expenditures which were above half of the total tour fare. Sightseeing and shopping were two main factors of their personal travel expenditures. Traffic and accommodation held the largest portion in expenditures of self-guided travellers as usual practice. Shopping and sightseeing also cost a lot for online travellers. While in package travel, travel agencies can arrange accommodation and traffic for Chinese outbound travellers with a deep discount. That explains why the tour fare of travel agency is relatively competitive.

With the development of e-commerce and the boom of Internet users, the ratios of Internet users′ online consumption in outbound travelling will become high. They can also easily find discounted hotel rooms and air tickets by online searching. Over 50% of the respondents usually book air/railway tickets and hotel rooms online, which reflect the fact that e-Tourism applications are mature and widely accepted among Chinese tourists.

In other words, online travellers will continue rely on the Internet for searching and making reservations. For most online travellers, convenience and favourable price are the key advantages of online booking and payment, but the adaptation of online payment methods is till a barrier. For example, many Chinese make online payments through popular UnionPay cards in China or third party payment intermediary (e.g., alipay.com). However, many international e-Tourism products do not support these Chinese tourists′ usual payment methods at this time.

Our survey results revealed that 85% of Chinese online travellers used to make online purchase of tourism products or services. With respect to the portion of online payment to the total travel expense, 47.6% of respondents made online payment less than 1/3 their total travel expenditures. While 26.8% of respondents made online payment between 1/3 and 2/3, and 10.5% of respondents will make the purchase of e-Tourism products or services more than 2/3 of their total travel expenditures. These figures obviously suggested that technology acceptance among Chinese online travellers is high. Also, it reveals that online searching and payment of e-Tourism products and services play a key role in tourism management. Many research and applications should be done in the future to fulfill the diversity, security, speed and comfort for all these online e-Tourism transactions.

4 Implications and Conclusions

In this empirical study, several thousand questionnaires are collected to analyze the unique outbound preferences and patterns of Chinese online travellers. Through some official statistics and combined with our survey estimates, China has over 400 million Internet users who show their interest to outbound tourism. Among these people, around 200 million Internet travellers without much travel abroad experience attempt to plan an outbound trip in the future.

Chinese outbound tourists have high spending willingness and abilities and can improve the trade balance with other countries. For this fast growing source market with a huge number of potential tourists, innovative research and applications are invited from international scope, especially for e-Tourism. Chinese online travellers

have some distinguishing features and preferences comparing with the general Chinese travellers.

In summary, several key research findings, drawing on the results of our survey, are:

- Online travellers are younger population with higher education level and income.
- Online travellers prefer independent trips rather than package tour.
- Online travellers anticipate more diversity, personality and privacy during their travel period.
- Online travellers are more favourable in European destinations rather than Asian destinations as their future first choice of travel places.
- Online travellers are more rationale and prefer travelling in Non-public holiday period to avoid peak season.
- Online travellers trust more online payment and word-of-mouth reputation of e-Tourism products and services.

Some implications of electronic marketing for Chinese outbound tourism are:

- With relatively richer travel experiences, online travellers for outbound travelling consider diversity of destinations and favour long-distance or privacy attractions. Online search of such tourism products can fulfil their curiosity.
- More Internet tourism contents are welcome to write in Chinese to attract more Chinese outbound audience and more diversity online payment methods are expected.
- Internet marketing for off-season outbound tourism can attract more Chinese online travellers to travel without many people meanwhile enjoy less travel budget.
- Electronic marketing is an excellent tool for destination marketing, such as European countries can adopt this channel to attract more Chinese tourists for visit. Tourism searching engines, tourism recommendation and intelligent systems, travel route maps and location-based services for tourism are promising applications that can serve millions of Chinese online travellers in the future.

Our empirical study indicated that Chinese online travellers also have high technology acceptance rate but the e-Tourism applications provided to them are still limited. Therefore, this article attempts to attract the attentions of tourism and IT research as well as industry participants to bring emerging and innovative e-Tourism applications and services for this prosperous area.

5 Limitations and Future Work

One limitation in this study is that online survey is the only way of primary data collection, mainly due to time and cost constraints. For future survey design, online survey as well as face-to-face interview by randomly selected interviewees through stratified sampling may provide improved samples and generalized results.

There is much research work to be done for the future e-Touism market of China. Considering the huge population of China tourists, we can further do segmentation of outbound tourists and analyze their unique characteristics and specific behavior, such as compare online travellers with versus without travel experiences, domestic versus outbound tourism. Then IT and hospitality providers can provide each segment proper products and services differentiation by knowing what kind of travel websites and travel information they are looking for and what e-Tourism applications they will use, and also other useful tourist web usage patterns that can assist their decision making.

References

Cai, L., Li, M. & Knutson, B. (2007). Research on China outbound market: A meta-review. *Journal of Hospitality and Leisure Marketing*, 16(1/2): 5-20.

China Internet Network Information Center (CNNIC). (2011). *Statistical Report on Internet Development in China*.

China National Tourism Administration (CNTA). (2011). *Annual Tourism Statistics Report*.

Hong Kong Tourism Board (HKTB). (2009). *A Statistical Review of Tourism*.

Hong Kong Tourism Board (HKTB). (2009). *Tourism Expenditure Associated to Inbound Tourism.*

Ivy Alliance & Pacific Asia Travel Association. (IA & PATA) (2010). *China Outbound Tourism Yearbook 2010*.

Jafar, B. (2000). *Encyclpedia of Tourism*. London: Routledge.

Law, R., Rong, J., Vu, H., Li, G. & Lee, A. (2010). Identifying changes and trends in Hong Kong outbound tourism. *Tourism Management* 32: 1-9.

Li, L. & Buhalis, D. (2006). E-Commerce in China: The case of travel. *International Journal of Information Management* 26: 153-166.

Li, X. (2006). The Changes and Trends of The Chinese Outbound travelling. *Human Geography* 3(89): 83-87.

Li, X., Harrill, R., Uysal, M., Burnett, T. & Zhan, X. (2010). Estimating the size of the Chinese outbound travelling market: A demand-side approach. *Tourism Management* 31: 250-259.

Li, X., Lai, C., Harrill, R., Kline, S. & Wang, L. (2011). When east meets west: An exploratory study on Chinese outbound tourists' travel expectations. *Tourism Management* 32: 741-749.

Sparks, B. & Pan, G. (2009). Chinese Outbound tourists: Understanding their attitudes, constraints and use of information sources. *Tourism Management* 30: 483–494.

Wang, X. (2007). The Increase, Reasons and Trend in Crowd Scale of Mainland Chinese Tourists' Outbound Tourism. *Population Journal* 6: 12-15.

World Travel & Tourism Council (WTTC). (2003). *The Impact of Travel & Tourism on Jobs and the Economy*.

Xiao, H. (2003). Leisure in China. In A. Lew, et al. (Eds.). *Tourism in China* (pp.263-275). New York: Haworth.

Zhang, H., Chong, K. & Ap, J. (1999). An analysis of tourism policy development in modern China. *Tourism Management* 20: 471-485.

Acknowledgements

The authors would like to thank NetEase Corporation and relevant people to provide helps and assistance in organizing the online survey. The authors also would like to acknowledge funding support from the Hong Kong Polytechnic University for the conference attendance.

A study of Hospitality and Travel-Related Deals on Hong Kong Group-buying Websites

Ada Lo,
Joey Wu, and
Rob Law

School of Hotel & Tourism Management
The Hong Kong Polytechnic University, Hong Kong
{hmada, hmjoey, hmroblaw}@polyu.edu.hk

Abstract

This study investigates the impact of the features of hospitality and travel-related group-buying deals and the group-buying websites on the sales of the deals. Empirical results reveal that the number of Facebook fans of the group-buying site is the most important factor in contributing to the sales. The special offer price for the deal has an inverse relationship with the sales. The numbers of aggregation group-buying sites featuring the deal and the length of coupon redemption period also have a significantly positive relationship with sales. If the deal belongs to the category of Chinese quick service restaurant, hotel/accommodation, travel package, and spa/beauty, this will have a negative impact to the sales of the deal.

Keywords: group-buying, group-buying deals, hospitality and travel, online purchase, website

1 Introduction

With the increase in popularity of online shopping, intermediaries continue to evolve to meet the demand of market. Hunting for the best deals is now a global phenomenon. Group-buying websites offer dynamic discounts for a wide range of products, from manufactured goods to hospitality services, to attract budget customers. The global market leader Groupon is regarded as "the world's fastest-growing company ever", with more than 4,000 staff, 51 million subscribers in 565 cities worldwide, and generated USD$760 million in revenue by the end of 2010" (The Economist, 2011). Now many websites follow the business formula of Groupon and have been launched in a number of major cities, thereby emphasizing their wide range of "deals of the day" offered in only a short period of time with a fixed dynamic price when enough buyers participate. As at June 2, 2011, Hong Kong has 45 group-buying websites with Yahoo Deals as the lasted entrant to the group-buying website arena (Leung, 2011). The main concept behind the business of group-buying websites is nothing new, but a combination of "online platform", "bulk purchase", and "team buying". Group-buying websites solicit merchants to sell products via their websites by offering a substantial discount off the regular price to customers. The deals (i.e. coupons for the discounted products) may be made available to customers for one or more days. The deal will only be "on" when a minimum threshold number of customers (tipping point) is reached (Wei, Straub, & Poddar, 2011). Thus, the deal is not offered to one individual who buys in bulk quantity, but is only available to numerous individual online shoppers who buy the same deal from the group-buying websites. Subscribers of the group-buying websites will receive information of the daily deals in their respective city via Email. Customers can simply purchase the deal coupons online and their credit cards will only be charged when the minimum number of customers is reached. The customer can then print an electronic coupon for the deal for redemption at the merchant (Groupon, 2011).

The group-buying websites feature deals from different categories of businesses including restaurants, health, beauty, events, recreation, services, and products. Information of each featured deal is made available on the group-buying sites to facilitate customers to make their decision to buy or not to buy. For example, the discount offered for the deal, the original and the special offer price, the tipping point in order for the deal to be on, start and expiry date for the deal, description of the product or services, terms and condition of the purchase may influence customers' decision. Moereover, characteristics of the group-buying websites may also play a role. Some of these group-buying sites have their own Facebook pages which also make it easier for their subscribers who are Facebook users to access and purchase the deals. Geographic coverage can also have different impacts on the quantity of the products or services sold. Since there are many group-buying websites offering numerous deals, finding the right deals may be relatively time consuming for customers. Thus, deal aggregator sites help to solve the problem by collecting all the deals from different sites and displaying them under different deal categories on one central website. Although these aggregator sites can deliver a huge number of deals from different sites to customers, they may not necessarily be what customers want (Stafford, 2011). Although it is logical to think that the deals with steeper discounts will attract more customers to purchase, and hence making the deals more successful, other features of the deal and the group-buying websites may also play a role in influencing the result. The purpose of this study is to investigate the impact of the features of hospitality and travel-related group-buying deals and group-buying websites on the sales of the deals (i.e. number of deal coupons sold). Results of this study will help hospitality and travel-related product vendors decide whether group-buying websites are an appropriate intermediary for their businesses, and if so, which type of group-buying websites is best for selling their products or services. This study will also help provide useful guidance on how to formulate the best selling features for hospitality and travel related deals on group-buying websites.

2 Related Studies

Before making a buying decision, customers may go through the process of searching for information and evaluating the alternatives available before they make the purchase (Kotler, Bowen, & Makens, 2006). A number of deal features can influence the sales on group-buying websites. Chiang, Lin and Chin (2010) found that online products sold better with appropriate promotion time limit being set. The difference (discount) between a reference price (face value) and a sale price (offer value) listed on advertisement may lead to consumers' perception of a greater value concerning the purchase of the product (Alford & Biswas, 2000; Compeau & Grewal, 1998). Kauffman and Wang (2001a) stated that demand increases as price dropped. Sufficient volumes of orders are the basis for a lower transaction price in group-buying activities (Kauffman, Lai, & Ho, 2010). As such, the minimum number of sales for the deals to be active can be another critical factor to entice a purchase. Brown, Pope and Voges (2003) studied that product type had an effect on online purchase intention. Krishna and Zhang (1999) found that longer coupon redemption duration not only generates more coupons sales but also attracts more customers. Studies also showed that the face value of an offer or its mode of presentation is

prominent for incentives with short redemption time frames but not for incentives with long ones (Roehm & Roehm, 2011).

Coehlho and Easingwood (2004) suggested that the exposure of products via multiple channels could lead to sales increase. Accordingly, some merchants place their deals in more than one group-buying website, while some group-buying websites link themselves to a few aggregation sites in order to extend market coverage. Companies can make use of the connections of customers from social networks to increase sales (Hill, Provost & Volinsky 2006). Thus, the number of social network subscribers to a company, and the availability of rewards and incentives for referring customers can also have impacts on its number of products or services sold.

Due to the perishable nature of hospitality and travel products/services, offering deals via group-buying websites certainly helps reduce the impact by generating sales to fill the business. As in the case of hotels, Law, Chan and Goh (2007) confirmed that most hotels used multiple distribution channels to increase market exposure. Similarly, Coehlo and Easingwood (2004) agreed that new channels provided access to market segments that were not previously reached by a company. Yoon, Yoon and Yang (2006) found that air travel markets would shift from traditional channels to Internet-based channels.

Prior studies on group-buying mainly focused on comparisons of pricing mechanisms of group-buying websites (Anand & Aron, 2003; Chen, Chen, & Song, 2006, Chen & Roma, 2011). Some prior studies concentrated on group-buyers including their benefits gained from group-buying (Chen, Chen, Kauffman, & Song, 2009) and their trust issues from group-buying deals (Kauffman, Lai, & Lin, 2010). Other studies looked into the characteristics of various types of group buying businesses (Kauffman & Wang, 2001b; Li, Zhao, Wang, & Wang, 2009; Schotanus & Telgen, 2007). No prior study, however, has been conducted to examine the impact of the features of individual deals and the group-buying website on the sales of deal coupons for hospitality and travel-related products or services.

3 The Study

Data for this study were collected from the top 25 group-buying websites from Hong Kong between 14 July and 11 August 2011. We only selected deals for hospitality (i.e. Chinese quick service restaurants, Chinese full service restaurants, non-Chinese quick service restaurants, non-Chinese full service restaurants, and spas), leisure (i.e. attractions), and travel related products or services (i.e. accommodations, transportation, and travel packages). Daily new deals appeared on these websites between 14 July and 28 July 2011 and were tracked on a daily basis for the duration until the end of the promotion period. The last promotion ended on 11 August. The number of coupons sold on each day for each active deal was recorded at 3:00 pm daily. A print screen capturing the details of the deal was made on the day when the deal was closed. A total of 246 deals were collected.

Based on the literature (see previous section), a number of deal features can influence the sales on group-buying websites. They include: type of the deal, promotion start date, promotion end date, duration of promotion, coupon redemption start date, coupon redemption end date, length of deal redemption period, discounted price, discount percentage, minimum number of sales for the deal to be active, number of

group buying sites featuring the deal, and number of aggregation sites featuring the deals. These features, together with the number of Facebook fans for the group-buying sites (which is a possible proxy for the number members of the sites), availability of rewards and incentives for referral, geographic coverage of the sites, and the total sales for each deal were also tracked. Information for these features on the print screen of each deal is coded and entered into SPSS.

4 Results

Two of the deals were removed from the data set as the deals disappeared after they were launched for a few days. Multiple regression analysis was employed to analyse the impact of different features of deals and the group-buying sites have on the number of deals sold. The number of coupons sold for each deal serves as the dependent variable. The independent variables include promotion duration, length of coupon redemption period, special offer price for the deal, discount percentage, number of group buying sites featuring the deal, number of aggregation sites featuring the deal, number of Facebook fans for the site, minimum number of customers to purchase the deal, rewards and incentives for referral, geographic coverage of the group-buying sites, and product/service category. Dummy variables were created for the nominal variables which include: rewards and incentives for referral, geographic coverage of the group-buying sites, and product/service category.

The data were assessed to ensure that assumptions of normality, linearity, homoscedasticity, and independence of residuals were met for the multiple regression analysis. Log transformation was applied to three independent variables (special offer price of the deal, number of group-buying sites featuring the deal, and minimum number of customers to purchase the deal) and the dependent variable (number of coupons sold). Outliers were checked and removed and finally 238 samples were used for the analysis. There is no correlation between the successive residuals as the Durban-Watson statistic of the final regression model is 1.820 which is close to the indicator of 2.0 (Hair, Black, Babin, & Anderson, 2010).

4.1 Characteristics of the Data

Table 1 shows the list of the group-buying websites where the deals were captured for the current study. These were the top 25 sites based on the most number of Facebook fans of each site. The 238 deals generated a total revenue of HK$14.9 million (US$1=HK$7.8) during the 29-day period. Seventeen percent of the deals were from Groupon, followed by 11.3% from Travelfish, and 10.9% of Beecrazy. Groupon has the highest revenue generated per deal (number of deals sold x the special offer price) (HK$170,605), followed by Beecrazy (HK$111,715), and Funshare (HK$106,835). Groupon generated close to half of the total revenue of all the deals (HK$6.8 million), while Beecrazy generated close to 20% (HK$2.9 million), and Funshare generated close to 9% (HK$1.3 million).

Among all deals, 171 (71.8%) covered only the Hong Kong market, 41 deals (17.9%) have a global coverage by being one of the geographic regions covered by the group-buying site, and 26 deals (10.9%) reached Hong Kong and Guangdong province in mainland China. Table 2 shows that majority of the deals were offers from non-Chinese full service restaurants (31.5%), followed by those offering spa and beauty services (25.2%), and travel package (13.4%). Only 1.7% of the deals were for

transportation arrangements. Deals from non-Chinese full service restaurants generated the most revenue (HK$6 million), followed by spa and beauty services (HK$2.5 million), travel packages (HK$2.3 million), and accommodation/hotel (HK$1.2 million). Transportation generated the least revenue (HK$49,536).

A total of 168 deals have information for both the start and end dates of the deals. As shown in Table 3, most of the deals started on Monday (18%) and Tuesday (18%), and weekend of Saturday (10.7%), while Sunday (4.2%) had the lowest number of new deals. Over 60% of the deals ended in the first half of the week (i.e. from Monday to Wednesday) and only 9% of the deals ended during weekend. Deals that started on Thursday sold the most number of coupons with an average of 921 coupons sold per deal. Deals starting on Saturdays generated the highest amount of revenue per deal (HK$82,372). Deals that started on Sundays sold the least number of coupons (i.e. 54 coupons) and the lowest revenue per deal (HK$24,916). Deals started on Monday generated the highest total revenue (HK$2.1 million), while those started on Sunday generated the least total revenue (HK$174,412). For deals that ended on Tuesday, they sold an average of 1,000 coupons per deal which is the highest among all the days of the week.

Table 1. No. of deals and revenue generated from the top 25 group-buying sites

Group-buying sites	No. of deals	% of deals	Avg rev per deal (HK$)	Ttl rev (HK$)	% of ttl rev
Groupon	40	16.8	170,605	6,824,215	45.77
Travelfish	27	11.3	21,330	575,910	3.86
Beecrazy	26	10.9	111,715	2,904,594	19.48
Couppie	22	9.2	9,322	205,076	1.38
Joycoupon	19	8.0	33,807	642,337	4.31
Babybamboo	19	8.0	12,755	242,337	1.63
Groupbuyer	13	5.5	69,808	907,504	6.09
Funshare	12	5.0	106,835	1,282,017	8.60
Hklashou	10	4.2	9,396	93,963	0.63
Gobuyer	9	3.8	28,105	252,946	1.70
Others	41	17.3	23,903	980,004	6.58
Total	**238**	**100**	**62,651**	**14,910,903**	**100**

Table 2. No. of deals and amount of revenues generated from each deal category

Deal category	No. of deal offers	%	Average revenue generated per deal (in HK$)	Total revenue (in HK$)
Non-Chinese quick service restaurant	17	7.1	55,394	941,698
Chinese quick service restaurant	6	2.5	9,254	55,524
Non-Chinese full service restaurant	75	31.5	80,791	6,059,338
Chinese full service restaurant	26	10.9	37,915	985,794
Accommodation/hotel	9	3.8	136,235	1,226,116
Transportation (air, ferry, or coach tickets)	4	1.7	12,384	49,536
Travel package	32	13.4	72,547	2,321,492
Leisure and attraction	9	3.8	74,949	674,545
Spa and beauty	60	25.2	43,281	2,596,860
Total	**238**	**100**	**62,651**	**14,910,903**

Deals ending on Saturdays generated the highest revenue per deal (HK$128,901). Again, deals ended on Sundays sold the least number of coupons per deal (128 coupons) and generated the least revenue per deal (HK$12,184). Deals ended on Wednesday generated the highest total revenue (HK$2.4 million) and again, those ending on Sunday generated the least total revenue (HK$48,736).

Table 3. Average number of transactions and average revenue generated for deals started and ended on different days of the week

Deals started on	No. of deals started on the day	%	Avg sales (avg no. of coupons sold per deal)	Avg rev generated per deal (in HK$)	Total rev (in HK$)
Mon	30	17.9	688.70	68,788	2,063,640
Tue	30	17.9	482.33	44,531	1,335,930
Wed	27	16.1	578.56	63,478	1,713,906
Thu	28	16.7	921.75	62,175	1,740,900
Fri	28	16.7	680.25	55,191	1,545,348
Sat	18	10.7	611.17	82,372	1,482,696
Sun	7	4.2	54.14	24,916	174,412
Total	**168**	**100**	**636.83**	**59,862**	**10,056,816**
Deals ended on	**No. of deals ended on the day**	**%**	**Avg sales (avg no. of coupons sold per deal)**	**Avg rev generated per deal (in HK$)**	**Total rev (in HK$)**
Mon	35	20.8	448.63	19,434	680,190
Tue	37	22.0	1,007.57	63,100	2,334,700
Wed	35	20.8	541.34	69,987	2,449,545
Thu	23	13.7	580.87	29,604	680,892
Fri	25	14.9	618.20	66,107	1,652,675
Sat	9	5.4	637.00	128,901	1,160,109
Sun	4	2.4	127.75	12,184	48,736
Total	**168**	**100.0**	**636.83**	**59,862**	**10,056,816**

4.2 Features of the Different Deal Categories

Table 4 shows the means of the different deal features for the different product/service categories. The average special offer price (discounted price) for all deals is HK$336. Deals for travel packages and accommodation were the most expensive ones, averaging HK$1,309 and HK$1,013, respectively. The least expensive deals were the ones from non-Chinese quick service restaurants (HK$24). All deals offered an average of 64% discount off their listed prices. Spa and beauty services offered the largest discount of close to 90%, while Chinese quick service restaurants only offered a 43% discount off their listed prices. On average, each deal was available on the group-buying sites for purchase for an average of 6 days. The redemption duration was on average 98 days with Spa and beauty products having the longest duration of 131 days and travel package of 48 days. The deals were featured in an average of 1.3 group-buying sites and 6.3 aggregation sites. The deals on average required a minimum of 3.4 coupons to be sold in order for the deal to be on. Deals for transportation only required one person to purchase a coupon in order to activate the deal. Most of the deals sold were from non-Chinese quick service restaurants (2,340 coupons), followed by leisure and attraction (1,047 coupons).

Table 4. Features of the deals

Deal category	Offer price (in HKD)	Dis-count %	Pro-motion duration	Redem-ption dura-tion	No. of sites featuring the deal	No. of aggregate sites featuring the deal	Min. no. of per-sons to sign up	Sales (no. of deals sold)
Non-Chinese quick service restaurant	23.94	55.98	7.08	94.53	1.71	6.18	2.82	2,339.65
Chinese quick service restaurant	41.33	43.38	7.17	102.17	2.33	6.00	2.50	202.50
Non-Chinese full service restaurant	109.91	59.68	5.42	99.69	1.143	6.83	4.07	991.29
Chinese full service restaurant	159.81	51.61	5.31	78.08	1.46	6.58	4.62	605.23
Accommodation/hotel	1013.22	53.98	5.67	105.00	1.00	6.33	1.44	184.67
Transportation	523.00	45.73	7.33	116.25	2.50	5.50	1.00	43.00
Travel package	1309.03	50.81	6.38	48.28	1.25	5.44	2.22	76.56
Leisure and attraction	152.67	53.44	6.57	95.56	1.00	7.22	3.00	1,047.33
Spa and beauty	210.00	89.98	6.18	131.42	1.07	6.13	3.38	241.55
Average	**336.67**	**64.09**	**6.02**	**98.01**	**1.34**	**6.34**	**3.39**	**669.22**

4.3 Factors Affecting Sales of the Deals

Results of multiple regression analyses are presented in Table 5. The R^2 shows that the model explains 63.7% of the observed variability in the number of transactions by the 19 independent variables. The significant F ratio suggested the satisfactory level of the goodness-of-fit of this regression model. Two of the variables, global geographic coverage of the site and non-Chinese full service restaurant are excluded in the model due to multicollinearity. Eight of the independent variables in the model were found statistically significant with an alpha level of >0.05. Thus, they were significant determinants of the sales of the deal. The standardized coefficient Beta values indicated the relative importance of each of the independent variables. According to the results, the number of Facebook fans a group-buying site has was the most important factor in contributing to the sales of the deals (β=.397), followed by the special offer price for the deal (β=-.378). The special offer price has an inverse relationship with the sales as a lower price is always related to a higher amount of sales. Moreover, the number of aggregation group-buying sites featuring the deal and the length of coupon redemption period also have a significant positive relationship to the sales with β =.199 and β=.159, respectively. The other four significant factors are related to the types of product/service which the deal belongs to. They include Chinese quick service restaurant (β=-.129), accommodation/hotel (β=-.179), travel package (β=-.207), and spa/beauty (β=-.315). This indicates that if the deal is any of these four types of products/services, the sales will be negatively affected and spa/beauty services have the highest magnitude on the negative impact to the sales of deals among all others.

Table 5. Factors affecting the number of deals sold

Independent variables	B	Standardized coefficient β	t	Sig.
(Constant)	6.431		4.680	.000
Geographic coverage: Hong Kong only	.195	.049	.254	.800
Geographic coverage: Guangdong province	.255	.047	.445	.657
Special offer price for the deal (Transformed)	-.580	-.378	-.4370	.000**
No. of group buying sites featuring the deal (Transformed)	-.304	-.061	-.901	.369
Minimum no. of customers to purchase the deal (Transformed)	-.160	-.087	-1.355	.178
Discount percentage	-.015	-.161	-1.553	.123
Promotion duration	.044	.068	1.013	.313
Length of coupon redemption period	.006	.159	2.390	.018*
Deal category: Non-Chinese QSR	-.082	-.012	-.187	.852
Deal category: Chinese QSR	-1.216	-.129	-2.056	.042*
Deal category: Chinese FSR	-.431	-.072	-1.208	.229
Deal category: accommodation/hotel only	-2.052	-.179	-3.156	.002**
Deal category: out-bound transportation	-6.53	-.040	-.712	.478
Deal category: travel package	-1.173	-2.07	-2.917	.004**
Deal category: leisure and attractions	.106	.012	.211	.833
Deal category: Spa and beauty	-1.267	-.315	-2.924	.004**
Number of aggregation sites featuring the deal	.221	.199	2.772	.006**
No. of Facebook members	6.967E-6	.397	2.270	.025*
Rewards and incentives for referral	-.003	.000	-.011	.991

*Significant at .05, **Significant at .01
R = .798 R^2 =.637 Adjusted R^2 = .587 F =12.743 Sig. =.000

5 Discussion

The global market leader Groupon continues its success in Hong Kong, generating a lot more sales in hospitality and travel related deals than its competitors including Travelfish which is well known as "travel arrangement specialist". Groupon's famous brand, its experience in group-buying and its worldwide geographic coverage could be their key to success. Restaurants (especially non-Chinese full service ones) as the most popular deal category in revenues earned and the number of deals offered may account for the habit and preference of dining out in Hong Kong as well as the aggressiveness of certain establishments in seeking new ways to attract customers.

Higher sales may be made on both the starting day and ending day of a deal due to the arrival of the new deal and the time pressure on purchase before the end of the deal. Group-buying websites seem to be aware that Sunday is the worst day of the week for both starting and ending a deal in terms of average number of sales transactions and average revenue per deal and decide to place the least number of deals starting or ending on this relatively unfortunate day. On the other hand, Saturday is the champion in terms of average revenues per deal as both a starting day and an ending day. The ups on Saturday could be due to pre-holiday mood for online shopping while the downs on Sunday may be caused by the habit of holiday outings where computer or internet is not available.

The high average discount and long redemption duration offered by spa and beauty services prove the stronger determination of the suppliers in attracting customers. Spa and beauty service suppliers do not seem to target on existing customers but the new ones as their deals are often tied with a number of terms and conditions. For example, coupons are only applicable for new customers, and a maximum of five coupons purchased per person but one coupon only usable by one person. Once the deal has enticed new customers to a spa and beauty center, there would be chances for their staff to recommend and sell additional services during consultation at their visit. As such, the center's lost in offering their group buying deal at low price may be offset.

The average minimum number of persons to sign up set for all deals is incredibly low and contradictory to the original concept of group-buying. Although the word "group-buying" appears on the websites, now sounds more like a gimmick than an actual criteria for dynamic discount offered by the merchants. A number of deals being confirmed as long as there is a single sales transaction made prove that suppliers are keen on attracting customers regardless of the quantity. Merchants (especially new businesses or brands that are not familiar to the public) may see luring customers to try their products or services and reducing excessive inventory as their higher priority in their group-buying deals comparing to maintaining a high profit. Besides, similar to spa and beauty services, personal selling can also take place in other suppliers like restaurants to help offset the lost at the group-buying discount. There is also a chance that some discounts are exaggerated just to create a 'wow' effect and suppliers may not really suffer even if their products or services are sold poorly at low price.

The number of Facebook fans the group-buying site has emerged as the most crucial factor to the sales. This result agrees with Hill *et al.*'s study (2006) on the power of social network connections on business increment and there is no exception to the hospitality and travel related products or services. With the growing popularity of smart phones, word-of-mouth and referrals can take place anywhere and anytime, so can sales transactions. Group-buying websites and their suppliers can particularly be benefited from this real time communication as group-buying deals are often sudden and temporary.

The significant positive impact of the length of coupon redemption period and the number of aggregation sites featuring the deal support the results of Krisnhna and Zhang (1999) on the influence of coupon redemption duration on coupon sales. Similar findings were reported in Coehlo and Easingwood (2004) on the impact of exposure of products on sales volume sold via multiple channels. Since hospitality and travel related products or services are perishable, customers may find such product deals with a long redemption period more valuable and attractive. Deals being featured in aggregation sites can increase their exposure to potential customers and lead to higher sales. By visiting an aggregation site, customers can view all the deals from a number of individual group-buying sites linked to it and some advanced aggregation sites can even categorize the deals by product types for the convenience of customers who are interested in specific types of deals. Time-conscious customers may prefer to visit the aggregation sites instead of viewing the individual group-buying sites so higher sales should usually be obtained when the deal being featured in aggregation sites.

Supported by Kauffman and Wang's study (2001a) on the increase of sales due to the falling of price, a strongly inverse relationship of the special offer price with the sales is found. Dynamic discount stated on group-buying websites can sometimes be exaggerated by magnifying the face value (reference price), but dynamic offer price never lies. As a significant determinant to sales volume, the special offer price may represent the true value of a deal in the mind of group-buyers. Budget consumers may dominate the market of group-buying and their spending power may limit the sales of expensive products or services on group-buying websites.

Supported by the study of Brown *et al.* (2003) on the impact of product type on online sales, spa and beauty, travel package, accommodation/hotel and Chinese quick service restaurant have negative impacts on the sales among the others. The price of spa and beauty deals on group-buying websites may be unrealistically too low (almost 90% off on average) considering the prestige nature of the service. Rosenbloom (2004) suggested that there were "price points" which consumers had become accustomed and consumers expected products sold at these customary prices. Customers may find a deal "too good to be true" and worry about deception if the price is much lower than their expectations. Another reason can be either that spa and beauty treatment consumers are loyal or have prepaid their spa and beauty centers they use, so they may not easily switch to another brand. The consumption of travel packages and accommodation/hotel is usually longer in duration and more distant in terms of location than other categories. Thus, longer planning may be required, especially when these deals are often offered to groups of two persons. The deal period may be too short for potential customers to make up their mind on purchasing such products or services. Finally, Chinese quick service restaurant deals having the relatively lowest discount among other categories. This finding agrees with the studies by Compeau and Grewal (1998) as well as Alford and Biswas (2000) on how the difference between reference price and sales price is influencing the sales. Potential customers may not find such deals attractive comparing to others.

6 Recommendations and Conclusions

When selecting a group-buying website, suppliers should not primarily base on whether the site is specialized in selling the suppliers' type of products but other determinants, like branding, experience and geographic coverage of the market if they want to ensure the promised sales. They can consider bargaining for a lower commission fee to group-buying websites for deals started and ended on Sunday due to the expected relatively poor average sales and revenue per deal. On thre other hand, suppliers offering products or services at a high price range or targeting a high volume of sales should request their deals to start and end on Saturday due to relatively high average revenue per deal.

With the number of Facebook fans being the most important determinant to the sales, group-buying websites should join the popular online social networks and recruit as many subscribers of social network as possible through incentives, such as lucky draw, sign up dollar credits or other promotional tools. Suppliers should select to place their deals on group-buying websites which have joined the popular social networks with a large number of subscribers in order to increase the exposure of their deals as well as their sales.

As the longer the redemption period, the higher the sales of coupons, hospitality and travel related product suppliers should arrange a longer redemption duration if possible to increase the product value in the mind of potential customers due to the perishable nature of the services. Due to the positive relationship of the number of aggregation sites featuring the deal with the sales volume, group-buying websites should link themselves to more aggregation sites. The commission fee paid to these aggregation sites should be rewarding as they provide convenience to time-conscious potential customers and help expose the deals to a bigger market.

With the special offer price having a strongly negative impact on sales, group-buying websites should emphasize more on the offer price on their websites as well as promotional messages to their subscribers. An exaggeration of discounts and face value misleading customers can, and should, be avoided. Rather, when selecting suppliers, group-buying websites should consider the ones with low-end products first as budget customers being the main group-buying market. High-end product suppliers should consider alternative intermediaries other than group-buying websites if they expect a high volume of sales.

Product types such as travel packages and accommodation/hotel with longer consumption periods and a more distant location for consumption have an inverse relationship with the sales due to long planning before purchase. Group-buying websites should thus, arrange longer promotion duration for such products for potential customers to allow thorough consideration. Single traveller package/accommodation can also be introduced if suppliers prefer to shorten the decision time of potential customers. Alternatively, group-buying websites can notify subscribers ahead via social network or e-mail when the next travel package or accommodation/hotel deal is coming so that longer planning by potential customers can take place. Spa and beauty services also have a negative impact on sales. Distribution channels emphasizing dynamic discounts may not be ideal to services with such prestige in nature. Finally, spa and beauty centers should consider other alternative middlemen.

7 Limitations and Future Research

The limitations of this study should be noted when interpreting the results. The small sample size of 238 deals may not be large enough to generalise the entire group-buying business on hospitality and travel related products or services. Future studies should, thus, consider collecting data from a bigger sample and for a longer period in multiple cities and further comparisons among different markets can be examined.

References

Alford, B.L. & Biswa, A. (2000). The effects of discount level, price consciousness and sale proneness on consumers' price perception and behavioral intention. *Journal of Business Research* 55: 775-783.

Anand, K.S. & Aron, R. (2003). Group buying on the web: A comparison of price-discovery mechanisms. *Management Science* 49(11): 1546-1562.

Brown, M., Pope, N. & Voges, K. (2003). Buying or browsing? An exploration of shopping orientations and online purchase intention. *European Journal of Marketing* 37(11/12): 1666-1684.

Chen, J., Chen, X., Kauffman, R.J. & Song, X. (2009). Should we collude? Analyzing the benefits of bidder cooperation in online group-buying auctions. *Electronic Commerce Research and Applications* 8: 191-102.

Chen, J., Chen, X. & Song, X. (2006). Comparison of the group-buying auction and the fixed pricing mechanism. *Decision Support System*s 43: 445-459.
Chen, R.R. & Roma, P. (2011). Group buying of competing retailers. *Production and Operations Management* 20(2): 181-197.
Chiang, C., Lin, C.S. & Chin, S.P. (2010). Optimizing time limits for maximum sales response in Internet shopping promotions. *Expert Systems with Applications* 38: 520-526.
Coehlo, F.J. & Easingwood, C. (2004). Multiple channel system in services: Pros, cons and issues. *Service Industries Journal* 24(5): 1-29.
Compeau, L.D. & Grewal, D. (1998). Comparative price advertising: An integrative review. *Journal of Public Policy and Marketing* 17(2): 257-273.
Groupon (2011). Retrieved 5 September 2011 from http://www.grouponworks.com/why-groupon/comparison-guide.
Hair, J., Black, W., Babin, B. & Anderson, R. (2010). *Multivariate Data Analysis*, 7Ed, Prentice Hall.
Hill, S., Provost, F. & Volinsky C. (2006). Network-based marketing: Identifying likely adopters via consumer networks. *Statistical Scie*nce 21(2): 256-276.
Kaufmann, R.J., Lai, H. & Lin, H.C. (2010). Consumer adoption of group-buying auctions: experimental study. *Information Technology and Management* 11: 191-211.
Kaufmann, R.J., Lai, H. & Ho, C.T. (2010). Incentive mechanisms, fairness and participation in online group-buying auctions. *Electronic Commerce Research and Applications* 9: 249-262.
Kaufmann, R.J. & Wang, B. (2001a). New buyers' arrival under dynamic pricing market microstructure: The case of group-buying discount on the Internet. *Proceedings of the 34th Hawaii International Conference on System Sciences 2001*, pp. 1-10.
Kaufmann, R.J. & Wang, B. (2001b). Bid together, buy together: On the efficacy of group-buying business models in the Internet-based selling. *The (R)evolution Goes Mobile, 5th Annual University of Minnesota Electronic Commerce Conference, March 27-28, Carlson School of Management, University of Minnesota, Minneapolis, MN*, pp. 1-44.
Krishna, A. & Zhang, Z.J. (1999). Short- or long-duration coupons: The effect of the expiration date on the profitability of coupon promotion. *Management Science* 45(8): 1041-1056.
Kotler, P. Bowen, J.T. & Makens, J.C. (2006). *Marketing for Hospitality and Tourism (4th edition)*, pp. 219-221. Pearson Education, New Jersey.
Law, R., Chan, I. & Goh, C. (2007). Where to find the lowest hotel room rates on the Internet? The case of Hong Kong. *International Journal of Contemporary Hospitality Management* 19(6): 495-506.
Leung, Y.H. (2011) Yahoo! Deals offering ten 50% off deals daily. Hong Kong Economic Times, June 2, 2011. Retrieved 2 September 2011 from http://www.hket.com/eti/article/bdbc05ea-e3e3-4db2-b467-7e4cb54c29d5-121985
Li, H., Zhao, P., Wang, Y. & Wang, G. (2009). A qualitative research of Tuangou: Modes, characteristics and roles of the new e-business model. *2009 International Symposium on Information Engineering and Electronic Commerce,* pp. 750-753.
Roehm, M.L. & Roehm, H.A. Jr. (2011). The influence of redemption time frame on responses to incentives, *Journal of the Academy of Marketing Science* 39: 363-375.
Rosenbloom, B. (2004). *Marketing Channels – A Management View (7th edition)*, pp. 343-345. South-Western, Cengage Learning.
Schotanus, E. & Telgen, J. (2007). Developing a typology of organisational forms of cooperative purchasing. *Journal of Purchasing & Supply Management* 13: 53-68.
Stafford, P. (2011). Microsoft enters group-buying game with discounts aggregator, Bing Deals. Retrieved 5 September 2011 from http://www.smartcompany.com.au/information-technology/20110304-microsoft-enters-group-buying-game-with-discounts-aggregator-bing-deals.html
The Economist (2011). Online-coupon firms: Groupon anxiety. *The Economist*, March 17, 2011. Retrieved 5 September from http://www.economist.com/node/18388904.
Wei, Y., Straub, D.W. & Poddar, A. (2011). The power of many: an assessment of managing internet group purchasing, *Journal of Electronic Commerce Research* 12(1): 19-43.
Yau, T. (2011). Group-buying business rises under inflation. *Wen Wei Po*, January 3, A6.
Yoon, M.G., Yoon D.Y & Yang, T.W. (2006). Impact of e-business on air travel market: Distribution of airline tickets in Korea. *Journal of Air Transport Management* 12: 253-260.

New Tools for Studying Visitor Behaviours in Museums: A Case Study at the Louvre

Yuji Yoshimura[a], Fabien Girardin[b], Juan Pablo Carrascal[a], Carlo Ratti[c], and Josep Blat[a]

[a] Information and Communication Technologies Department
Universitat Pompeu Fabra, Spain
{yuji.yoshimura, jp.carrascal, josep.blat}@upf.edu

[b] Lift Lab
Switzerland
fabien@liftlab.com

[c] SENSEable City Laboratory
Massachusetts Institute of Technology, USA
ratti@mit.edu

Abstract

In this paper we discuss the exploitation of data originated from Bluetooth-enabled devices to understand visitor's behaviour in the Louvre museum in Paris, France. The collected samples are analysed to examine frequent patterns in visitor's behaviours, their trajectory, length of stay and some relationships, offering new details on behaviour than previously available. Our work reinforces the emergence of a new methodology to study visitors. It is part of recent lines of investigation that exploit the presence of pervasive data networks to complement more traditional methods in tourism studies, such as surveys based on observation or interviews. However, most past experiments have explored quantitative data coming from mobile phones, GPS, or even geo-tagged user generated content to understand behaviour in a region, or a city, at a larger scale than that of our current work.

Keywords: Bluetooth sensing; human behaviour; museum study; real time management tool.

1 Introduction

In recent decades, tourism has developed to become one of the biggest industries. The World Tourism Organization foresees that the number of tourists will reach 1,600 million around 2020 and the World Travel and Tourism Council predicts that direct/indirect economic impact generated by the touristic industry will amount to 9.6% of Gross Domestic Product (GDP) and generate 9.7% of employment all over the world in 2012[1].

The increase of its economical, cultural and social impact on urban areas requires more precise and dynamic understanding of tourist behaviours and movements at micro (e.g. district, city) and macro (e.g. region, country) scales. Some emerging technologies make it possible to record and analyse them at city and district level (e.g. GPS, mobile phones with or without GPS (Asakura & Iryob, 2007); the passive mobile positioning data (Ratti, Pulselli, Williams, & Frenchman, 2006, Ahas, Aasa, Roose, Mark, & Silm, 2008); user-generated data (Girardin, Dal Fiore, Ratti, & Blat,

[1] See: http://www.wttc.org/eng/Tourism_Research/Economic_Research/

2008, Pereira, Vaccari, Girardin, Chiu, & Ratti, 2012, Girardin, Calabrese, Dal Fiore, Ratti, & Blat, 2008). In museums, the observation, and interview-based surveys have been used mostly to understand the social use of the environment and evaluate its use (see Hooper-Greenhill, 2006 for a review and Yalowitz & Bronnenkant, 2009, Hillier & Tzortzi, 2006). The information collected by these "traditional" methods provides support for the management of the spaces, which have a proved value. However they often provide a snapshot on the life of a built environment, and the interviews and questionnaires can have self-reporting bias. Moreover, they fail to record empirical evidences and measures (e.g. visiting time, sequences of visits, time of stay, density) key to produce a more complete picture on people use of a space.

The purpose of this paper is to discuss a Bluetooth proximity detection approach (previously developed for a traffic data collection system (Sanfeliu, Llácer, Gramunt, Punsola, & Yoshimura, 2010) to gather insights on visiting behaviours in a museum context and to demonstrate its relevance to support the management of environments that must respond to the increasing tourism demand. For instance, the analysis reveals the dynamic description of different use of museum spaces, the visiting profiles and the spatio-temporal patterns of visitors' behaviours.

In section 2 a brief summary of related works, their contributions and their main limitations is provided. In section 3 a Bluetooth proximity approach to detect visitor's presence and sequential movements is proposed. In section 4 the dataset processing for the analysis is discussed and key concepts for our research introduced. Section 5 presents some initial findings from our field trials, the frequent pattern and visitors' spatial uses. Finally, we summarize our on-going work on developing methods and tools for analysing the museum and urban environments.

2 Strategies to Collect Empirical Visitor Data

With the emergence of location technologies, a variety of methodologies have been proposed to locate a person, specifically for the collection of empirical data in the context of tourism (Asakura & Iryob, 2007, Ratti et al., 2006, Ahas et al., 2008, Girardin, Calabrese et al., 2008, Girardin, Dal Fiore et al., 2008, Yalowitz & Bronnenkant, 2009, Hillier & Tzortzi, 2006, Kanda et al., 2007). They are classified into 3 groups, and remark the burden which each method imposes on the persons involved (Table 1).

The first group of more traditional techniques includes observation, and interviews. With the latter or with user diaries, one can obtain specimens of detailed visitor's behaviour. However, the data can be subjectively biased and the methods are costly requiring a lot of human resources (Girardin, Dillenbourg, & Nova, 2009). Something similar happens with direct observation, which could be difficult to sustain for long periods as it poses a heavy burden to the observer. The representativeness of the sample in interviews and questionnaires can be an issue too.

The second group is based on technologies such as GPS or RFID, which can supply more objective and precise time, location and route data (Asakura & Iryob, 2007, Kanda et al., 2007) than the traditional methods – however, without the motivations which can appear in interviews. Currently, these techniques demand the users to carry

specifically enhanced devices that are not widespread. They make the data collection more cumbersome, and they may bias the user behaviour, and thus the collected data.

A third group includes using image sensing devices or passive mobile positioning data, which give little burden to the users – but no motivations are available either. Their main limitation is spatial; for instance, image sensing devices can record visitor's behaviour with spatio-temporal accuracy (Antonini, Bierlaire, & Weber, 2006), but the recording area covered by a single camera is limited (Yalowitz & Bronnenkant, 2009). Passive mobile positioning data have started to be used in tourism studies (Ratti et al., 2006, Ahas et al., 2008) as it can provide better empirical data on movement of people at global scale; nevertheless, the estimation of the presence and movement of people is limited by the cell size (i.e. the area of coverage of the base station that serves the mobile service).

Table 1. Data capture techniques showing their main strengths and weaknesses in the context of tourism and urbanism studies

Data capture	Strengths	Weaknesses	Application example
Manual surveys	Capture motivations	Very costly and applied to a limited time period	Timing and Tracking Survey (Yalowitz & Bronnenkant, 2009)
GPS and Cell phone (device-based)	Timely mobility data (potentially augmented with in-situ survey)	Survey limited in time and participants. It does not work inside the buildings	Describe social and spatial characteristics with limited samples (Asakura & Iryob, 2007)
RFID	Precise real-time mobility data	Survey limited in time and participants. Infrastructure deployment needed	Describe social and spatial characteristics with limited samples (Kanda et al., 2007)
Cell phone (aggregated network-based)	Use existing infrastructure to provide real-time mobility data	Does not work at the building and room scale	Real-time urban dynamics (Ratti et al., 2006)
Bluetooth detection	Precise real-time mobility data, non-intrusive to participants	Infrastructure deployment needed	Describe social and spatial characteristics (Kostakos, O'Neill, Penn, Roussos, & Papadongonas, 2010)

This paper presents several contributions in the development of data collection tools and methodologies for the analysis of large samples describing visitor's behaviour at small spatial scale using Bluetooth. The recent wide spread of mobile devices implies that many people have their Bluetooth switched on passively, thus providing an important source of useful data. A variety of projects have exploited Bluetooth data for measuring the social network relationships between people (Eagle & Pentland, 2005, Paulos & Goodman, 2004, Nicolai, Yoneki, Behrens, & Kenn, 2006), mobility of vehicles (Yalowitz & Bronnenkant, 2009, Barceló, Montero, Marqués, & Carmona, 2010) and mobility of pedestrians and their relationships (O'Neill et al., 2006, Kostakos et al., 2010). However these investigations have not considered a specific analysis of pedestrians and their use of space. This paper aims at reducing this shortcoming.

3 Data Collection Settings

A large majority of mobile devices currently on the market embed Bluetooth, and a significant proportion of users have them turned on in passive mode (Kostakos et al., 2010). The presence of these Bluetooth-enabled devices can be detected by means of sensors that scan the wireless spectrum. This section and the following describe the settings of our study and how the collected data were structured to handle privacy issues and allow for the spatial behavioural analysis.

3.1 Context of the Study

The Louvre is the most visited museum in the world with 8.5 million visitors in 2009 and more than 40,000 visitors at peak days[1]. This context of "cultural enthusiasm" has direct consequences on the quality of the visitor experience as well as on the organization and management of the Museum (e.g. application of flow management strategies and increased stress level of the surveillance staff). In response to the increasing tourism demand and the necessity to setup and evaluate museum strategies, we proposed to collect and analyse empirical data on the flows and occupancy levels of visitors in key areas of the Louvre.

3.2 Study Settings and Characteristics of the Bluetooth Sensors

Because of the context, our study is particularly focused on one of the busiest areas of the museum identified by Le Louvre officials, namely a trajectory that leads visitors from the entrance (Pyramid) to the Venus de Milo. 10 Bluetooth sensors were deployed and they were sufficient to gather measures of visiting sequences and staying times at representative locations along the path (Figure 1). Two were on floor -1 (0 or *Hall*, 1 or *Denon access*); five were on floor 0 (2 or *Denon* 0, 3 or *Samothrace* 0, 4 or *Venus de Milo*, o 5 or *Caryatides*, 6 or *Sphinx*), and 3 on floor 1 (7 or *Big Gallery*, 8 or *Samothrace* 1, 9 or *Glass*).

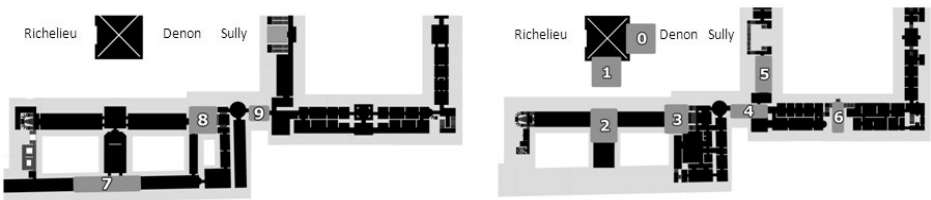


Fig. 1. Location of 10 sensors (No.0-No.9) indicating their approximate sensing range

The administrative and technical restrictions (e.g. protection against robbery, areas unreachable to visitors, no sources of electrical power, safety and health concerns) guided the deployment of the devices, sometimes preventing the installation in ideal locations for optimal detection. These special circumstances required the use of an ad-hoc battery with 10 days of autonomy for each sensor. This temporal limitation constrained the accumulation of empirical data, but our analyses show that this period is sufficient to extract relevant evidences.

[1] See: http://www.theartnewspaper.com/attfig/attfig10.pdf

The sensors gathered a unique encrypted identifier distinguishing each mobile device that supports Bluetooth and is set to be discoverable, as well as 2 time stamps for check-in and check-out times within the range of each sensor. Assuming that a mobile device belongs to a person, the movement of the device can be related to that of the visitor.

3.3 Data and Privacy Issues

Based on a previous research that focused on the privacy issues related to the use of Bluetooth scanners (Sanfeliu et al., 2010), we adopted a solution that 1) does not allow the identification of individuals, 2) keeps the anonymity of trajectory data even after recording and archiving. This is achieved with the application of Secure Hash Algorithm (SHA) to the Bluetooth unique IDs detected by our system.

3.4 Sensor Detectable Area and the Definition of its Node

The spatial definition of the detectable area by a Bluetooth sensor is a critical issue for any research which uses this type of sensors. The shape of the area is similar to a flower with four petals of different length and width. In an optimal setting, the largest petal is an ellipse of almost 40 meters long by 15 meters wide, while the smallest is approximately 15 by 10 meters. The other two have a similar shape and a size of 15 by 10 meters. However, it could be customized for an indoors space with the largest petal dimensions being 20 by 7-8 meters. We identify the area detectable by a sensor as a node which represents the corresponding location, and use this definition through the rest of the paper. The detectable area estimations fluctuate according to the museum settings, due to the location of the sensors (e.g. within wooden boxes or administrative desks) and other factors, but we made sure that they would cover the targeted areas along the studied visitor trail (Fig. 2).

Fig. 2. Conceptual diagram of Bluetooth sensor's detectable area

4 Collected Data and Measures

Based on the methodologies proposed in previous sections, the sample data during a specific audit period (more on this in section 4.2) was collected. This section describe how these amounts of collected sample data were organized to extract the desired values, and two measure concepts, *length of stay* and *trajectory* are defined.

4.1 Database

The raw dataset collected from all sensors is huge and requires pre-processing for us to be able to extract meaningful information from it. These data basically consist of a unique encrypted identifier for every mobile device and two timestamps, which correspond to the first and last times such a device has been detected by the sensor. Then a database and a query engine were built to reorganize these data for our analyses (see table 2 for an example). Let us indicate some of the tags used for organizing and analysing the raw data. *Rffr* is the unique encrypted identifier. *Date* is the year, month and day when the data were collected. *Path* indicates the nodes that a mobile device has visited and it is represented by a sequence of node numbers, from 0 to 9, separated by a colon (":"). *Nodes* represents the total number of nodes that a device has visited during its whole trajectory, while *distinct nodes* indicates the number of different nodes which a mobile device has visited. *Checkin* is the moment when the signal of a mobile device is first detected in the museum (i.e. at the first node of the trajectory) and *checkout* is the moment when it disappears from the last node (i.e., when the device has left the museum). *Staylength* is the time difference between *Checkout* and *Checkin* and it represents the total duration of stay of a mobile device inside the museum.

Table 2. Example of the data set.

Rffr	Date	Path	Distinct nodes	Nodes	Checkin	Checkout	Staylength
Unique ID	2010-04-30	0:3:8:7:0	4	5	09:04:35	11:07:52	02:03:17

4.2 Collected Sample

A high frequentation 10-day period in May 2010 was selected to perform a first analysis of visitor's behaviour. During this audit period, our installation recorded the presence of 12,944 unique devices. Through the data cleaning process we removed the logs from security and museum staff by looking at their recurrence, and the time of their presence (e.g. outside visiting times). Also, it was found out that the logs from two sensors had erroneous time synchronization and had to be discarded. Indeed, synchronization is a key element of our approach, for instance to infer the sequence of visit.

4.3 Measures Definition

A sensor log reveals the visitor's presence at a node: once a Bluetooth-enabled mobile device enters the detectable area, the sensor continues to receive the signal emitted from the device until it disappears from its range. Each sensor records the first time the device appears as a check-in time at then records the time when the signal of the device disappears, as the checkout time. The difference between both time stamps is the length of the stay at the node. If the nodes visited are ordered by time, and then the check-in time at the first node in the trajectory and the checkout time at the last are selected, the values of the total duration of the visit to the museum will be obtained. As it can be seen, synchronization of sensors plays a key role for the collected data to be meaningful.

On the other hand when a unique Bluetooth identifier is logged out with time stamp (t1) by sensor A and some time later is logged in with time stamp (t2) by sensor B, the difference between t2 and t1 measures the travel time. The sequential movement of a mobile device detected by a pair of sensors (e.g. A-B) is defined as a trajectory with the travel time of t2-t1 minutes. The concept of trajectory in our research is different from that obtained with GPS systems, which indicate precise locations. The trajectories are obtained from Bluetooth detection through the time stamped sequential transition of a mobile device detected through different nodes (e.g. sequence of A-B-D), while GPS can describe the precise movement of the device. Our measures are, in this way, indirect ones.

5 Results

Using the Bluetooth data and concepts described previously, a novel approach is developed to analyse the spatial use in the Louvre museum. The following subsections present the on-going analysis efforts built around these concepts and the initial findings in order to obtain indicators for crowd management and to extract the frequent patterns in visitors' behaviour. In 5-1 the representativeness of the sample captured by our sensors is discussed. In 5-2 an analysis of the use of the Pyramid space related to the visitors' trajectories, which may reveal the distribution of visitors' presence and its basic flow in the museum, as all the visitors use the Pyramid as entrance and exit is presented. In 5-3 an analysis of visitors' trajectories and the time spent in each route is presented, revealing the existence of frequent patterns according to these two parameters.

5.1 Representativeness of the Collected Sample

Only a part of the visitors have got devices with Bluetooth, and only a part of them are enabled so that they are detectable. The number of devices detected at the entrance is compared with the official museum head counts and ticket sales to understand the representativeness of our Bluetooth data. The sample represented between 5.9% and 8.7% of the visitors with a strong, positive correlation of +80%, providing support for its representativeness. Figure 3 shows the linear regression fit of the numbers of detected devices and official counts of each day, with data of 101 days.

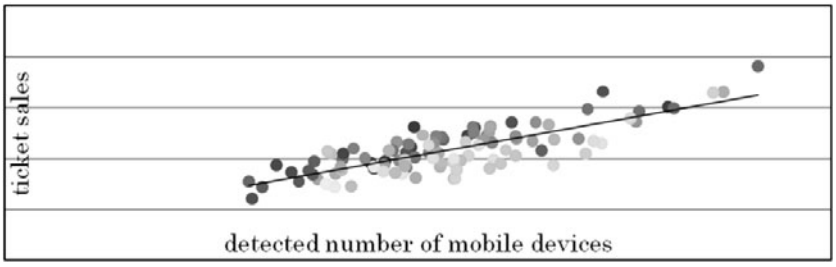


Fig. 3. Correlation between detected devices and visitors estimations per day

5.2 Use of the Pyramid Space

The Pyramid space serves for distributing the visitors through three museum accesses, named *Denon*, *Sully* and *Richelieu*. As it is the starting point for almost all the

museum visitors, it is important to identify and analyse their spatial use in order to devise more efficient and flexible policies for the museum. Since all the sensors are installed on the trails that lead to the Venus de Milo, along the Denon area, mobile devices detected by sensors in nodes 0 to 8 represent people who visited such area. These are the data we deal with in this section.

Sensor 0 is the Hall, and due to the museum's spatial layout, routes 0-3, 0-4, 0-7 and 0-8 mean that the *Denon access* has been used; routes 0-5, 0-6 and 0-9 mean use of either *Sully* or *Richelieu access*; route 0-0 means that only the Sully or Richelieu areas were visited. Our data indicate that 76% of visitors used the *Denon access* while only 23% used either of the other two.

However, if one focuses on visitor's exit behaviour, i.e. moving towards node 0 or the Hall, the spatial use tendency changes. The most used route leading to node 0 was the 3-0 (25%), followed by the 7-0 (around 17%) and the 5-0 (around 16%); around 40% of the visitors left the museum through the *Sully* or *Richelieu* access, while 60% used *Denon* as their exit route, which means a decrease of an absolute 16% of the latter.

5.3 Visitor's Trajectories

In this subsection, visitors' trajectories, their average length of stay and their relationships are analysed to discover frequent patterns or trends in visitors' behaviours. Sequential pattern mining (Agrawal & Srikant, 1995) has received much attention in the recent decade to find frequent sequences of events in data with a temporal component, with transition probabilities between events. However, extracting meaningful patterns requires appropriate algorithms and parameters. In the following, the grounding work and initial findings of our analysis are presented.

Most used trajectory and visitor's transition rates between nodes. Clarifying visitors' most used trajectory and their nodes transition rates helps to uncover hidden rules behind the seemingly disordered dataset. The data correspond to the route starting by 0-8, which is used by 60.6% of the visitors (7721 devices). Within the people who took this route, 71.1% moved to node 3 (0-8-3), while 13.5% went back to node 0 (0-8-0); 7.2% do not have any further records after node 8 (0-8), and it is assumed that they should have finished their itineraries without being registered by node 0 again. This is because if they would have continued visiting the museum, some of the other sensors should have detected them. Thus, 20.8% of these visitors (12.6% of the total) came to the museum just to take the 0-8-0 route. The popularity of this route might be due to the presence of two major works, Mona Lisa, located between nodes 7 and 8, and Winged Victory of Samothrace (in node 8). Moreover, the spatial structure of the museum strengthens the link between those two works, and thus an important sequential pattern including a visit to the Mona Lisa, followed by the Winged Victory of Samothrace, and then the Italian Gallery might exist. Let us perform a more detailed analysis considering the objects in the spatial structure, to reveal patterns.

The visitors' distribution rate from each node to every subsequent node is iteratively computed until the route finished at 0. The percentages that appear with each arrow (Fig.4) are these transition rates from each node to the next ones. This makes it easier to understand quantitatively the visitors' flows. For instance, the upper part of fig.5

shows the most used trajectory (0-8-3-7-3-4-5-0) while the lower one shows the second one (0-8-3-7-3-4-5-6-0) both with thick arrows. Thin arrows express the second or third higher transition rates from a node to the following ones.

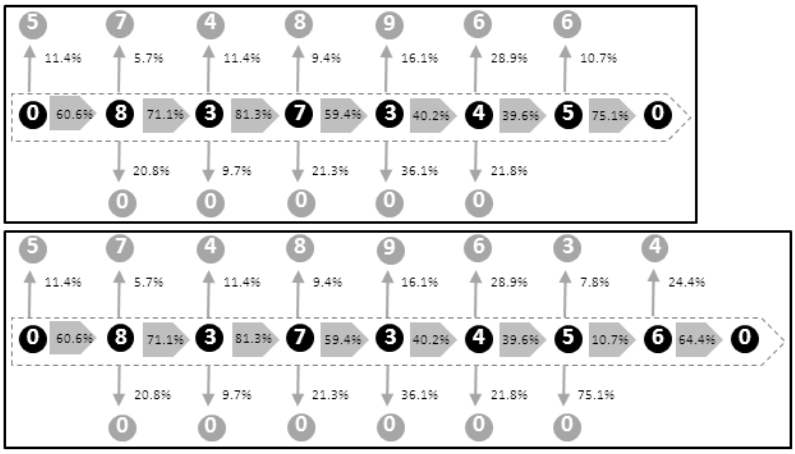


Fig. 4. Diagram of nodes, and percentages of visitors moving between them. Above, the most used trajectory is shown; below, the second most used one.

Various findings from the diagram can be extracted, but the most visible outcome is the strong connection between nodes 8 and 3, and between nodes 3 and 7. While 71.1% of visitors moved to node 3 after visiting node 8, 81.3% of them went to node 7 after visiting node 3. Concerning node 3, the analysis shows that people tend to use it to make a change of direction as it is in the same way from the Pyramid space. In a similar way as the Pyramid space, which distributes people for three accesses, node 3 also serves for distributing visitors to other places.

All of these findings demonstrate that our methodologies can reveal unknown aspects of visitors' spatial use, which observation and traditional interview-based approaches could not clarify at small scale in spatiotemporal terms.

Relationship between length of stay and number of visited nodes. Next, the relationship between the average stay length and the number of visited nodes of each trajectory is analysed, as it can provide another pattern of visitors' behaviour and spatial use in the museum (table 3).

Table 3. Average length of stay corresponding to each trajectory

Trajectory	Average length of stay
0-8-0	2:46:27
0-8-3-0	2:49:11
0-8-3-7-0	2:45:21
0-8-3-7-3-0	2:35:17
0-8-3-7-3-4-0	2:19:58
0-8-3-7-3-4-5-0 (most used trajectory)	2:19:14

The regression line of the number of nodes in a trajectory versus the average staying time has a negative slope, meaning that the larger the number of nodes in the trajectory, the shorter the visit to the museum lasts (with an r2 value of 0.85) – which opposes to the most obvious assumption.

6 Discussion, Conclusions and Future Work

The paper shows, through different analyses, that Bluetooth data can throw new light on spatial use and visitors' behaviours at the building scale. Namely, evidence on the use of the different accesses as the entrance route by visitors and an exit pattern, which is different from the entrance pattern, has been given. This evidence would be difficult and costly to obtain from observations and surveys. Again, these traditional methods would have had difficulties to offer estimates of percentages of visitors that have followed different trajectories, to detect the importance of the 0-8-0 route, or to easily detect, from raw data, the role of node 3 as a crossroads. Another example of the power of the simple analysis on the data is the inverse relation found between length of stay and number of visited nodes.

These initial findings suggest that the methodology proposed has a great potential to clarify the features of the space and its use by visitors in small spatiotemporal scales with unprecedented accuracy. For example, comparing the results of audits of different periods would offer the possibility of obtaining results of seasonal nature; analysing audits of a large amount of data collected, finer detail of patterns and relationships could be obtained – beyond the crude relationship of average visit length and number of nodes. Collecting audits during longer periods requires only very small extra empirical effort besides the one already described; and the analysis would only mean extending and refining the analytical tools.

Before discussing more details of our current work, let discuss some aspects of the data obtained. As seen, the data correspond to a small sample of the visitors – although very large compared to the typical sample used in surveys, and without subjective bias -, but seems to be reasonably valid in terms of the correlation shown. However, more work should be done to clarify the extent to which the sample is representative, as carrying a Bluetooth device set as enabled might be a significant bias. With respect to other data collection strategies involving users carrying specific devices and consequently being aware of them, data appear to be free from potential bias, and the dataset obtained is larger.

Secondly, the data obtained are usually noisy; however, examples of strategies for checking data consistency from the data themselves, and for cleaning it have been given. Larger audits, which would offer larger datasets, can help to strengthen this aspect.

Let turn now to current and future work. Exploring some of the aspects mentioned before based on larger audits of the Louvre has started. And based on the current results, indicators for crowd management and an algorithm for sequential mining are being developed to discover frequent patterns and the underlying association rules. The dynamic estimation of the density and flow of visitors in and between nodes could be associated with the indicator of the relation between pedestrian flow and its

density (Seyfried, Steffen, Klingsch, & Boltes, 2005) for more dynamic crowd management. While several attempts have been made to extract meaningful frequent trajectory patterns and predict further movements of objects at a variety of scales from region and city (Giannotti, Nanni, Pedreschi, & Pinelli, 2007) to retail shop (Larson, Bradlow, & Fader, 2005), improved mining techniques and parameter settings depending on the nature of the data would be needed in order to achieve these goals. Explore similar pedestrian data collected in unconstrained environments has started, and it should help to substantiate the previous statement.

As a final point, one should remark that the understanding of the patterns in visitors' behaviour and its prediction will enable to optimize the spatial layout of objects, human resources and facilities, including advertising and visitor information points, to respond to the increasing tourism demand. It could become a strong management tool not only for museums but also for urban environments in the tourism flourishing age.

References

Agrawal, R. & Srikant, R. (1995). Mining sequential patterns. In *ICDE '95: Proceedings of the Eleventh International Conference on Data Engineering.* Washington DC: IEEE Computer Society.

Ahas, R., Aasa, A., Roose, A, Mark, U. & Silm, S. (2008). Evaluating passive mobile positioning data for tourism surveys: An Estonian case study. *Tourism Management* 29(3): 469-486.

Antonini, G., Bierlaire, M. & Weber, M. (2006). Discrete choice models of pedestrian walking behaviour. *Transportation Research Part B Methodological* 40(8): 667-687.

Asakura, Y. & Iryob, T. (2007). Analysis of tourist behaviour based on the tracking data collected using a mobile communication instrument. *Transportation Research Part A: Policy and Practice* 41(7): 684-690.

Barceló, J., Montero, L., Marqués, L. & Carmona, C. (2010). Travel Time Forecasting and Dynamic Origin-Destination Estimation for Freeways Based on Bluetooth Traffic Monitoring. Transportation Research Record. *Journal of the Transportation Research Board* 2175: 19-27.

Eagle, N. & Pentland, A. (2005) Reality mining: sensing complex social systems. *Personal and Ubiquitous Computing* 10(4): 255-268.

Giannotti, F., Nanni, M., Pedreschi, D. & Pinelli, F. (2007). Trajectory Pattern Mining. In *Proceedings of the 13th ACM SIGKDD international conference on knowledge discovery and data mining KDD 07*, San Jose: 330-339.

Girardin, F., Calabrese, F., Dal Fiore, F., Ratti, C. & Blat, J. (2008). Digital footprinting: Uncovering tourists with user-generated content. *Pervasive Computing, IEEE* 7(4): 36-43.

Girardin, F., Dal Fiore, F., Ratti, C. & Blat, J. (2008). Leveraging explicitly disclosed location information to understand tourist dynamics: a case study. *Journal of Location Based Services* 2(1): 41-56.

Girardin, F., Dillenbourg, P. & Nova, N. (2009). Detecting air travel to survey passengers on a worldwide scale. *Journal of Location Based Services* 3(3): 210-226.

Hillier, B. & Tzortzi, K. (2006). Space Syntax: The Language of Museum Space. In S. MacDonald (Ed.), *A Companion to Museum Studies.* London: Blackwell Publishing.

Hooper-Greenhill, E. (2006). Studying visitors. In S. MacDonald (Ed.), *A Companion to Museum Studies*. London: Blackwell Publishing.

Kanda, T., Shiomi, M., Perrin, L., Nomura,T., Ishiguro, H. & Hagita, N. (2007). Analysis of people trajectories with ubiquitous sensors in a science museum. In *Proceedings 2007 IEEE International Conference on Robotics and Automation (ICRA'07)*: 4846-4853.

Kostakos, V., O'Neill, E., Penn, A., Roussos, G. & Papadongonas, D. (2010). Brief encounters: sensing, modelling and visualizing urban mobility and copresence networks. *ACM Transactions on Computer Human Interaction* 17(1): 1-38.

Larson, J.S., Bradlow, E.T. & Fader, P.S. (2005). An exploratory look at supermarket shopping paths. *International Journal of Research in Marketing* 22: 395-414.

Nicolai, T., Yoneki, E., Behrens, N. & Kenn, H. (2006). Exploring social context with the Wireless Rope. In R. Meersman, Z. Tari and P. Herrero (Eds.), *On the move to meaningful internet systems 2006: LNCS, vol 427*. Heidelberg: Springer.

O'Neill, E., Kostakos, V., Kindberg, T., Fatah gen. Schieck, A., Penn, A., Stanton Fraser,D. & Jones,T. (2006). Instrumenting the city: Developing methods for observing and understanding the digital cityscape. In P. Dourish and A. Friday (Eds.), Ubicomp 2006, LNCS 4206: 315-332.

Paulos, E. & Goodman, E. (2004). The familiar stranger: anxiety, comfort, and play in public places In. *Proceedings of the SIGCHI conference on Human factors in computing systems* 6(1): 223-230.

Pereira, F. C., Vaccari, A., Giardin, F., Chiu, C. & Ratti, C. (2012) Crowdsensing in the web: analysing the citizen experience in the urban space. In M. Foth, L. Forlano, C. Satchell, and M. Gibbs (Eds.), *From Social Butterfly to Engaged Citizen: Urban Informatics, Social Media, Ubiquitous Computing, and Mobile Technology to Support Citizen Engagement*. Cambridge (MA): MIT Press.

Ratti, C., Pulselli, R., Williams, S. & Frenchman, D. (2006). Mobile Landscapes: using location data from cell phones for urban analysis. *Environment and Planning B: Planning and Design* 33(5): 727-748.

Sanfeliu, A., Llácer, M.R., Gramunt, M.D., Punsola, A. & Yoshimura, Y. (2010). Influence of the privacy issue in the Deployment and Design of Networking Robots in European Urban Areas. *Advanced Robotics* 24(13): 1873-1899.

Seyfried, A., Steffen, B., Klingsch, W. & Boltes, M. (2005). The fundamental diagram of pedestrian movement revisited. *Journal of Statistical Mechanics: Theory and Experiment* 2005(10): 10002-10002.

Yalowitz, S. S. & Bronnenkant, Kerry. (2009) Timing and Tracking: Unlocking Visitor Behavior. *Visitor Studies* 12(1): 47-64.

Acknowledgements

We would like to thank the Studies and Research Department of the Louvre Museum as well as BitCarrier for their support. We are indebted to many people at Universitat Pompeu Fabra and Lift Lab for the stimulating research environments and their generous feedback. Of course, any shortcomings are our sole responsibility.

A New Hotel Online Reputation Framework for Ontology Learning Development

Sofia Reino[a],
Aihnoa Serna[b],
Carlos Lamsfus, and
Aurkene Azua-Sorzabal[c]

[a] School of Arts, Social Sciences and Management
Queen Margaret University, UK
sreino@qmu.ac.uk.

[a] Faculty of Engineering
Mondragon Unibertsitatea, Spain
aserna@mondragon.edu

[c] Cooperative Research Centre in Tourism, CICtourGUNE, Spain
{CarlosLamsfus, AurkeneAlzua}@tourgune.org

Abstract

Managing online reviews is fundamental to the marketing of hotels. Reputation Management Systems need strengthening their methodologies to support their aim efficiently. This paper develops a conceptual framework for Hotel Online Reputation and the kernel for Ontology Learning. This is done through a combination of a literature review of Corporate Reputation and the content analysis of Hotel Online Reviews. The results provided an insight into the perception that consumers have of hotels, the way these perceptions are portrayed through online reviews, and the elements and attributes which constitute the concept of Hotel Online Reputation. A detailed explanation of the elements involved in the development of Hotel Online Reputation, and how this framework supports the development of an ontology has also been provided. Implications for those involved in the software development of Reputation Management Systems, managers and scholars have also been introduced through the conclusions sections.

Keywords: Hotels, online reviews, reputation management systems, ontology learning.

1. Introduction

User Generated Content (UGC), and in particular, online reviews have enabled substantial changes to the dynamics of entire industries. Such is the case of travel and tourism, and specifically of the hotel sector. Consumers no longer need to rely on official sources of information only when trying to learn about new products. Thanks to new technologies, and to the increasing number of platforms that support the communication among different stakeholders, consumers can nowadays have access to larger amounts of information about products from different sources. Thus, the American Association of Travel Industry estimates that 67% of the travellers with an Internet connection use the Web to search information related to tourism destinations (TIA, 2005). According to some studies, tourists rely more on this UGC because all comments are done on the base of self-experience (Pan et al., 2007).

This enables consumers to obtain a more comprehensive picture of new products. This is also the case in online hotel reputation. Tourists increasingly check comments before making any decision. The structural changes to the industry above described have opened opportunities to some players, especially to consumers, but they have also entailed challenges to others. This is the case of those establishments or hotel chains which experience difficulties to deal with the online reviews about their businesses. Their lack of ability to respond to negative comments may jeopardise other marketing efforts.

However, managing online reviews entails substantial efforts, and not every single business may have these resources. In order to facilitate this task a number of solutions, also called Reputation Management Systems, have been developed. Their function is to provide businesses with a summary of their online reputation, which would facilitate the task of managing online reviews. Popular examples of Reputation Management Systems for Hotels are ReviewAnalyst, ebuzzconnect and ReviewPro, etc. However, limited information regarding their selected methods is provided, which suggests that a robust methodology may be lacking.

For instance, what is known about their approach is that these systems search through online sources looking for the keywords representing the online reputation of hotels. Ebuzzconnect, in particular claims that their system searches for the 10 keywords which are most commonly used in online reviews. These keywords are then classified these into positive and negative attributes. These 10 terms are all self-explanatory, such as “nice hotel”, “excellent”, “highly recommended”, “noise”, etc. However, no explanation of the process by which only self-explanatory terms have been selected is provided. Additionally, many relevant topics related to the hotel online reputation could also be missing from these reviews.

And finally, no explanation of the process is provided by which key terms are classified into positive and negative categories. For example customers could be commenting on how noisy the hotel is or they could be suggesting that there is complete absence of noise, in which case it may be a positive attribute. An effort to develop a robust methodology for Online Reputation Systems in the context of the travel and tourism industry was undertaken by Marchiori et al. (2010). However, their framework focuses on tourism destinations instead of specifically on the reputation of the hotel sector. Although there could be some overlap, important differences could also be expected.

In order to support Reputation Management Systems overcoming their current methodological limitations, this paper develops a new conceptual framework for Hotel Online Reputation. Additionally, it aims to provide guidelines for the analysis of online hotel reputation, the organisation of data, and the generation of an ontology kernel. This process will subsequently support the semiautomatic semantic-based content analysis of unstructured texts, through the application of techniques such as Information Retrieval and Natural Language Processing (Wing Li & Law, 2007; Alexakos et al., 2009; and Xiao Wu et al., 2008). The resulting framework is based on a combination of a literature review of Corporate Reputation, together with the results from the content analysis of Hotel Online Reviews.

2. Literature Review

Most of the work about online reviews in the area of eTourism seems to explore the impact of online reviews on consumer decision-making (Gretzel & Hyan Yoo, 2008; Vermeulen & Seegers, 2009), motivations to write reviews (Gretzel, Hyan-Yoo, & Purifoy, 2007), and their social dynamics (Gretzel & Hyan-Yoo, 2008; O´Connor, 2008). Also within the context of eTourism, the work by Marchiori et al. (2010) developed a framework to classify UGC for a given destination. The framework named DORM (Destination Online Reputation Model) is a combination of the Reputation Quotient and the RepTrak models developed by the Reputation Institute (Vidaver-Cohen, 2007). However, their work focuses on the concept of online reputation for tourism destinations and not for hotels. And as suggested above lines, there is potential for some overlap, but substantial differences could also be expected, related to the different attributes which may be used to describe tourism destinations and hotels.

Reputation is defined as a "collective assessment of a company's ability to provide valued outcomes to a representative group of stakeholders" (Fombrun, Gardberg, & Sever, 1999). The concept of reputation is closely related to that one of image, which was defined by Kotler et al. (1994) in the context of tourism destinations as "the sum of believes, ideas, and impressions that a person holds of it". An additional concept which is closely related to that one of reputation is "brand equity", suggested by Aaker (1996). This refers to "the set of human characteristics associated with a brand", and with the credibility of a firm (Nguyen & Leblanc, 2001). This author identifies 5 personality traits including sincerity, excitement, competence, sophistication and ruggedness, suggesting that different personality traits influence on the preference for different brand attributes.

According to Nguyen and Leblanc (2001) reputation has attracted the attention of scholars from different disciplines. These include economist, organisational researchers and marketing scholars. The former group has been primarily interested in understanding the relationship between reputation and quality or price; from the organisational researchers´ perspective, the focus has also been placed on understanding the role of reputation in contributing to the performance of the company; finally, from the perspective of marketing, which is the one of interest to this paper, research has aimed to identify the essential elements of reputation and the process by which this is formed.

Reputation Quotient and the RepTrak models developed by the Reputation Institute (Vidaver-Cohen, 2007) define trustful organisations in relation to the *attributes* that the different *stakeholders* consider *positive* or *negative*. They relate to a set of specific criteria which stakeholders may expect from a well-behaved organisation, i.e. those whose behaviours comply with what is considered as socially acceptable. Herbig and Milewicz (1993) consider that reputation is "an estimation of the consistency *over time* of an attribute of an entity", which suggests the important role played by time in the consolidation of certain opinions. Britt (1971) identified what he called the "corporate imagery mix", a concept which aimed to highlight potential differences across channels in their representation of a product's image or reputation.

Examples of different channels include national origin (of corporation and products), product appearance, container, name (of company and brands), graphics, and the place where the products are sold, employees and advertising and promotion.

Figure 1 integrates these elements into a framework of Corporate Online Reputation. This includes the stakeholders who have multidimensional attributes, positive or negative about specific subjects (Vidaver-Cohen, 2007), e.g. hotels. In turn, these take place in a specific context, (Britt, 1997) e.g. online, and are formed over time (Herbig & Milewicz, 1993).

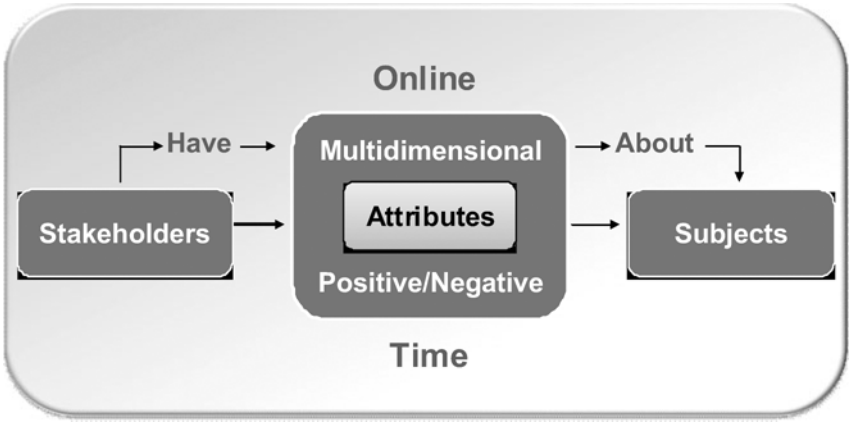


Fig. 1. Framework of Corporate Online Reputation

The remaining sections of the paper will focus on the identification of the different *attributes* which are part of this framework in the context of Hotel Online Reputation.

3. Methodology

A qualitative approach based upon a Grounded Theory methodology was adopted to support the identification of the multidimensional opinions of Hotel Online Reputation. Grounded Theory relates to a qualitative methodology which supports the development of theory directly emerging from the data, through a systematic process of tracking, checking and validation (Goulding, 2002).

The selection of this methodological approach responds to the nature of the research, which aims to develop conceptual constructs, i.e. theory, directly emerging from the data (Guba & Lincoln, 2005). Other qualitative approaches were considered, but Grounded Theory was selected because this supports the development of theory directly emerging from the data, while others are mainly descriptive, such as phenomenology and ethnography, or focus only on limited number of cases, such as case study (Goulding, 2002).

Content analysis of 274 hotel online reviews from 15 hotels in San Sebatian (Spain) was undertaken. Content Analysis is a technique commonly used to support Grounded Theory (Suddaby, 2006). It consists of making inferences from text data; each word or phrase in a text is categorised by applying labels that reflect concepts (Goulding,

2002). Counting, which may or may not be done in content analysis was not adopted for this research. The reason for this relates to the aim of the work which claimed for inclusivity of all possible terms rather than selection. In order to select the hotels to be included in the study, a search with the keywords "Hoteles in San Sebastián-Donostia" was undertaken in TripAdvisor.com. The search produced forty one establishments with reviews. In order to obtain a representation of the different types of reviews, the first 5 resulting hotels of the list, five from the middle (posts 18 to 22) and the last five ones were selected for the analysis, as shown in Table 1.

Table 1. Number of Online Reviews per Hotel and Rank of Listing

Top Listed		Medium Listed		Bottom Listed	
Hotel Name	**Reviews**	**Hotel Name**	**Reviews**	**Hotel Name**	**Reviews**
Villa Soro	30	Arrizul Gros	12	Pension Kaia	4
Bellas Artes	30	Ezeiza	11	Terminus	6
Astoria	30	Avenida	30	Residencia La Estrella	1
Maria Cristina	30	Bahía	23	Anoeta	4
Silken Amara	30	Edorta	30	Monte Ulia	3

Some hotels produced a large number of reviews. Therefore, the 30 first online reviews of hotels where selected for the analysis. Data collection took place in July and August 2011, and theme coding was the strategy adopted for the analysis.

As mentioned at the beginning of the paper, the analysis of hotel reputation was undertaken by means of ontologies. Ontologies (Gruber, 1993) have traditionally been used for various purposes, including information analysis (Corcho et al., 2007). They aim to represent domains of knowledge (Guarino, 1995), which can be populated by extracting relevant instances from information sources, creating thus knowledge bases.

This process is known as ontology population or markup. Ontology learning from text is then essentially the process of deriving high-level concepts and relations as well as the occasional axioms from information to form an ontology (Wong et al., 2010).

In the case of the research presented in this paper, the new online hotel reputation framework originated an ontology that has to be populated with instances. The general structure of the ontology and its different categories have been created with the concepts extracted from the 274 hotel online reviews. The instances are taken from the unstructured texts in the different information sources (in this case, TripAdvisor) dealing with online hotel reputation. Each of these instances is then automatically processed. By means of this processing, the system determines the class

of the ontology onto which the instance has to be loaded, thus creating the knowledge bases that will provide the means to measure reputation. The process of building the ontology in this case has been top down. After analysing the 250 hotels online, the most common concepts have been identified and grouped in classes of concepts, as shown in Figure 2. A core ontology is then built based on these concepts.

4. Results

Figure 2 shows the core concepts extracted from the 274 hotel online reviews. This includes both *Physical* (those included in the grey area of Figure 2) and *Abstract* elements (those in the green area of Figure 2).

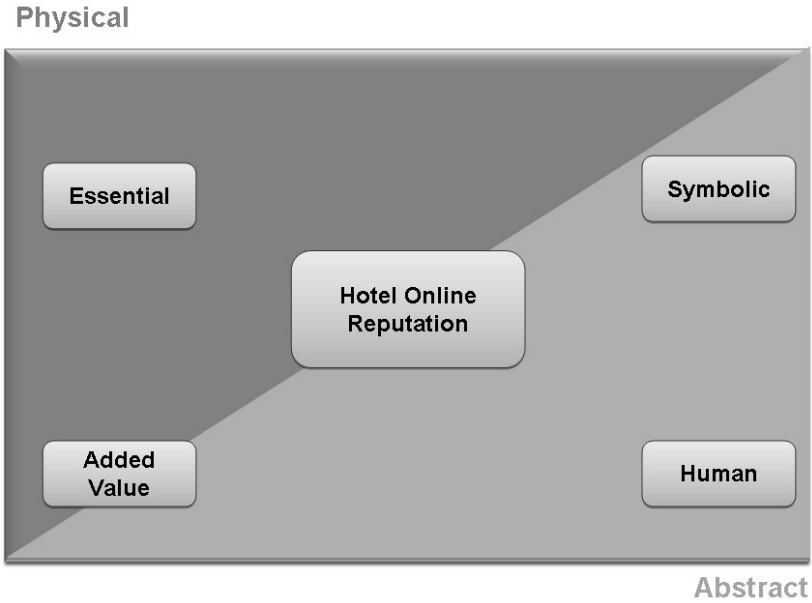


Fig. 2. Hotel Online Reputation Core

Elements which are *Essential* to the concept of hotel have all been classified as *Physical*, and those which are *Symbolic* or of a *Human* nature have all fallen within the category of *Abstract*. However, *Added Value* elements include examples of both *Physical* and *Abstract* concepts.

Going into detail on the category of *Essential* elements shown in Figure 2, this can be decomposed into classes (i.e. *Maintenance*, *Location*, *Room*) and attributes (i.e. *Design*, *Sound*, *Illumination*).

Examples of *Added Value* and *Physical* elements relate to "Free WiFi" and car park facilities", while examples of *Abstract* elements relate to "information about the destination". This latter subsequently includes "maps and itineraries". Furthermore, the concept *Details*, which is related to *Physical* elements, can also subdivided into

different categories, including “flowers”, “chocolates”, “sweets”, “bottle of champaign” and “hotel amenities”.

Symbolic concepts are all *Abstract* and include terms which make reference to the feelings associated with the establishment, like “travelling in time, to Belle Epoque”, being “like at home” or at a “magic B&B”, staying at an establishment “with class”, “luxurious”, “elegant”, “romantic”, “with contemporary comfort”, or which suggest that the establishment is a “point of reference in San Sebastian”, and the “flavour of the past”.

Human elements include concepts like the *Host–Guest Relationship*, *Service* and *Logistics*. Concepts related to the *Host-Guest Relationship* include references to “Smiles”, “Hugs”, “Kisses”, “Exceptional Attention”, “Willingness to offer you everything”, “Care”, “Pampering”, “Assistance” and “Special Treatment”. *Service* is defined in the sense of “Impeccable”, “Amusing”, “Bad”, “Very good”, “Very kind”, “Dreadful”, “Cold”, “Dry”, “Charming” and “Polite”. And finally, references to *Logistics* are made through reference to “Fake Deals” and “Payment Options”.

Furthermore, it should be noted that meta-attributes, such as “I wished that all the hotels which are supposed to be better than this one would treat you as well as this one”, “You´ll never find a better one” are all introduced in a comparative format. This suggests that the influence of previous experience in forming the *expectations* of consumers plays an important role in the generation of hotel online reputation.

Figure 3 combines together the Framework of Corporate Online Reputation (Figure 1), and the Hotel Online Reputation Core (Figure 2), complemented with further insights arising from the data. These altogether generate the Conceptual Framework of Online Hotel Reputation, which was the main aim of this paper.

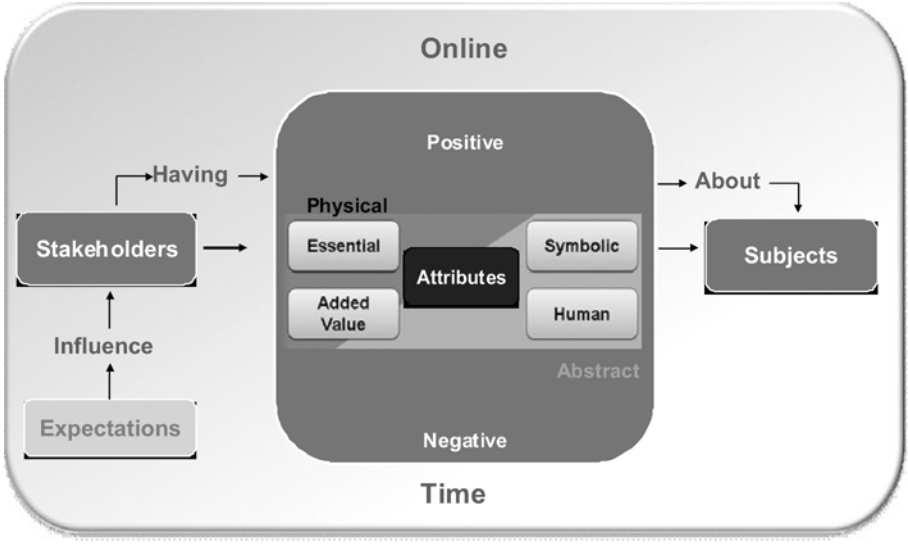


Fig. 3. Conceptual Framework of Online Hotel Reputation

As previously shown in Figure 1, it includes the stakeholders, the multidimensionality and positive/negative dichotomy of opinions, and the subjects upon which opinions are displayed (Vidaver-Cohen, 2007), i.e. hotels; the channel of communication (Britt, 1997), i.e. online; and the element of time (Herbig & Milewicz, 1993). But it also captures the *Physical*/*Abstract* dichotomy and the four categories which were captured through the data analysis stage, namely *Essential*, *Added Value*, *Symbolic* and *Human*. Furthermore, it also captures the suggestions from the data that the *expectations of* stakeholders also play an essential role in determining hotel online reputation.

Additional findings need further consideration towards developing an ontology of Hotel Online Reputation. With regards to the *Positive*/*Negative* dichotomy, it should be noted that normally, the *Positive* or *Negative* connotation of attributes are easily differentiated. However, some exceptions have been noted through the analysis. Some attributes were given different connotations in separate reviews. This is particularly the case of "Old design" and "Old furniture" which was occasionally regarded as a *Positive* element with symbolic meanings such as "travelling back on time…" or providing the hotel with a "romantic" characteristic.

On the other hand, this same attribute is regarded as a *Negative* attribute when suggesting that the hotel "needed refurbishment". Another example where the difference between *Positive* and *Negative* connotations was blurred related to the element of *Location*. For example a "15-20 minutes walk" has been considered to place the hotel "far away" in some reviews, and "conveniently located, only a short 15-20 minutes walk away". These results present important implications, both theoretical and methodological.

The theoretical implications relate to the concept of brand equity suggested by Aaker (1996) within the context of brand reputation. The different perception of an attribute, as positive and negative, suggests that the concept of brand equity suggested by Aaker (1996) is also applicable to the context of Online Hotel Reputation, and therefore, the characteristics of users can be expected to have a direct influence on their positive or negative perception of an attribute. This suggests the need to learn about users and the type of influence that their personality traits have the type of characteristics that they consider of relevance when selecting and evaluating a hotel.

With regards to the methodological implications, these relate to the suggested approach for the development of ontology learning. The ambiguity of these terms reinforces the suggestion that human intervention may always be required in the evaluation of online reputation and a completely automatic system may not suffice. Therefore, a semi-automatic process for Ontology Learning is here suggested and its architecture is represented in Figure 4.

The development of the suggested architecture would involve the following steps: an initial manual process in which the analyzed data is represented by a small ontology (the core of the final ontology to the pre-processing of the text), XML tagged text, information extraction, use of the WordNet (Fellbaum, 2005) lexical database, and the discovery of new relationships and labels. The terms and concepts extracted from the plain text during the data analysis of this work, complemented with details of triplets and axioms, will provide the baseline for the core ontology.

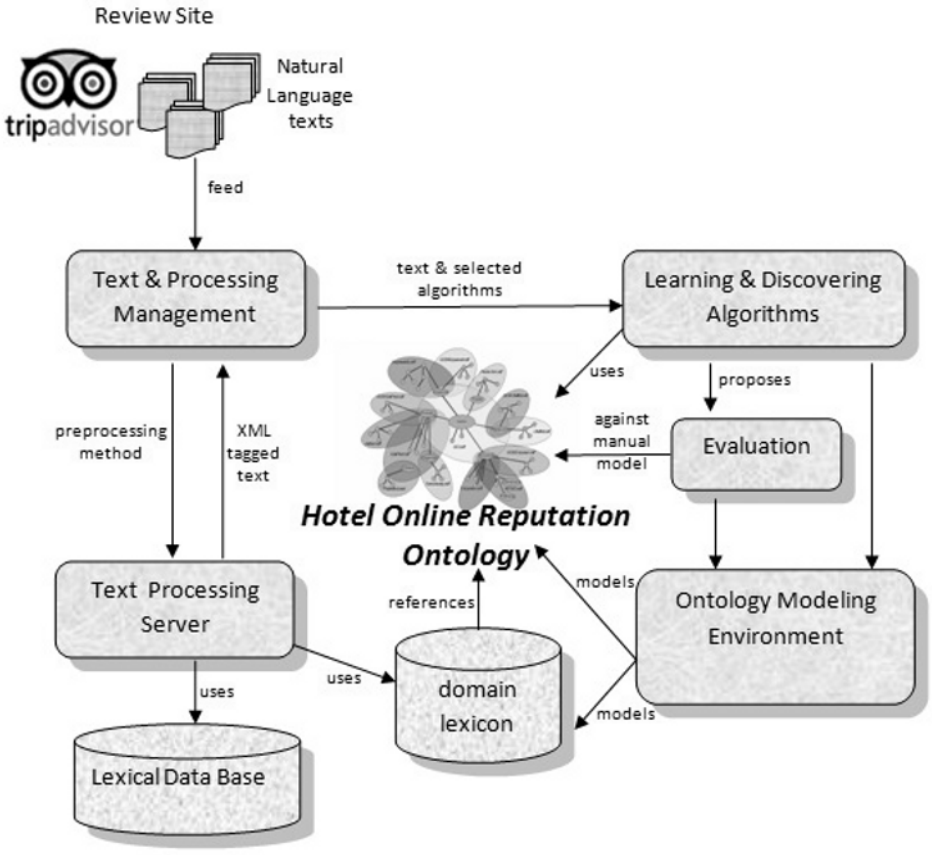


Fig. 4. Semi-automatic Ontology Learning Architecture for Hotel Online Reputation System

The semi-automatic Ontology Learning process is integrated into an architecture that has several key features described as a kind of pipeline in the following general scheme. As the knowledge model matures during the semi-automatic learning process, the conceptual framework represented in the Figure 3 is validated with new data through a feedbacking process.

5. Conclusion

The need to strengthen the methodologies of Reputation Management Systems for Hotels has urged the need to develop a conceptual framework. Ontology learning can help us discovering and labelling relationships between concepts, but it requires an initial framework. This paper aimed to undertake the task of developing this framework through a combination of a literature review and primary data collection, i.e. content analysis of hotel online reviews.

The literature review of Corporate Reputation supported the identification of the key elements of corporate reputation, as suggested by scholars, and contextualised this within the scope of online communication (Figure 1). Following a Grounded Theory approach, the process of content analysing hotel online reviews provided an insight into the way travellers portray hotels online and the elements that support the development of hotel online reputation through UGC. Additionally, it enabled the identification of the different dimensions which are intrinsic to the concept of hotel online reputation. These can be categorised through the dichotomy *Positive/Negative*, *Physical/Abstract*, and four additional categories labelled as *Essential, Added Value, Symbolic* and *Human* elements. All these together provided the Hotel Online Reputation Core (Figure 2), and subsequently, the Conceptual Framework for Hotel Online Reputation (Figure 3).

The results have also suggested some warning sings with regards to the use of completely automatic systems as the baseline methodology for Reputation Management Systems. This is due to the ambiguous connotations which were found for some terms, and it suggests that semi-automatic learning ontologies are more suitable approaches for generating ontology learning than those which are completely automated systems. An illustration of the process by which the conceptual framework supports the development of a semi-automatic Ontology Learning Architecture for Hotel Online Reputation Systems (Figure 4) has also been provided.

Finally, the work presents some limitations which should be addressed through further research. The first one relates to the application of these principles for the development of an ontology and to pilot this. This process will support testing the validity of the framework. The second limitation relates to the geographical scope of the data collection. The data from which the results have been generated correspond only to hotel online reviews of hotels based in San Sebastian.

It is suggested that further research should validate the framework geographically by extending the data collection to hotels from additional destinations worldwide. Finally, the influence of consumers´ personality traits in the importance that they give to different attributes should also be explored in additional studies.

References

Aaker, D. (1996). *Building Strong Brands*. New York: The Free Press.

Alexakos, C., Kalogeras, A., Konstantinopoulos, P., Charatsis, K., Koubias, S. & Tomara, V. (2009). Semantically-supported information retrieval in inter-enterprise cluster for the tourist sector. *International Journal of Digital Culture and Electronic Tourism* 1(4): 325 – 345.

Britt, S.H. (1971). The right marketing mix for the corporate imagery mix. Role of various publics and marketing basics. *Business Horizons* 14 (February): 87-94.

Corcho, O., Fernández-López, M. & Gómez-Pérez, A. (2007). Ontological Engineering: What are Ontologies and How can We build them?. In J. Cardoso (Ed.), *Semantic Web Services: Theory, Tools and Applications*. Idea Group Inc.

Crompton, J.L. (1979). An Assessment of the Image of Mexico as a Vacation Destination and the Influence of Geographical Location upon the Image. *Journal of Travel Research* 18(4): 18–23.

Fellbaum, C. (2005). WordNet and wordnets. In K. Brown (eds.), *Encyclopedia of Language and Linguistics* (2nd ed.) Oxford: Elsevier.

Fombrun, C.J., Gardberg, N. & Sever, J.M. (1999). The Reputation of Quotientsm: A multistakeholder measure of corporate reputation. *The Journal of Brand Management* 7(4): 241-255.

Goulding, C. (2002). *Grounded theory: a practical guide for management, business and market researchers*. London: SAGE Publications Ltd.

Gretzel, U. & Hyann-Yoo, K. (2008). Use and Impact of Online Travel Reviews. In P. O´Connor, W. Höpken and U. Gretzel (Eds.), *Information and Communication Technologies in Tourism* 2008. Proceedings of the International Conference. Springer: Vienna.

Gretzel, U., Hyan-Yoo, K. & Purifoy, M. (2007). Online Travel Review Study: Role & Impact of Online Travel Reviews. Retrieved August 10, 2011, from http://www.tripadvisor.com/pdfs/OnlineTravelReviewReport.pdf.

Gruber, T. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition* 5: 199-220.

Guarino, N. (1995). Formal ontology, conceptual analysis and knowledge representation. *International Journal of Human and Computer Studies* 43: 625-640.

Guba, E. G. & Lincoln, Y. S. (2005). Paradigmatic controversies, contradictions, and emerging influences. In N. K. Denzin and Y. S. Lincoln (Eds.), *The Sage Handbook of Qualitative Research* (3rd ed.), pp. 191-215. Thousand Oaks, CA: Sage.

Herbig, P. & Milewicz, J., 1993. The relationship of reputation and credibility to brand success. *Journal of Consumer Marketing* 10(1): 5-10.

Hjelm, H. & Volk, M. (2011). Cross-language ontology learning. In W. Wong, W. Liu and M. Bennamoun (Eds.), *Ontology Learning and Knowledge Discovery Using the Web: Challenges and Recent Advances*. NY: IGI Global.

Marchiori, E., Inversini, A. & Cantoni, L. (2010). Me*asuring the Online Reputation of Sustainable Tourism Destinations*. Presented at 14th International Conference on Corporate Reputation, Brand, Identity and Competitiveness, Rio de Janeiro – Brazil.

Nguyen, N. & Leblanc, G. (2001). Corporate image and corporate reputation in customers´retention decisions in services. *Journal of Retailing and Consumer Services* 8: 227-236.

O´Connor, P. (2008). User-Generated Content and Travel: A Case Study on Tripadvisor.com. In P. O´Connor, W. Höpken and U. Gretzel (Eds.), *Information and Communication Technologies in Tourism 2008*. Proceedings of the International Conference. Vienna: Springer.

OWL DL (2004). Retrieved from http://www.w3.org/TR/owl-ref/#OWLDL

Pan, B., MacLaurin, T. & Crotts, C.J. (2007). Travel blogs and the implications for destination marketing. *Journal of Travel Research*. 46: 35-45.

Shamsfard, M. & Barforoush, A.A. (2004). Learning ontologies from natural language texts. *Int. J. Hum.-Comput. Stud*. 60(1): 17-63.

Suddaby, R. (2006). From The Editors: What Grounded Theory is Not. *Academy of Management Journal* 49(4): 633-642.

TIA. (2005). *Executive Summaries - Travelers' Use of the Internet*, 2004 Edition. Retrieved November 15, 2005.

Vermeulen, I.E. & Seegers, D. (2009). Tried and tested: The impact of online hotel reviews on consumer consideration. *Tourism Management* 30: 123-127.

Vidaver-Cohen, D. (2007). Reputation Beyond the Rankings: A Conceptual Framework for Business School Research. *Corporate Reputation Review* 10: 278-304.

Wing Li, K. & Law, R. (2007). A novel English/Chinese information retrieval approach in hotel website searching. *Tourism Management Journal* 28(3): 651-950.

Wong, W. (2009). *Learning Lightweight Ontologies from Text across Different Domains using the Web as Background Knowledge.* Doctor of Philosophy thesis, University of Western Australia.

Wong, W., Liu, W. & Bennamoun, M. (In Press).Ontology Learning from Text: A Look back and into the Future. To appear in *ACM Computing Surveys*.

Wong, W., Liu, W. & Bennamoun, M. (In Press). Ontology Learning from Text: A Look back and into the Future. *ACM Computing Surveys*.

Wu, X., Li, J., Zhang, Y., Tang, S. & Neo, S. (2008). Personalized multimedia web summarizer for tourist. In *Proceeding of the 17th international conference on World Wide Web*, April 21-25, 2008, Beijing, China.

When and Where Tourists are Viewing Exhibitions: Toward Sophistication of GPS-Assisted Tourist Activity Surveys

Junya Kawase,
Yohei Kurata, and
Naoto Yabe

Department Tourism Science
Tokyo Metropolitan University, Japan
j.kawase0922@gmail.com
{ykurata, yabebeya}@tmu.ac.jp

Abstract

In recent years, many tourist activity surveys using GPS devices have been conducted, but the survey methodology needs still to be improved. One critical problem is that we can know where tourists visit and how long they stay there, but not what they actually do there. Thus, we investigated the relationship between tourists' activities recoded by a video camera and their GPS logs. The result shows that the location history of a tourist itself is not sufficient for estimating how long he/she enjoys each attraction, and his/her walking speed seems critical for this estimation.

Keywords: GPS, tourist activity survey, zoological park, logistic regression model

1 Introduction

Global Positioning System (GPS) provides accurate, continuous, worldwide, three-dimensional position and velocity information to users with appropriate receiving equipment (Kaplan, 1996). GPS expands the possibility of tourist activity surveys, because with an effective use of GPS devices we can investigate where tourists visit and how long they stay there, without recording their activities manually. As a matter of course, there are also some drawbacks; first of all, GPS receivers work properly only in outdoor environments. Measurement error and battery lifetime are other issues that need to be considered. In tourist activity surveys, however, there remains one more critical issue—GPS logs tell where tourists have been to, but not what they have been doing there. Tourism is an activity that involves a large variety of activities. Thus, for discussing marketing strategies, renovating tourist spaces, and providing appropriate information for tourists, it is essential to clarify tourists' behaviours in conjunction with their attributes, such as age and gender. For promoting tourist activity surveys to know tourists' behaviours, an intelligent technique for inferring tourist activities from their GPS logs will be highly desirable.

As a first step toward this technique, we investigated the relationship between the tourists' activities and their GPS logs, focusing on the simplest activity of tourism—*viewing something*. Then, we developed a statistical model with which we can discern from a GPS log whether a tourist had been viewing exhibitions or not at each moment. This model can be applied to GPS logs in previous surveys to estimate the time spent by tourists for viewing each exhibition and eventually, to estimate the attractiveness of each exhibition.

The remainder of this paper is organized as follows: Section 2 explains the background of this research. Section 3 reports our initial GPS-assisted survey at a zoological park and points out its problems. Section 4 describes our supplementary experiment that aims at developing a new technique for estimating tourists' viewing time. Finally, section 5 concludes with a discussion of future work.

2 Background

Recently, GPS loggers have become smaller, lighter, and less expensive. In addition, mobile communication devices equipped with a GPS sensor, such as smartphones, have become rapidly widespread all over the world. Along with this trend, activity surveys using GPS devices have been conducted in various fields (Yabe et al., 2010). Especially in tourism studies, location history data recorded by GPS devices are highly useful to clarify what attractions tourists have visited in what order and how long they have stayed there (Kurata et al., 2010).

If tourists' spatio-temporal behaviours are well-understood, it would be possible to optimize transportation, operations of tourist attractions, and marketing strategies, all in line with their actual needs (Shoval & Issacson, 2007a). Of course, such information can be obtained from a questionnaire or an observation survey, but usually such surveys are costly and difficult to be conducted regularly for monitoring tourists' behaviours. On the other hand, GPS-assisted surveys impose almost no burden on tourists and are relatively inexpensive if GPS devices are used repeatedly (Kurata et al., 2010).

GPS-assisted activity surveys have been conducted extensively in transportation studies of motor vehicles. One of its reasons is that it is easy to install GPS devices on vehicles (Shoval & Issacson, 2007a). For example, Nagao et al. (2004) analyzed tourists' movement patterns in a macro scale based on the data of GPS loggers installed on rental vehicles. Some GPS-assisted tourist activity surveys targeted pedestrians to clarify their movement patterns and characteristics (Asakura & Hato, 2004; Asakura & Iryo, 2007; Shoval & Issacson, 2007b), as well as the spatial use of a city by tourists (Shoval, 2008). Some surveys attempt to extract interesting locations for tourists from their GPS logs (Zheng et al., 2011). However, it still remains as a research challenge to develop a technique for inferring tourists' activities from their GPS logs, while it is important to know what tourists actually do at each place in order to optimize various operations in a tourist area.

3 GPS-Assisted Survey in Tama Zoological Park

As an example of GPS-assisted tourist activity surveys, we report our survey at Tama Zoological Park (Hino, Tokyo) on 3th (Fri) and 4th (Sat) of September, 2010. There are three advantages in conducting a GPS-assisted survey in a zoological park:

(i) most exhibitions are located outside

(ii) the number of its entrances and exits, where GPS loggers should be distributed and collected, are limited, and

(iii) the time spent by each visitor does not exceed the battery lifetime of GPS loggers. In addition, in Tama Zoological Park, animal exhibitions are located separately and thus, we can expect clear difference between visitors' viewing and walking states.

Fig. 1 shows a map of Tama Zoological Park. The plan is complicated due to its hilly location. This is why its staff has wanted to know the visitors' movement patterns in the park.

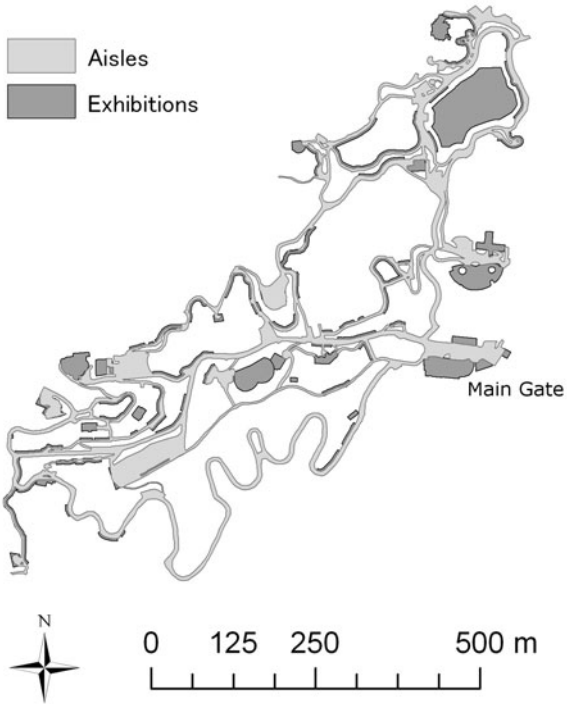


Fig. 1. A map of Tama Zoological Park

Our survey was conducted as follows: First, we stood at the main gate and distributed GPS loggers to the visitors who agreed to participate in our survey. When the visitors came back to the same place to leave the park, we collected the loggers from them and asked them to answer a questionnaire to obtain their attribute data, such as age, gender, and accompanying persons.

The GPS loggers we used were Qstarz Black Gold 1300 (Fig.2), whose weight was only 22g. We asked the visitors to put the GPS loggers on their necks, bags, baby carriages, or wherever they like. Each GPS logger was set to record the geo-coordinates of its location once in every second.

Fig. 2. A GPS logger used for our survey (Qstarz Black Gold 1300)

We obtained 190 valid sets of GPS logs and visitors' attribute data. The GPS logs were refined by removing error points. In this study 'error points' were defined as those that refer to a location more than three meters away from park aisles, because the possible horizontal error of our GPS loggers was three meters (catalogue value).

From the refined data, we first calculated each visitor's *staying time* at each animal exhibition (i.e., how long each visitor had spent there) by counting the number of points located within three meters from the boundary line between the exhibition and the aisle (Fig. 3) in ArcGIS 9.1. This is because we assumed that the visitors viewed animals just in front of each exhibition and the loggers' possible measurement error was three meters.

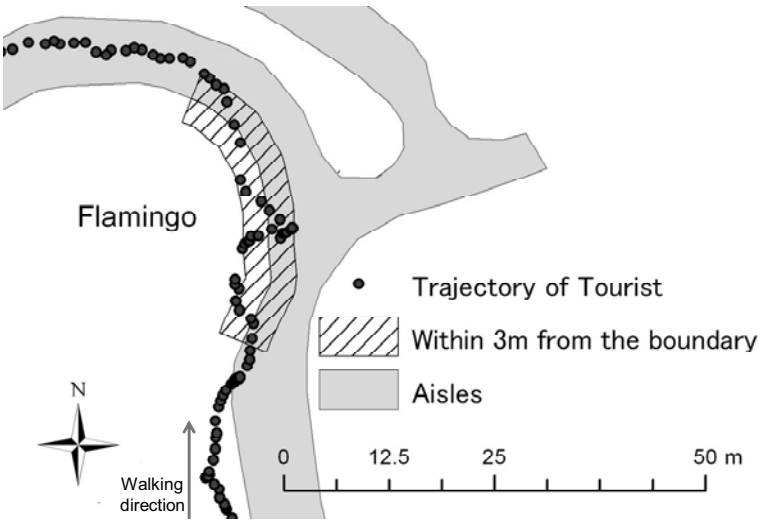


Fig. 3. A sample GPS log, together with an area within three meters from the boundary between Flamingo exhibition and an aisle

Table 1 shows ten top animal exhibitions in terms of average staying time. In addition, Table 2 shows the rankings of animal exhibitions by generations. The exhibition where all generations of people spend the longest time was Insect Museum. Children, especially boys, were caught by this exhibition (Table 2). Insect Museum has a long fixed route in it and accordingly, visitors who went into this museum naturally have spent long time there.

In general, most of Asian animals were not viewed for long time (Table 1). One of zoo staff commented that many visitors wanted to see African and Australian animals, while Asian animals were commonplace for them.

Table 1. Ten top animal exhibitions in terms of average staying time

Rank	Animal Exhibition	Staying time
1	Insect Museum	11’23”
2	Lions	8’06”
3	Giraffes & Zebras	6’15”
4	Japanese Monkeys	4’30”
5	African Elephants	4’08”
6	Chimpanzees	3’49”
7	Walk-in Bird Cage	3’45”
8	Siberian Tigers & European Grey Wolves	3’03”
9	Orang-utans	2’30”
10	Asian Elephants	2’23”

Our result shows that the animal exhibitions people visited, as well as the time they spent there, were quite different by their age, gender and accompanying persons. For instance, we found that people in their 40s spent more time in front of Japanese monkeys and Chimpanzees than those in their 20s, 30s, and 50s (Table 2). This is probably because they were typically interested in monkeys’ sociality. Data about which ages of people are interested in each exhibition is suggestive for providing appropriate information through guide boards at the exhibition.

Another unique finding was that families with small children actually stayed longer time in cafeterias than in animal exhibitions.

We also calculated the percentage of visitors who visited each animal exhibition. Here we assumed that a visitor had ‘visited’ an exhibition if he/she spent more than five seconds there.

Table 2. Ten top animal exhibitions in terms of average staying times by generations

Rank	10s	20s	30s	40s	50s
1	Insect Museum 25’49”	Insect Museum 8’45"	Insect Museum 10’16”	Insect Museum 12’21”	Insect Museum 7’23”
2	Japanese Monkeys 20’36”	Giraffes & Zebras 5’35”	Lions 8’16”	Lions 7’56”	Wolves & Tigers 7’13”
3	Giraffes & Zebras 6’20”	Lion 4’53”	Giraffes & Zebras 7’10”	Japanese Monkeys 6’03”	Giraffes & Zebras 6’54”
4	Brown Bears 5’18”	Chimpanzees 3’06”	Walk-in Bird Cage 4’56”	Giraffes & Zebras 5’49”	Lions 4’03”
5	Lions 5”11’	Walk-in Bird Cage 2”47’	Chimpanzees 4’23”	Wolves & Tigers 4’47”	Japanese Monkeys 3’53”

Table 3 shows the ten top highly-visited exhibitions. The result indicates that animal exhibitions along major corridors and those with large cages were visited by a larger number of people. However, we also found that higher percentage of visits did not necessarily mean longer average staying time (Table 1).

Table 3. Ten top highly-visited animal exhibitions

Rank	Animal exhibition	Percentage
1	Giraffes & Zebras	88%
2	Lions	87%
3	Oriental Storks	73%
4	African Elephants	69%
5	Great Indian Rhinoceroses	68%
6	Siberian Tigers & European Grey Wolves	66%
7	Raptors	66%
8	Orang-utans	63%
9	Parma Wallabies & Wombats	62%
10	Reindeers & Emus	62%

In this way, we successfully clarified the spatio-temporal behaviour of park visitors based on a GPS-assisted activity survey. We, however, considered only ‘staying time’, but not ‘viewing time’. Viewing time is a direct measure for evaluating the attractiveness of each exhibition. On the other hand, staying time is easily affected by other factors. For instance, one of zoo staff pointed out the possibility that visitors had

spent time resting on a bench without viewing an exhibition nearby (e.g., chimpanzees), and also the possibility that the staying time becomes inevitably longer if the exhibition has a long aisle along it (e.g., giraffes). Our assumption that people viewed animals just in front of exhibitions was also questionable, because large animals can be viewed from more than three meters away, while small animals cannot be seen even in front of the exhibition when they are hiding. Thus, in order to judge visitors' viewing states from their GPS logs, we decided to look into additional available data in GPS logs; that is, walking speed.

4 Relationship between Tourist's State and Walking Speed

We conducted a supplementary experiment to investigate the relationship between people's activity (whether viewing or not) and his/her walking speed. This experiment was conducted from 15th of May (Sun) to 28th of June (Tue), 2011 in the same Tama Zoological Park. At this moment we employed five undergraduate and graduate students (three males and two females) as subjects. All of them are in the early 20s.

In this experiment, the subjects were asked to go around a zoological park, carrying a GPS logger. In addition, our staff followed each subject and filmed the subject's activity by a video camera. From this video we made the subject's activity log, which recorded precisely when the subject had been viewing an animal exhibition. The GPS log and activity log of each tourist were then matched. Note that the GPS logs were refined by removing the points recorded when the subject had been in a bus or an indoor exhibition. Fig. 4 shows a sample GPS log with the subject's activity record. It seems that the subject has walked into the cage of Flamingo, but this is actually not (recall that our GPS logger has a measurement error). The figure also shows that the subject started viewing an animal exhibition even from more than three meters away when he is approaching the exhibition. Considering that the possible horizontal error of our GPS loggers is three meters, this indicates that our previous assumption that people view an animal exhibition just in front of it seems not appropriate.

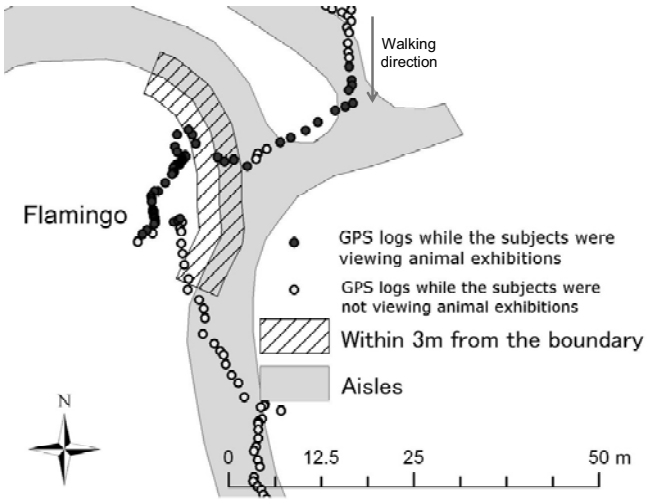


Fig. 4. A sample GPS log together with the subject's activity record

We then calculated the *probability of viewing* (i.e., the probability that people view an animal exhibition) in relation to walking speed, as the relative frequency of viewing states at each speed (every 0.1km/h). The resulting graph (Fig. 5) clearly shows a linear inverse relation between walking speed and the probability of viewing; that is, the probability of viewing increases as people walk more slowly, but even while not moving he is not always viewing an exhibition.

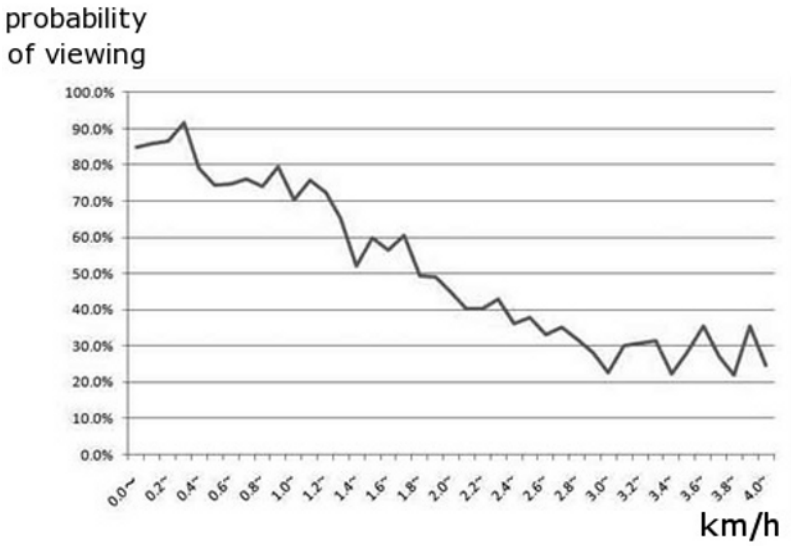


Fig. 5. Relation between probability of viewing and walking speed

From the GPS logs with the record of the subjects' activities, we built a logistic regression model in Eq. 1, with which we can estimate from walking speed whether people are viewing an exhibition or not. In Eq. 1, P_i means the probability of viewing at point i, and v_i means walking speed at point i. When applied to the GPS logs obtained in our experiment, this model correctly discerned the subjects' viewing states over 68.4 percent of the entire logs. On the other hand, if the previous rule (i.e., we regard people as viewing exhibitions if they are within three meters from the exhibition) is applied to the same GPS logs, the success rate drops down to 56.7%. From this fact, we can conclude that the time people have spent in front of an exhibition is not sufficient to estimate the time they have viewed it.

$$\text{logit}(P_i) = 0.6194 - 0.3237 v_i \tag{1}$$

Estimate	Standard deviation	Chi-squared	p-value	Exp(estimate)
0.6194	0.0226	752.1810	0.0000	
-0.3237	0.0105	948.5306	0.0000	0.7235

The logistic regression model in Eq. 1 can be applied to the GPS logs recorded in our initial survey. For example, Fig. 6 shows the GPS log of a visitor in our initial survey, together with the presumption of his viewing state. In this case, he probably viewed the exhibition only first half of his stay in front of this exhibition. By counting the number of these points where he is presumed viewing an exhibition, we can estimate his viewing time on this exhibition, as each point corresponds to one second. Indeed, by repeating this process for everybody, we calculated the average of estimated viewing time for each animal.

For the calculation of average estimated viewing time, we used the GPS logs of 26 visitors who were in their 20s and accompanied by a single person, because the model in Eq. 1 was developed based on the data of subjects in their 20s followed by a staff. In addition, we assumed that each subject views the nearest animal exhibition when the subject was presumed viewing something.

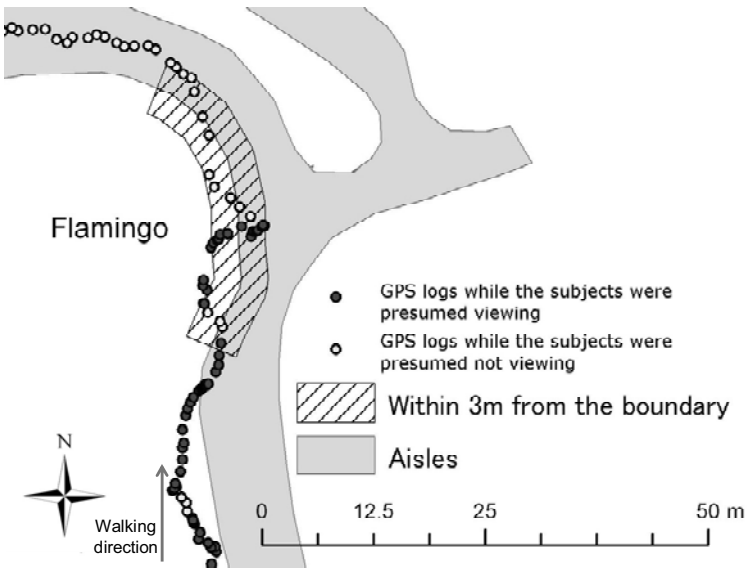


Fig. 6. A sample GPS log with the visitor's presumed viewing states at each location

Table 4 compares two rankings of animal exhibitions: the left one features average staying time, which was calculated by our previous method (i.e., counting the number of log points within three meters from the boundary between an exhibition and an aisle), while the right one features average viewing time estimated by our new method. The result shows several differences between staying time and viewing time. First, the viewing time of lions (5'08") is much shorter than its staying time (7'20"). This is because the lions' exhibition has a long aisle along it and accordingly, people have spent long time for walking the aisle often without viewing the lions. The same reason applies to Insect museum which has long fixed route in it. On the other hand, the viewing time of wolves and tigers (4'41") is estimated much longer than its staying time (2'56").

This is probably because many visitors have viewed these animals from a distance, particularly the space under trees, rather than the area near their cages. In sum, in order to evaluate how much people are involved in each exhibition, it is better to consider not only their location, but also their moving speed. In addition, other elements such as distance and angle to each exhibition may work as additional clues for discerning whether people view it or not. Thus, we are currently working on the refinement of our logistic regression model by considering additional parameters. In addition, we are currently conducting an additional experiment with non-student subjects (e.g., families with children) to expand the target of our model.

Table 4. Two rankings of animal exhibitions in terms of average staying and viewing times, estimated from the GPS logs of 26 visitors in 20s coming with another person

Rank	Animal Exhibition	Staying Time	Rank	Animal Exhibition	Viewing Time
1	Lions	7'20"	1	Lions	5'08"
2	Insect Museum	5'43"	2	Giraffes & Zebras	4'43"
3	Giraffes & Zebras	5'14"	3	Wolves & Tigers	4'41"
4	Japanese Monkeys	4'06"	4	Insect Museum	4'34"
5	Chimpanzees	3'09"	5	Japanese Monkeys	3'40"

5 Conclusion

The result of our experiment shows that not only the location history of visitors, but also their walking speed is essential for judging their viewing states and eventually for evaluating the exhibition's attractiveness. The information about exhibitions' attractiveness will be important for considering how to improve exhibitions, spatial designs, and tourist information of the zoological park. For instance, animal exhibitions which small children like can be relocated near the entrance, because children cannot walk long distance. On the other hand, animal exhibitions which attract adults rather than children may better be equipped with an adult-oriented guide board that explains the social aspect of the animal, etc.

In this work, the subjects are limited to students in their 20s. To increase generality, we are currently conducting a similar experiment where the subjects are general visitors accompanying small children, as they are the main target of zoological parks. In addition, we are refining our model by adding other parameters than location and walking speed, such as acceleration and walking angle to the exhibition.

In future work, we will refine the model, such that we can discern a larger variety of tourists' activities. In addition, we will conduct similar surveys in other sorts of tourist facilitates, such as amusement parks and open-air museums, and examine whether our technique can be applied generally to these facilities.

References

Asakura, Y. & Hato, E. (2004). Tracking survey for individual travel behaviour using mobile communication instruments. *Transportation Research Part C: Emerging Technologies* 12(3-4): 273-291.

Asakura, Y. & Iryo, T. (2007). Analysis of tourist behaviour based on the tracking data collected using a mobile communication instrument. *Transportation Research Part A: Policy and Practice* 41(7): 684-690.

Kaplan, E D. (1996). *Understanding GPS, Principles and Applications*. Artech House.

Kurata, Y., Yabe, N., Komaki, N., Arima, T. & Sugimoto, K. (2010). What, When, and How Much Do They See, and What Are They Interested In?: Toward More Detailed Survey on Inbound Tourists' Behaviors. In *Research Conference of Society for Tourism Informatics*, 43-48 (in Japanese).

Nagao, M., Kawamura, H., Yamamoto,M. & Ohuchi, A.(2004). Acquisition of Personal Tourism Activity Information Based on GPS Log Mining Method. In *Fifth Asia-Pacific Industrial Engineering and Management Systems Conference*, CD-ROM.

Shoval, N. (2008). Tracking Technologies and Urban Analysis. *Cities* 25(1): 21-28.

Shoval, N. & Isaacson, M. (2007a). Tracking Tourist in the Digital Age. *Annals of Tourism Research* 34(2): 141-159.

Shoval, N. & Isaacson, M. (2007b). Sequence Alignment as a Method for Human Activity Analysis. *Annals of the Association of American Geographers* 97(2): 282-297.

Yabe, N., Arima, T., Okamura, Y. & Kadono, A. (2010). An Agenda on Tourist Activity Survey Using GPS and Investigation of Analysis Methods. *International Journal of Tourism Science* 3: 17-30 (in Japanese).

Zheng, Y., Li, Y., Zha, Z. & Chua, T. (2011). Mining Travel Patterns from GPS-Tagged Photos. In K. Lee, W. Tsai, H. Liao, T. Chen, J. Hsieh and C. Tseng (Eds.), *Advances in Multimedia Modelling*, Lecture Notes in Computer Science 6523: 262-272, Springer.

Acknowledgements

We appreciate the visitors to Tama Zoological Park who participated in our surveys, as well as its staff who kindly helped our surveys and gave valuable comments to us. This work was supported by JST RISTEX's research program on Needs-oriented Service Science, Management and Engineering (project title: "Architecting service with customer participation based on the analysis of customer experience and design processes: sophisticating tour design processes as a case study") and Tokyo Metropolitan University's internal research fund.

Identification of Innovative Users for New Service Development in Tourism

Rita Faullant,
Ines Krajger, and
Markus Zanker[1]

Alpen-Adria-Universität Klagenfurt,
9020 Klagenfurt, Austria
{firstname.secondname}@aau.at

Abstract

The integration of customers into the development process of new products and services is crucial for companies regardless of the industry in which they operate. The advent of the Internet and new media applications has transformed value creation processes of the tourism industry to an extent unseen in other industry sectors. Electronic information mediators, such as information and booking portals of DMOs (destination management organisations) are not only becoming an increasingly dominant source of information and opinions for tourists but also for tourism service providers. The latter motivates our work to identify the innovative users from all visitors of an electronic information platform. In this paper we present findings from a research project, where we identified innovative users based on lead user characteristics as well as by a virtual idea contest. The paper's findings are controversial with respect to mainstream innovation research as they indicate that scales used in traditional innovation research disappoint in the tourism context. With our research we contribute to the existing literature in three ways: 1) we explore whether the lead user concept can also be applied for new service development in spa tourism; 2) we investigate the innovative potential of the identified lead users by assessing their creativity through an idea contest; and 3) thereby evaluate whether lead users in spa tourism are able to substantially contribute to new service development.

Keywords: Open innovation, lead user identification, virtual idea generation, new service development, classification.

1 Introduction

For most services in general, and for most tourism offers in particular, the customer traditionally entails a critical and integral role. The service itself is seen as highly dependent upon the interaction quality between service provider and customer (Bolton & Drew, 1991; Grönroos, 1993; Parasuraman et al., 1988). Despite the acknowledgement of the customers' central role for tourism offers, it is surprising that customers are rarely integrated into new service development for tourism offers. In sharp contrast, customer integration into new product development has been established on a systematic basis already in the early 1980s for physical goods (Urban & von Hippel, 1988; von Hippel, 1978). In particular, the most demanding and advanced customers, so called lead users (von Hippel, 1986), have been tried to be filtered out and integrated into new product development. Several studies suggest that the integration of users into new product development is an appropriate means for companies to come up with faster and more customer-centered innovations (for example Gruner & Homburg, 2000). Also for the service industry potential for

[1] Corresponding author

various stages and modes of user involvement has been highlighted (Alam, 2002). With our study we aim at testing whether the concept of lead-user integration is in principle also applicable for new service development in tourism. We invited users of a Web Portal focusing on spa vacations to participate in a virtual idea competition and to submit ideas for new service offers in spa tourism. Furthermore we assessed their lead user characteristics through a web-based questionnaire. Based on creativity theory an independent jury of spa experts evaluated these ideas, which were then correlated with the users' lead user characteristics.

2 Literature review

2.1 Lead Users in New Product Development

Within the literature on innovation management, user integration into new product development has become an important research field. Instead of solely considering users as information providers, the potential of users to actively engage in the new product development process gains more and more attention (Evardsson et al. in press; von Hippel, 2005b). Previous research confirms the ability of users to contribute to the NPD (new product development) process (Alam, 2006; Füller et al., 2007; Lilien et al., 2002; Oliveira & von Hippel, 2011; Skiba & Herstatt, 2009). However, only a small proportion of users namely between 10 and 40 percent, has the know-how, creativity and expertise for truly innovative problem solutions that are not restricted to line-extensions or incremental innovations (von Hippel, 2005a). Since the value of customer contributions in the development of new products and services varies significantly, it is crucial to carefully select the right users to be integrated into new product or service development (Enkel et al., 2005; Gruner & Homburg, 2000). One group of users that has been shown to be able to deliver highly innovative suggestions for new product development are lead users.

Lead Users are different from other users because they (a) have needs that will become commonplace in a market before the bulk of the other users encounters them and (b) they expect to benefit significantly from obtaining a solution to those needs (von Hippel, 1986). These characteristics are also known as the 'Ahead of Trend' (AT) dimension and the 'High Expected Benefit' (HEB) dimension, respectively (Franke et al., 2006). Many scientific contributions and practical experiences have been published, confirming that products developed in cooperation with lead users are appreciated as highly innovative by firms. (Franke et al., 2006; Lilien et al., 2002; Lüthje, 2000). The ability to bear innovative solutions is fundamentally linked to a person's individual creativity (Faullant et al., 2009). In psychology creativity is generally defined as "the production of novel, useful ideas or problem solutions" (Amabile et al., 2005, p. 368). The first aspect emphasises the originality or unexpectedness of an idea (Sternberg & Lubart, 1999). The second aspect stresses that an idea must be of value, or "appropriate (i.e., useful, adaptive concerning task constraints)" (Sternberg & Lubart, 1999) which is especially important for new product development. In the past, many studies confirmed that lead users are able to produce both novel and useful ideas.

Analyzing the markets in which these studies have been conducted, the literature review shows that initial lead user studies concentrated predominantly on industrial goods markets (Franke & von Hippel, 2003; Herstatt & von Hippel, 1992; Lüthje,

2003; Morrison et al., 2000; Olson & Bakke, 2001; Urban & von Hippel, 1988). Meanwhile a bulk of research confirms that the identification of lead users is also promising for user integration in consumer mass markets such as kite surfing, extreme sporting equipment, technical diving, kitchen appliances (Franke & Shah, 2003; Füller et al., 2006; Füller et al., 2007; Lüthje, 2004; Lüthje et al., 2005; Schwarz et al., 2009). Within new service development systematic lead user identification and their integration for service innovation has been widely neglected (Skiba & Herstatt, 2009). In their recent work Edvardson et al. (forthcoming) proposed a conceptual framework of methods of customer integration into new service development. The lead user method was classified as a method being able to generate highly novel service solutions, but at the same time requiring high methodological competences. Indeed, recent empirical evidence confirms the potential of user innovation also for the service sector (Oliveira & von Hippel, 2011). In tourism so far, little is known about lead user identification and their involvement in new service development.

2.2 Virtual User Integration for New Product and Service Development

Recent studies highlight the potential of the Internet to use the creative potential of consumers. The use of the World Wide Web allows companies to reach potential users world-wide for new product development (Hienerth & Füller, 2004; Sawhney et al., 2005). This is accompanied by the development of new tools and methods for virtual user integration (e.g. Dahan & Hauser, 2002; Dahan & Srinivasan, 2000; Franke & Piller, 2004; Füller et al., 2007; Jeppesen, 2005; Verona et al., 2006). Web-based methods such as idea competitions, toolkits for user innovation, virtual worlds, virtual stock markets and virtual communities have already diffused in practice supporting collaborative new product development (Bullinger et al., 2010; Ebner et al. 2009).

For the service sector in general Sigala (2010) provided insights from the Starbucks community that virtual user communities are able to generate, shape, and co-create ideas for new service development. The shared interpretation of an idea throughout the community can lead to different cultural interpretations of what a new service might constitute. In tourism, the potential of user communities for the development of new touristic products has been recognized already in the early 2000s (Wang et al.. 2002).

Recent studies however confirm that the main activities in user communities and blogs are still limited to information exchange such as sharing and documenting travel experiences and ratings of tourism products (Dippelreiter et al., 2008; Waldhör & Rind, 2008; Yoo & Gretzel, 2008). An empirical study of Greek tourism companies illustrates that companies predominantly use online tools for marketing purposes such as monitoring user reviews and comments, managing complaints as well as for market research and segmentation. For new service development users' feedback or comments are not used. However, even in tourism companies have to shift from product designing for users to co-creation with users (Sigala, 2008). With our study we aim at investigating whether users have the potential to substantially contribute to new service development and whether those users can be identified by web-based means.

3 Methodology

3.1 Data Collection

An empirical study was executed in the field of spa service offers in collaboration with www.thermencheck.com, a major virtual spa platform in German speaking countries. Visitors of this platform receive information on and can book offers of several hundred different spas in Germany, Austria, and Slovenia. At the front page a banner with the invitation to submit innovative ideas for spa development and new service creation was placed. We assumed to attract lead users as well as ordinary users with this invitation. The banner directly linked to the briefing page of the questionnaire. Among all submitted ideas a jury of tourism experts in the spa domain determined the winning ideas, and as an incentive spa vouchers worth over 1000 Euros were distributed. After describing their idea in a text field as detailed as possible, participants were asked to complete a standardized web-based questionnaire. Data collection took place from 1 January to 30 April 2011. During this period over 74.000 distinct accesses (excluding bots) to the platform were recorded.

3.2 Measures and Jury Evaluation

To assess participants' extent of being a lead user the two hallmark characteristics of being *ahead of trend* (AT) and *high expected benefit* (HEB) were used from existing measures in the literature (Lüthje, 2000). The scales were adapted to the spa context and were measured as a continuous variable on a 7-point Likert scale as denoted in Table 1. Furthermore, we collected demographics like gender, age, nationality or education as well as data about the actual spa usage of participants (SPA).

Table 1. Codes of questionnaire items

Code	Scale
AT1	I'm regarded as being well informed in the field of spa offers.
AT2	I usually determine new spa offers earlier than most other people.
AT3	I try to visit just recently opened spas.
HEB1	I have needs and preferences which are not satisfied by spa offers.
HEB2	During my past visits of spa resorts I noticed shortcomings several times.
HEB3	I'm dissatisfied with the existing spa resort offers.
SPA1	How often do you visit spa resorts each year.
SPA2	How many different spa resorts have you visited up to now?
SPA3	Which of the following eight recently opened spa resorts have you already visited?

To evaluate and rank the submitted ideas the Consensual Assessment Technique (CAT) (Amabile, 1982) was applied. According to this method "a product or response is creative to the extent that appropriate observers independently agree it is creative." (Amabile, 1982, p. 1001). 4 independent experts (managers of spas and spa regions,

and responsible persons in spa distribution channels) have been asked to assess the quality of submitted ideas on three dimensions: originality of the idea (short: originality), customer value of the idea (utility), overall impression (overall) (Kristensson, et al. 2002; Walcher, 2007).

The number of jury members is in line with the state-of-practice of CAT in other domains (Piller & Walcher, 2006). These dimensions were presented on a 5-point scale from 1 (very low) to 5 (very high). The experts rated all ideas independently from each other (and in different randomly assigned orders), and based on their evaluation on their own subjective definition of the dimensions. All experts rated some ideas with 0, that is interpreted as *not even very low*. Consequently, ideas are actually rated on a 6-point scale.

4 Analysis and Results

4.1 Sample Description

In total 161 participants filled out the self-administered questionnaire, and 122 submitted ideas or suggestions for spa service development (i.e. ~0.15% of all unique client accesses during that period participated and submitted their ideas). For the further analysis only fully matching cases have been included, i.e. completed the questionnaire and submitted an idea. After data purification the sample sized resulted in 110 cases (74% females, median age group between 21 and 30, over 50% higher education or university degree, nearly 75% Austrians and the remaining participants, with the exception of one Italian, from Germany).

4.2 Collected Data

Descriptive statistics of replies to the adapted questionnaire on AT and HEB items (scale 1-7, 1: I do not agree at all, 7: I fully agree) are given in Table 2. Submitted ideas have an average length of 555 characters (standard deviation: 1411.5, min: 17, max: 14501) and Figure 1 depicts their length distribution with one extremely long outlier. The most dominant themes for innovation are the different needs of adults and children that should be addressed in separate locations and ideas for designing the relaxation and recreation areas. On average respondents have 7.18 spa visits per year (SPA1), know 6.25 different spa resorts (SPA2) and have already tried 1.4 out of the 8 newly opened spa resorts that were named in the questionnaire (SPA3).

Table 2. Statistics on questionnaire items

Item	Mean	Standard deviation
AT1	4.03	1.87
AT2	4.4	1.89
AT3	4.36	1.78
HEB1	2.53	1.68
HEB2	4.44	1.69
HEB3	2.72	1.66
Scale reliability: Cronbach alpha = 0.74		

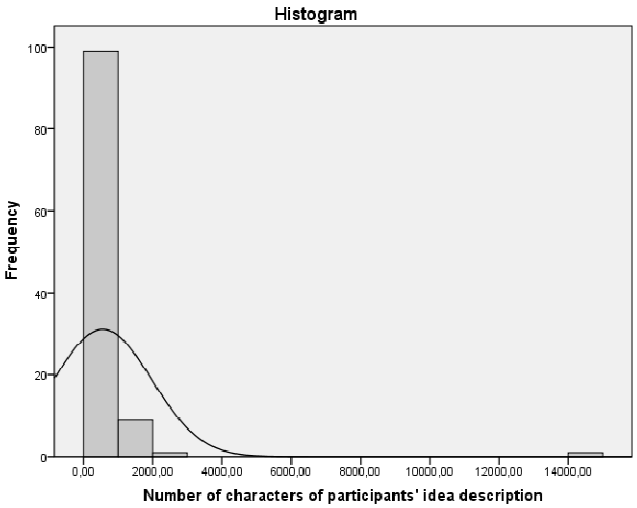


Fig. 1: Distribution of character length of submitted ideas

Third, we describe the rating feedback that has been collected from the expert panel (see Table 3). Ideas received an overall rating as well as ratings for their originality and their presumed utility for customers. Between 1 and 21 ideas received the highest possible overall rating from at least one domain expert.

Table 3: Statistics on idea ratings

				Std.	Frequency of ratings					
Expert	**Scale**	**N**	**Mean**	**dev.**	**0**	**1**	**2**	**3**	**4**	**5**
A	*Overall*	110	2.26	1.64	25	9	25	28	9	14
	Originality	110	2.31	1.61	24	9	21	34	9	13
	Utility	110	2.7	1.73	23	3	17	29	17	21
B	*Overall*	108	2.79	0.938	4	2	31	47	24	0
	Originality	108	2.11	1.02	4	26	44	22	12	0
	Utility	108	2.58	1.02	4	11	30	44	19	0
C	*Overall*	110	1.87	1.23	16	29	29	27	7	2
	Originality	110	1.51	1.48	40	18	25	15	7	5
	Utility	110	1.99	1.3	16	26	26	30	9	3
D	*Overall*	110	2.15	0.927	1	28	43	31	6	1
	Originality	110	1.95	1.04	1	43	41	13	10	2
	Utility	110	2.43	1.04	0	24	34	35	15	2

In order to adjust ratings for different mean and deviation, we standardized ratings by computing mean-deviation divided by standard deviation leading to standard deviation units. Next, Pearson correlation coefficients between standardized ratings of experts/judges are pairwise computed and Cronbachs alpha for scale reliability is given. Table 4 and Table 5 give details. All correlation coefficients are significant (2-sided) at the $p = 0.01$ level.

Table 4: Interjudge reliability of standardized overall ratings

Overall	A	B	C	D
A	1	0.396**	0.442**	0.428**
B		1	0.348**	0.414**
C			1	0.678**
D				1

Scale reliability: Cronbachs alpha = 0.777

Table 5: Interjudge reliability of standardized originality and utility ratings

Originality	A	B	C	D
A	1	0.424**	0.629**	0.599**
B		1	0.348**	0.414**
C			1	0.775**
D				1

Scale reliability: Cronbachs alpha = 0.833

Utility	A	B	C	D
A	1	0.367**	0.489**	0.389**
B		1	0.320**	0.382**
C			1	0.680**
D				1

Scale reliability: Cronbachs alpha = 0.766

4.3 Analysis

The goal of the analysis is to assess how well the two hallmark characteristics for determining lead users (Lüthje, 2000), namely being ahead of trend (AT) and high expected benefit (HEB), conform to the opinions (i.e. ratings) of our expert panel on the participants' ideas. For that purpose, we again apply correlation analysis to research how well the idea ratings' variance corresponds to the variance in the participants' answers to the lead user characteristics. In Table 6 we compare the averaged standardized ratings of our judges with each of standardized replies to the AT and HEB scales as well as to the participants' spa experience (SPA1-SPA3).

Table 6: Pearson correlation between standardized questionnaire items and ratings

	Overall	Originality	Utility
HEB1	0.067	0.001	0.058
HEB2	**0.277****	**0.260****	**0.267****
HEB3	-0.050	-0.062	-0.148
AT1	0.229	0.216	0.155
AT2	**0.256****	**0.261****	0.149
AT3	0.057	0.086	-0.036
SPA1	0.136	0.093	0.131
SPA2	0.097	0.112	0.063
SPA3	0.088	0.075	0.052

With respect to the controlled demographic criteria such as age, gender, nationality or education we can state that none of them had a significant influence on the ratings. However, not unexpectedly, the standardized length of the idea descriptions showed high correlation coefficients with the standardized ratings (length/overall: 0.449**,

length/originality: 0.435** and length/utility: 0.421**). Note, that 2-sided statistical significance at the p=0.01 level is denoted by **. To summarize, the results of our analysis indicate that existing scales for determining lead users based on HEB and AT scales do not seem to be appropriate for the tourism domain. Only two questionnaire items (namely HEB2: *During my past visits of spa resorts I noticed shortcomings several times*. and AT2: *I usually determine new spa offers earlier than most other people*.) showed significant Pearson correlation coefficients with experts' judgments on submitted ideas.

However, low correlation coefficients of r~0.25 indicate that only around 6% of the variance (i.e. r^2) is shared. In contrast, the length of submitted idea descriptions shared around 20% of the variance with the averaged expert ratings. Thus, we can conclude that the utility of the described HEB and AT scales for identifying innovative users in the tourism context is rather low and that further research is needed to develop characteristics and measures that help to predict the degree of innovativeness of online users for new tourism service development. Furthermore, experts' ratings of participants ideas did not significantly correlate with the standardized participants' experience parameters such as spa visits per year (SPA1) or number of known spa resorts (SPA2). However, average spa experience of our sample was already relatively high, due to respondents' self selection to participate in the idea contest. Furthermore, we have good reasons to reject the conjecture that respondents might have overestimated their own status in terms of being *ahead of the trend*, because participants' experience with different spa resorts (SPA2) did significantly correlate with all questionnaire items of the *ahead of trend* scale.

5 Conclusions

The presented work constitutes one step towards identifying indicators and predictors for determining innovative customers in the tourism domain. We adapted the two hallmark scales (*high expected benefit* and *ahead of trend*) from traditional innovation research and evaluated if they predict participants' high quality ideas in the opinions of expert judges. The findings of our research are that these scales for measuring lead user characteristics are inappropriate for determining lead users in the spa domain. Only few questionnaire items can explain some of the variance in the experts' ratings variance. For these unexpected findings several explanations can be offered: 1) in the classic lead user method users are integrated from the early stage of idea generation throughout the whole process of new product development in several workshops with firm-intern experts. Our users had only limited time (constrained by their own time at disposition) to describe their ideas, and had no possibility to get feedback and improve their suggestions. Therefore the quality of their ideas might be inferior compared to extended workshops; 2) the evaluation of ideas was carried out by 4 experts in the spa domain. This procedure reflects the common practice used to judge the creative potential of solutions (Amabilie, 1996; Füller et al., 2011). It is however a judgment based on the opinion of a few, and it might well be, that a bulk of ordinary users would judge the generated ideas differently and lead to diverging results; 3) finally it might be that the scales for screening lead users in mass markets are not thus easily transferable to service markets although initial studies suggested good potential (Skiba & Herstatt, 2009; Oliveira & von Hippel, 2011).

From a managerial perspective when trying to involve users into new service development a pre-selection of users based on lead-user characteristics cannot be recommended at this stage. Further research towards identifying the characteristics of innovative tourists is required. We will therefore research if the digital traces of participants will help to identify the innovative users in the large online crowds as part of our ongoing work.

References

Alam, I. (2002). An exploratory investigation of user involvement in new service development. *Journal of the Academy of Marketing Science* 30 (3): 250-261.

Alam, I. (2006). Removing the fuzziness from the fuzzy front-end of service innovations through customer interactions. *Industrial Marketing Management* 35: 468-480.

Amabile, T., Barsade, S. G., Mueller, J. S., et al. (2005). Affect and Creativity at Work. *Administrative Science Quarterly* 50: 367-403.

Amabile, T. M. (1982). Social Psychology of Creativity: A Consensual Assessment Technique. *Journal of Personality and Social Psychology* 43 (5): 997-1013.

Bolton, R. N. & Drew, J. H. (1991). A multistage model of customers' assessments of service quality and value, *Journal of Consumer Research* 17 (March): 375-384.

Bullinger, A. C., Neyer, A.-K., Rass, M., et al. (2010). Community-Based Innovation Contests: Where Competition Meets Cooperation. *Creativity & Innovation Management* 19 (3): 290-303.

Dahan, E. & Hauser, J. (2002). The virtual customer. *Journal of Product Innovation Management* 19 (5): 332-53.

Dahan, E. & Srinivasan, V. (2000). The predictive power of internet-based product concept testing using visual depiction and animation. *Journal of Product Innovation Management* 17: 99-109.

Dippelreiter, B., Grün, C., Pöttler, M., et al. (2008). Online tourism communities: On the path to web 2.0: an evaluation. *Information Technology & Tourism* 10: 329-353.

Ebner, W., Leimeister, J. M., and Krcmar, H. (2009). Community engineering for innovations: the ideas competition as a method to nurture a virtual community for innovations. *R&D Management* 39 (4): 342-356.

Enkel, E., Perez-Freije, J. & Gassmann, O. (2005). Minimizing Market Risks Through Customer Integration in New Product Development: Learning from Bad Practice. *Creativity and Innovation Management* 14 (4): 425-437.

Evardsson, B., Kristensson, P., Magnusson, P., et al. (in press). Customer integration within service development - a review of methods and analysis of insitu and exsitu contributions. *Technovation*.

Faullant, R., Schwarz, E., Krajger, I., et al. (2009). Towards a comprehensive understanding of lead userness: the role of personality and creativity. In *16th International Product Development Management Conference*, 7 - 9 June 2009, Enschede/NL. Enschede/NL: EIASM.

Franke, N. & Piller, F. (2004). Value Creation by Toolkits for User Innovation and Design: The Case of the Watch Market. *Journal of product innovation management* 21: 401-415.

Franke, N. & Shah, S. (2003). How communities support innovative activities: an exploration of assistance and sharing among end-users. *Research Policy* 32 (1): 157-178.

Franke, N. & von Hippel, E. (2003). Satisfying heterogeneous user needs via innovation toolkits: the case of Apache security software. *Research Policy* 32: 1199-1215.

Franke, N. & Hippel, E. & Schreier, M. (2006). Finding commercially attractive user innovations: A test of the Lead-User Theory. *Journal of product innovation management* 23: 301-315.

Füller, J., Bartl, M., Ernst, H., et al. (2006). Community Based Innovation: How to integrate members of virtual communities into new product development. *Electronic Commerce Research* 6 (1): 57-73.

Füller, J., Hutter, K. & Faullant, R. (2011). Why co-creation experience matters? Creative experience its impact on the quantity and quality of creative contributions. *R&D Management* 41 (3): 259-273.

Füller, J., Jawecki, G. & Mühlbacher, H. (2007). Innovation creation by online basketball communities. *Journal of Business Research* 60: 60-71.

Grönroos, C. (1993). Toward a third phase in service quality research: challenges and future directions. In T. A. Swartz, D. E. Bowen and Stephen Brown (Eds.), *Advances in services marketing and management: research and practice*. London.

Gruner, K., E. & Homburg, C. (2000). Does customer interaction enhance new product success? *Journal of business research* 49: 1-14.

Herstatt, C. & von Hippel, E. (1992). From Experience: Developing New Product Concepts via the Lead User Method: A Case Study in a "Low-Tech" Field. *Journal of Product Innovation Management* (9): 213-221.

Hienerth, C. & Füller, J. (2004). Engaging the creative consumer. *European Business Forum* 19: 54-58.

Jeppesen, L. B. (2005). User Toolkits for Innovation: Consumers Support Each Other. *Journal of Product Innovation Management* 22 (4): 347-362.

Kristensson, P., Magnusson, P. R. & Matthing, J. (2002). Users as a Hidden Resource for Creativity: Findings from an Experimental Study on User Involvement. *Creativity and Innovation Management* 11 (1): 55-61.

Lilien, G. L., Morrison, P. D., Searls, K., et al. (2002). Performance Assessment of the Lead User Idea-Generation Process for New Product Development. *Management Science* 48 (8): 1042-1059.

Lüthje, C. (2004). Characteristics of innovating users in a consumer goods field. An empirical study of sport-related product consumers. *Technovation* 24: 683-695.

Lüthje, C. (2003). Customers as Co-Inventors: An empirical analysis of the antecedents of customer-driven Innovations in the field of medical equipment. In *Proceedings of the 32th EMAC Conference*, Glasgow.

Lüthje, C. (2000). *Kundenorientierung im Innovationsprozess*. Wiesbaden: Universitäts-Verlag.

Lüthje, C., Herstatt, C. & von Hippel, E. (2005). User innovators and "local" information: The case of mountain biking. *Research Policy* 34: 951-965.

Morrison, P. D., Roberts, J. H., von Hippel, E. (2000). Determinants of user innovation and innovation sharing in a local market. *Management Science* 46 (12): 1513-1527.

Oliveira, P. and von Hippel, E. (2011). Users as service innovators: The case of banking services. *Research Policy*, 40, 806-818.

Olson, E. L. & Bakke, G. (2001). Implementing the lead user method in a high technology firm: a longitudinal study of intentions versus actions. *Journal of product innovation management* 18: 388-395.

Piller, F. & Walcher, D. (2006). Toolkits for idea competitions: a novel method to integrate users in new product development. *R&D Management* 36(3): 307-318.

Parasuraman, A., Zeithaml, V. A. & Berry, L. L. (1988). SERVQUAL: A multi-item scale for measuring consumer perceptions of service quality. *Journal of Retailing* 64 (1): 12-40.

Sawhney, M., Verona, G. & Prandelli, E. (2005). Collaborating to create: The Internet as a platform for customer engagement in product innovation. *Journal of Interactive Marketing* 19: 4-17.

Schwarz, E., Faullant, R., Krajger, I., et al. (2009). Are leading edge users creative? In *Proceedings of the 38th EMAC Conference*. Brussels: European Marketing Academy (EMAC).

Sigala, M. (2010). Managing customer involvement and roles in New Service Development (NSD): lessons learnt from www.mystarbucksidea.com. In *28th EuroCHRIE Conference Amsterdam*, The Netherlands.

Sigala, M. (2008). Developing and Implementing an eCRM 2.0 Strategy: Usage and Readiness of Greek Tourism Firms. In *Enter Proceedings.*

Skiba, F. & Herstatt, C. (2009). Users as sources for radical service innovations: opportunities from collaboration with service lead users. *International Journal of Services Technology & Management* 12 (3): 317-337.

Sternberg, R. J. & Lubart, T. I. (1999). The Concept of Creativity: Prospects and Paradigms. In Robert J. Sternberg *(Ed.), Handbook of Creativity.* Cambridge: Cambridge University Press.

Urban, G. L. & von Hippel, E. (1988). Lead user analyses for the development of new industrial products. *Management Science* 34 (5): 569-582.

Verona, G., Prandelli, E. & Sawhney, M. (2006). Innovation and virtual environments: towards virtual knowledge brokers. *Organization Studies* 27 (6): 765-788.

von Hippel, E. (2005a). *Democratizing Innovation*. The MIT Press.

von Hippel, E. (2005b). Democratizing innovation. The evolving phenomenon of user innovation. *Journal für Betriebswirtschaft* 55: 63–78.

von Hippel, E. (1986). Lead Users: A source of novel product concepts. *Management Science* 32 (7): 791-805.

von Hippel, E. (1978). Successful industrial products from customer ideas. *Journal of Marketing* 42 (1): 39-49.

Walcher, D. (2007). *Der Ideenwettbewerb als Methode der aktiven Kundenintegration: Theorie, empirische Analyse und Implikationen für den Innovationsprozess*. Wiesbaden: Deutscher Universitätsverlag.

Waldhör, K. & Rind, A. (2008). eBlogAnlaysis - Mining Virtual Communities using Statistical and Linguistic Methods for Quality Control in Tourism. In *Enter Proceedings*.

Wang, Y., Yu, Qu. & Fesenmaier, D. R. (2002). Defining the virtual tourist community: implications for tourism marketing. *Tourism Management* 23: 407-417.

Yoo, K. H. & Gretzel, U. (2008). What motivates consumers to write online travel reviews? *Information Technology & Tourism* 10: 283-295.

Acknowledgements

Parts of this work have been financed by the ÖNB grant no. 13.000 of the Austrian National Bank Jubilee Foundation. Furthermore, the authors would like to thank New Mediacheck GmbH (http://www.thermencheck.com) and domain experts from different spa resorts for their collaboration and support.

Differences in Social Presence Perceptions

Myunghwa Kang [a], and
Ulrike Gretzel [b]

[a] Department of Agricultural Leadership, Education & Communication
University of Nebraska, U.S.A.
mkang5@unl.edu

[b] Institute for Innovation in Business and Social Research
University of Wollongong, Australia
ugretzel@uow.edu.au

Abstract

The influence of technology-related characteristics on social presence perceptions has been established in the literature while the influence of the social setting in which the technology use occurs has been largely neglected. This paper looks specifically at the influence of opportunities for social interactions and motivations to engage in social interactions as measured by travel party size and composition and individual need for affiliation. Data was collected in a national park among visitors who listened to an interpretive podcast tour. The research finds support for differences based on need for affiliation. The effect of travel group size was only marginally significant and no support was found for travel parties with children versus those without. The findings indicate that need for affiliation is an important factor to consider and that the influence of social settings of technology use need to be further explored.

Keywords: need for affiliation; social presence; travel group structure; podcast tours.

1 Introduction

Recent studies have shown the role of information and communication technology (ICT) in facilitating interactions between tourists and tourism destinations (Benckendorff, Moscardo, & Murphy, 2006; O'Dell, 2007; Stamboulis & Skyannis, 2003). Quasi-social interactions represent an important aspect of such mediated tourism experiences. Their quality is often measured in terms of perceived social presence. Social presence refers to the sense of being with another human or human-like intelligence (Biocca, Harms, & Burgoon, 2003). Perceptions of social presence play a role in enhancing dimensions of tourist experiences such as learning, enjoyment, and escape (Kang & Gretzel, 2011). The ability of ICT to instil social presence perceptions is especially important for those destinations that, due to budget restrictions, can only offer limited human contact, for instance through visitor information staff or interpretive personnel.

Prior studies regarding social presence have mostly focused on ICT characteristics that foster these perceptions, with certain types of ICT being able to inspire greater feelings of social presence than others, e.g. due to their richness. Only a small number of studies has investigated differences in social presence perceptions as influenced by user characteristics. Specifically, differences in terms of socio-demographics, experiences, learning style, personality, and emotions have been suggested (Lee, 2002; Lombard & Ditton, 1997; Slater & Usoh, 1993; Witmer & Singer, 1998). Surprisingly, the social setting in which the ICT is used and the personal preference

for social versus solitary settings have not been investigated in terms of their influence on technology-induced perceptions of social presence. Therefore, this study focuses on the travel party size and structure as well as the personal need for affiliation to examine if and how they impact social presence perceptions.

2 Theoretical Foundations

2.1 Social Presence

The extent to which ICTs embody the real world physically and socially as it is has been a central question for many scholars as well as practitioners (Lee, 2002). Social presence is one aspect of this question that has received substantial attention in the literature. The concept has been recognized as a crucial part in understanding ICT environments from a social perspective (Whitelock, Romano, Jelfs, & Brna, 2000).

Social presence theory was originally proposed over 30 years ago by Short, Williams, and Christie (1976). For a recent review of social presence literature see Lowenthal (2009). The term social presence was defined as the "degree of salience of the other person in the interaction and the consequent salience of the interpersonal relationships" (Short, et al., 1976, p. 65). This conceptualization of social presence is therefore used mainly to assess how much socially rich interpersonal interaction a particular medium can support (Fulk, Schmitz, & Power, 1987; Walther, 1992; Short et al., 1976). In this context, the possibility for intimacy and immediacy is an important aspect of social presence (Reysen, Lloyd, Katzarska-Miller, Lemker, & Foss, 2010). The approach has some limitations in understanding human-computer interaction since social presence as defined above is applicable only to media users' experiences of a real human through communication media (Lee, 2002).

ICTs have developed quite substantially, allowing for a wide variety of interactions that can be social and "para-social", i.e. involving interactions with virtual personas/agents or other forms of intelligent systems (Weisberg, Te'eni & Arman, 2011). The concept of perceived social presence should therefore be rooted in the understanding of virtual experiences in not only ICT mediated environments but also in human and computer interaction. Accordingly, as Lee (2002) argued, two different virtual experiences, para-authentic and artificial experiences, should be taken into account when conceptualizing social presence. First virtual experience is technology-mediated experiencing. That is, what users experience "is not an actual object but the mediated version of it (para-authentic objects)" (Lee, 2002, p. 56). Second virtual experience can include the experience of artificial objects. That is, experienced objects do not actually exist in the world, yet are experienced as if they were (artificial objects).

Based on what the objects are that are experienced in virtual environments, Lee (2002) discussed three different types of virtual experiences: physical, social and self. The focus of this paper is on the social. According to Lee (2002), social experience is the mediated experience of social actors (e.g., humans or human-like intelligence). In the same way as people "mentally simulate virtual physical objects based on sensory cues provided by technology, people automatically simulate virtual social actors upon receiving cues manifesting humanness" (Lee, 2002, p. 63). Similarly, Biocca (1997)

connected these three different types of experiences to three different types of presence: physical, social and self presence. The concept of presence involves perceptual subjective feelings of immediacy and has been explored as an important determinant which influences the overall ICT use experience (Lee & Nass, 2004). As such, the concept of presence has great relevance to the design and evaluation of ICT interfaces in a variety of fields such as entertainment (e.g., games), education (e.g., online teaching), and marketing (e.g., online shopping). It is argued here that it is especially important in evaluating ICTs in the context of tourism, which involves rich physical, social and self experiences.

For human-computer interaction research, Lee (2002) defined social presence as the psychological state in which the virtuality of the social experience is not noticed. The latter conceptualization of social presence emphasises that individuals interact with a virtual or artificial entity or agent (artificial entity) or with a person who does not communicate in real time (para-authentic entity) (Lee, 2002). In this sense, social presence occurs when technology users do not notice the mediated nature of experiencing other humans (para-authentic experiences) and/or artificially simulated social actors (artificial experiences) (Lee & Nass, 2004).

Previous empirical studies have identified the positive effect of perceived social presence on ICT users' behaviour and attitudes in various fields such as marketing, education and tourism (Kang & Gretzel, 2011; Lee, 2002; Skalski, 2004). In particular, Kang and Gretzel (2011) found that perceptions of social presence can be enhanced by ICT (in that case mobile interpretive tours) at a tourism destination (e.g., a nature-based national park), and that the perception of social presence can positively influence tourist experiences in environments that otherwise lack social interactions with staff (e.g. park rangers). They concluded that even if communicated through audio-only media (i.e., podcast tours on MP3 players), podcast tours can create a positive social context which influences tourist experiences.

2.2 Determinants of Perceived Social Presence

Technological advances have increasingly enabled computer-generated entities to mimic both the appearance and behaviour of human beings by integrating voice, text, graphics and videos (Brent & Thompson, 1999). Web sites can increase social presence by adding a personal touch such as a personalized greeting or through pictures and texts that convey a personal presence in the same way as photographs and personal letters (Gefen & Straub, 2003). With mobile technologies, human entities rely more on sound/voice because of screen size and mobility. Voices, as audio cues, can also be used to foster social presence (Kang & Gretzel, 2011; Lee, 2002). Greater fidelity, higher screen resolutions, greater speed of transmission and higher degrees of personalization create a sound technological basis for fostering social presence.

As mentioned above, technological drivers of social presence perceptions have been explored quite extensively in the literature. Witmer and Singer (1998) present four different categories – control (degree, mode and immediacy of control; anticipation of events), sensory modality (environmental richness; multi-modal presentation; consistency of multi-modal information; degree of movement perception; active search), distraction (isolation; selective attention; interface awareness), and realism

(sense realism; information consistency with objective world; meaningfulness of experience; disorientation). Lombard and his colleagues (Lombard, 2004; Lombard & Ditton, 1997) go beyond the purely technological and emphasize three types of factors – media form (image size and quality; audio fidelity; camera techniques), contents (social physical realism; use of media convention) and media user (willingness to suspend disbelief; prior experience with medium; age; gender). Slater and Usoh (1993) broadly discussed two main factors – technology-dependent factors and participant-dependent factors. Lee (2002) distinguishes among three factors: technology (objective quality of technology), user (individual differences), and social (social characteristics of technology). Lee (2002) lists experiences and familiarity with technology, adoption and learning of system features, socio-demographics, other psychological factors (mood, perceived risk) and personality characteristics as individual user-relevant factors. Despite the acknowledgment of personality characteristics as influence factors, need for affiliation has not yet been studied in the context of social presence.

Effect of Need for Affiliation On Social Presence. Personality is generally referred to as consisting of invariant personal characteristics (i.e., dispositions) that distinguish different individuals and lead to consistencies in behaviour across situations and over time (Villani & Wind, 1975). Personality has been used to predict individual differences in a number of areas, including product choice (Lastovicka & Joachimsthaler, 1988). In tourism, personality has been identified as a predictor of travellers' motivations, preferences, intentions, and behaviours, and is generally seen as a solid criterion for segmenting travellers along with other psychographics (e.g., attitudes, values, lifestyle), and demographics and geographic variables (Weaver, Weber, & McCleary, 2007).

Personality theorists have shown that need for affiliation explains individual differences in social motives and have explored its influence on social behaviour. The construct has been generally conceptualized as a personality trait (Heyns, Atkinson, & Veroff, 1958; Hill, 1997; Murray, 1938) and appears in the literature under various terms, such as affiliative motive, need for affiliation, affiliative need, and affiliation motivation. While motivation theorists see the concept as a motive (O'Connor & Rosenblood, 1996), personality and trait theorists have studied it broadly as a personality type or sub-dimension of personality. This study uses the term need for affiliation in the sense of a stable disposition that organizes or explains what an individual says, does and feels based on Hill's (1987) conceptualization.

Need for Affiliation (n-Aff) is defined as showing a tendency of establishing, maintaining, or restoring a positive affective relationship with another person or group of people (Atkinson, et al., 1954). In the Murray framework (Murray, 1938), n-Aff was described as gratification derived from harmonious relationships and a sense of communication. n-Aff is generally believed to motivate individuals to seek out social contact (O'Connor & Rosenblood, 1996), to form friendships and associations, and to enjoy being with friends and people in general (Murray, 1938). n-Aff drives feelings of affection, the need for liking and love as a major social reward and incentive (Buss, 1983; Foa & Foa, 1974), as well as feelings of belongingness and intimacy (Veroff & Veroff, 1980). Hill (1991) showed that the intrinsic interest in being around and interacting with others is a major influence contributing to social

thoughts or behaviours. For this reason persons with strong affiliative motives are expected to be more likely to engage in direct action aimed at attaining available emotional support and to actively encourage and facilitate positive interaction with another individual (Hill, 1991; 1997).

Human beings are more motivating to process and interpret information in ways that fits their dispositions. Accordingly, we hypothesize that:

> H1: Individuals with greater n-Aff have a greater likelihood of interpreting cues in interactions with technology as social and, thus, will perceive greater social presence.

Effect of Travel Party Characteristics On Social Presence. Along with travellers' demographics and socio-demographics, trip characteristics have been identified as important factors influencing tourism experiences. Tourism scholars have examined the influence of trip characteristics on travellers' purchase decisions, travel motivation, preferences, destination image creation and perceptions, satisfaction and loyalty, post-trip evaluation, expenditure, etc. (Baloglu, 1997; Heung, Qu, & Chu, 2001; Hsu & Kang, 2007; Kim & Prideaux, 2005; Scott-Halsell, Palakurthi, Dunn, & Saiprasert, 2010; Simon Wong & Gladys Liu, 2011; Sung, Morrison, Hong, & O'Leary, 2001; Uysal, Jurowski, Noe, & McDonald, 1994; Weaver, Weber, & McCleary, 2007). Travel party size is one trip characteristic that has been studied quite extensively in this context. Travel party composition, especially whether children are included or not, is also often investigated. It is argued that these travel party related factors determine the social setting of a trip.

Social psychologists have shown that individuals in different groups show different levels of performance in organizational studies (Ang & O'Connor, 1991) and different learning effects in educational contexts (Hew & Cheung, 2011) due to group interaction processes and group member influence. Most importantly, depending on group size, individuals show different attitudes toward group interaction (Ahmad, 2011; Hew & Cheung, 2011; O'Dell, 1968). Therefore, we can assume that travel group size also influences motivations for social interactions.

Due to cognitive limitations related to maintaining social relationships (Dunbar, 1992), social interactions are looser in large group settings and more intense in smaller groups. This could influence the motivation to process social cues and establish social relationships with a virtual social actor. Accordingly, solitary travellers have the greatest capacity to engage with virtual social actors. Also, it can be assumed that social interactions are more focused on the immediate group if children are present who need attention. We therefore hypothesize the following:

> H2: Solitary travellers and travellers in large social groups are more motivated to process social cues and, thus, feel greater social presence of virtual social actors.
>
> H3: Travellers in groups with children are less motivated to process social cues communicated through technologies.

3 Methodology

3.1 Data Collection

Data for the study were collected from visitors to a national park in the United States. The visitors were invited to listen to interpretive podcast tours on an MP3 player. They completed a pre-test questionnaire about their trip characteristics and demographics before commencing the podcast tours and a post-test questionnaire regarding their experience when they returned the MP3 player. A total of 221 valid responses were collected.

3.2 Measures

Social presence was measured using three items following the scale developed by Kang and Gretzel (2011) which was based on previous studies by Gefen and Straub (2003) and Lee (2002). Table 1 shows the individual items, which were measured using an answer scale ranging from 1 (Not at all) to 5 (A lot). n-Aff was measured with three items based on the need for affiliation items included in Hill's interpersonal oriented scale (Hill, 1987). Similar to social presence, n-Aff was measured with a 5-point Likert Scale (1=strongly disagree to 5=strongly agree).

Table 1. Constructs and Items

Construct		Item
Social Presence	1	I felt cared for in the park even though there was no human guide.
	2	I felt involved with the narrator(s).
	3	I perceived the narrator(s)' messages as being personal
Need-for-Affiliation	1	I think being close to others, listening to them, and relating to them is one of my favourite and most satisfying pastimes.
	2	I would find it very satisfying to be able to form new friendships with whomever I liked.
	3	Just being around others and finding out about them is one of the most interesting things.

Table 2. Scale Properties

Construct/Item	Loadings	Mean	Standard Deviation	Reliability
Social Presence		*3.71*	*0.86*	*0.86*
Social Presence 1	0.76	3.89	0.91	
Social Presence 2	0.90	3.58	0.97	
Social Presence 3	0.81	3.64	1.04	
Need-for-Affiliation		*3.91*	*0.82*	*0.87*
Need-for-Affiliation 1	0.87	3.86	0.96	
Need-for-Affiliation 2	0.88	4.00	0.84	
Need-for-Affiliation 3	0.92	3.84	0.95	

The properties of both scales were investigated using Cronbach Alpha to test their internal consistency and factor analysis to examine their dimensionality. Both scales exhibited satisfactory properties (Table 2). Participants were divided into a high or low n-Aff group based on the scale mean. A small travel group was defined as consisting of 2 or 3 travellers and a big group of 4 or more travellers.

3.3 Sample Characteristics

Over half of the participants were female (56.4%); between 35 and 54 (50.5%) years of age; and having at least a four year university education (56.4%). About 59% of them visited the national park for the first time. Over 40% made a decision to visit the park earlier the same day or the day before while approximately 15.9% planned to visit one to five months earlier. Approximately 22% of the study participants answered that they would stay overnight and 45% responded that they would stay 2 or less hours.

With regard to the size of the travel group, about a third travelled in pairs. Only 2% travelled alone. Over 51% had no children in their travel group; 12.7% of those with children in their group had children under five years old. Table 3 shows the detailed profile.

Table 3. Profile of Respondents

Profile of Respondents	%	Profile of Respondents	%
Gender		Size of travel group	
Female	56.4	Travelled alone	2.3
Male	43.6	2	35.3
		3	18.6
		4	21.3
		5 or more	22.6
Age		Frequency of Visit	
18-24	11.4	First time	58.9
25-34	11.4	2 – 3 times	28.5
35-44	25.5	4 – 5 times	5.9
45-54	25.0	5 times or more	7.2
55-64	18.2		
65 or older	8.6		
Education		Time of Decision to Visit	
High School or less	13.2	Early today	24.5
Some college	21.4	Yesterday	17.3
A degree from a 2 yr. college	9.1	Within past week	14.5
Graduate from 4 yr. university	23.2	1-3 weeks ago	20.5
Some graduate school	7.7	1-5 months ago	15.9
A graduate degree	25.5	More than 6 months ago	7.3
Children in Travel Group		Length of Stay	
Yes	48.9	< 1 hours	1.4
Age of Children		1 – 2 hours	12.3
0-5 years	12.7	2 – 5 hours	45.0
6-10 years	29.4	5 – 24 hours	19.1
11-17 years	24.0	Overnight	22.3

3.4 Manipulation Checks

To make sure that subjects actually listened to the podcast tour and to check whether the technology or podcast quality might have interfered with their ability to experience social presence, a number of additional questions were asked. The answers to these questions indicate that almost all listened to 4 or more of the 6 podcast tour segments offered. Approximately 58% of participants owned an MP3 player and nearly 62% of them had experience with listening to podcasts. The majority (91%) found the MP3 player easy to use and 85% thought it was easy to select the specific podcast segments. Almost all (92%) found the voice on the audio tour pleasant, 93% found the information presented helpful, 91% thought the content was interesting, 82% indicated that they perceived the length of the audio segments to be appropriate and 89% found the amount of information presented to be proper. The majority (88%) was very satisfied with the audio tour.

4 Results and Discussions

This study assessed whether there is a difference between high and low n-Aff groups in perceptions of social presence using a t-test. The result revealed that there is a significant difference between travellers with high and low n-Aff in perceiving social presence when listening to interpretive podcast tours ($t=-3.8,\ p<0.000$). The group with high n-Aff showed perceived greater social presence (mean=4.20) than the group with low n-Aff (mean=2.83) (Figure 1). Therefore, H1 was supported.

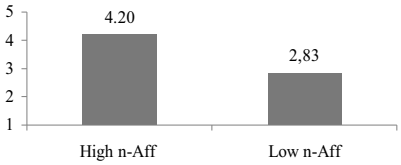


Fig. 1. Mean difference of social presence between high and low n-Aff groups

To see the effect of different travel group size in perceiving social presence, this study conducted again a t-test. Due to the small incidence of solitary travellers (n=5), only the assumptions regarding small and big groups could be tested. The results revealed that there is a difference between small travel groups (two or three: n=119) and big travel group (four or more: n=97) in perceiving social presence, but this difference is only marginally significant ($t=1.55;\ p=0.120$). However, the effect is in the right direction with travellers in bigger travel groups perceiving greater social presence (mean = 3.82) than those in small groups (mean = 3.64) when they listened to the podcast tour. Finally, there is no statistical difference between travel groups with children and those without children in perceptions of social presence ($t=0.22;\ p=0.880$) (Figure 2). Thus, the current study finds only some support for H2 and no support for H3.

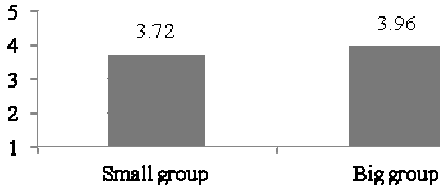


Fig. 2. Mean difference of social presence between small and big travel groups

5 Conclusion

As media richness continues to grow, the potential of experiencing social presence increases. Perceptions of social presence were found by Kang and Gretzel (2011) to have significant positive impacts on tourism experiences and, thus, should be fostered as much as possible. To do so, one needs to better understand influences on social presence beyond the technology design aspects. This paper tested three constructs that influence travellers' motivations to engage in social interactions. The significant findings for n-Aff indicate that social presence is perceived differently by individuals depending on their personality. For individuals with low n-Aff it might be more important to facilitate other forms of presence to enhance the experience. The findings further indicate that more research is needed regarding travel group size and that other variables related to the social setting need to be investigated.

Overall, the study confirms that differences in social presence perceptions exist and have yet to be completely understood. This is important from a practical perspective, especially when designing travel technologies to replace social interactions.

References

Ahmad, K. Z. B. (2011). Group size as a moderator of effect of equity sensitivity on employee job satisfaction. *International Journal of Management* 28(3): 716-729.

Ang, S. & O'Connor, M. (1991). The effect of group interaction processes on performance in time series extrapolation. *International Journal of Forecasting* 7(1): 141-149.

Atkinson, J. W., Heyns, R. W. & Veroff, J. (1954). The effect of experimental arousal of the affiliation motive on thematic apperception. *The Journal of Abnormal and Social Psycholog* 49(3): 405-410.

Baloglu, S. (1997). The relationship between destination images and sociodemographic and trip characteristics of international travellers *Journal of Vacation Marketing* 3(3): 221-233.

Benckendorff, P., Moscardo, G. & Murphy, L. (2006). *Visitor Perceptions of Technology Use in Tourist Attractions*. Paper presented at the Cutting Edge Research in Tourism `06, Surrey, UK.

Biocca, F. (1997). The cyborg's dilemma: Progressive embodiment in virtual enviorments. *Journal of Computer Mediated Communication* 3(2).

Biocca, F., Harms, C. & Burgoon, J. K. (2003). Toward a more robust theory and measure of social presence: a review and suggested criteria. *Presence: Teleoperators and Virtual Environments* 12(5): 456-480.

Brent, E. & Thompson, G. A. (1999). Sociology: modeling social interaction with autonomous agents. *Social Science Computer Review* 17(3): 313-322.

Buss, A. H. (1983). Social rewards and personality *Journal of personality and social psychology* 44(3): 553-563.

Dunbar, R. I. M. (1992). Neocortex size as a constraint on group size in primates. *Journal of Human Evolution* 22(6): 469-493.

Foa, U. G. & Foa, E. B. (1974). *Societal structures of the mind.* Springfield, IL: Charles C. Thomas.

Gefen, D. & Straub, D. (2003). Managing user trust in B2C e-service. *e-Service Journal* 2(2): 7-24.

Heung, V. C. S., Qu, H. & Chu, R. (2001). The relationship between vacation factors and socio-demographic and travelling characteristics: the case of Japanese leisure travellers. *Tourism Management* 22(3): 259-269.

Hew, K. F. & Chueung, W. S. (2011). Higer-level knowledge construction in asyncrhonous online discusssions: An analysis of group size, duration of online dicussion, and student facilitation techniques. *Industrial Science* 39(3): 303-319

Heyns, R., Atkinson, J. W. & Veroff, J. (1958). the effect of experimental arousal of the affiliation motive on thematic apperception. In J. W. Atkinson (Ed.), *Motives in fantasy, action and society*. London: Van Nostrand.

Hill, C. A. (1987). A affiliation motivation: People who need people....But in different ways. *Journal of Personality and social psychology* 52(5): 1008-1018.

Hill, C. A. (1991). Seeking emotional support: The influence of affiliative need and partner warmth. *Journal of Personality and Social Psychology* 60(1): 112-121.

Hill, C. A. (1997). Relationship of expressive and affiliative personality despoisitons to perceptions of social support. *Basic and applied social psychology* 19(2): 133-161.

Hsu, C. H. C. & Kang, S. K. (2007). CHAID-based Segmentation: International Visitors' Trip Characteristics and Perceptions. *Journal of Travel Research* 46(2): 207-216.

Kang, M. & Gretzel, U. (2011). Effect of podcast tours on tourist experiences in a national park. *Tourism Management*. doi:10.1016/j.tourman.2011.05.005.

Kim, S. S. & Prideaux, B. (2005). Marketing implications arising from a comparative study of international pleasure tourist motivations and other travel-related characteristics of visitors to Korea. *Tourism Management* 26(3): 347-357.

Lastovicka, J. L. & Joachimsthaler, E. A. (1988). Improving the Detection of Personality-Behavior Relationships in Consumer Research. *Journal of Consumer Research* 14(4): 583-587.

Lee, K. M. (2002). *Social responses to synthesized speech: theory and Application.* Doctor of Philosophy Stanford University (UMI Microform 3048561).

Lee, K. M. & Nass, C. (2004). The multiple source effect and synthesized speech: doubly disembodied language as a conceptual framework. *Human Communication Research* 30(2) 182-207.

Lombard, M. (2004). Resources for the study of presence: Presence explication Retrieved May 1, 2004, from http://www.temple.edu/isu/frame_explicat.htm

Lombard, M. & Ditton, T. (1997). At the heart of it all: Concept of Presence. *Journal of Computer Mediated Communication* 3(4) http://www.ascusc.org/jcmc/vol3/issue2/lombard.html.

Lowenthal, P. R. (2009). Social presence. In P. Rogers, G. Berg, J. Boettcher, C. Howard, L. Justice and K. Schenk (Eds.), *Encyclopedia of distance education and online learning* (2nd ed., pp. 932–936). Hershey, PA: IGI Global.

Murray, H. A. (1938). *Explorations in personality.* New York: Oxford University Press.

O'Connor, S. C. & Rosenblood, L. K. (1996). Affiliation motivation in everyday experience: A theoretical comparison. *Journal of Personality and Social Psychology* 70(3): 513-522.

O'Dell, T. (2007). Tourist experiences and academic junctures. *Scandinavian Journal of Hospitality and Tourism* 7(1): 34-45.

Reysen, S., Lloyd, J. D., Katzarska-Miller, I., Lemker, B. M., Foss, R. L. (2010). Intragroup status and social presence in online fan groups. *Computer in Human Behavior* 26(6): 1314-1317.

Scott-Halsell, S. A., Palakurthi, R., Dunn, G. & Saiprasert, W. (2010). Trip Characteristics of Casino and Racino Visitors In A Midwestern State. *UNLV Gaming Research & Review Journal* 14(1): 27-35.

Short, J., Williams, E. & Christie, B. (1976). *The social psychology of telecommunications*. London: Wiley.

Simon Wong, C. K. & Gladys Liu, F. C. (2011). A study of pre-trip use of travel guidebooks by leisure travelers. *Tourism Management* 32(3): 616-628.

Skalski, P. (2004). *Interactive social agent and social presence effects on informationprocessing and persuasion.* Doctoral of Philosophy, Michigan State University, Lancing, MI. (AAT 3159006).

Slater, M. & Usoh, M. (1993). Represntations systems, perceptua position, and presence in immersive virtual enviroments. *Presence: Teleoperators and Virtual Enviroment* 2(3): 221-233.

Stamboulis, Y. & Skyannis, P. (2003). Innovation strategies and technology for experience based tourism *Tourism Managemen,* 24(1): 35-43.

Sung, H. H., Morrison, A. M., Hong, G.-S. & O'Leary, J. T. (2001). The Effects of Household and Trip Characteristics on Trip Types: A Consumer Behavioral Approach for Segmenting the U.S. Domestic Leisure Travel Market *Journal of Hospitality & Tourism Research* 25(1): 46-68.

Uysal, M., Jurowski, C., Noe, F. P. & McDonald, C. D. (1994). Environmental attitude by trip and visitor characteristics : US Virgin Islands National Park. *Tourism Management* 15(4): 284-294.

Veroff, J. & Veroff, J. B. (1980). *Social incentives: A life-span developmenal approach*. New York: Academic press.

Villani, K. E. A. & Wind, Y. (1975). On the Usage of "Modified" Personality Trait Measures in Consumer Research. *Journal of Consumer Research, 2*(3), 223-228.

Weaver, P. A., Weber, K. & McCleary, K. W. (2007). Destination Evaluation: The Role of Previous Travel Experience and Trip Characteristics. *Journal of Travel Research* 45(3): 333-344.

Whitelock, D. Romano, D. M., Jelfs, A. & Brna, P. (2000). Perfect presence. What does this mean for the design of virtual learning environments? *Education and Information technologies* 6(4): 277-289.

Weisberg, J., Te'eni, D. & Arman, L. (2011). Past purchase and intention to purchase in e-commerce: The mediation of social presence and trust. *Internet Research* 21(1): 82-96.

Witmer, B. G. & Singer, M. J. (1998). Measuring presence in virtual environments: a presence questionnaire. *Presence* 7(3): 225-240.

Factors influencing Travellers' e-Ratings and e-Complaints about Hotel Services: Insights from an Italian Tourism Destination

Giacomo Del Chiappa[a] , and
Stefano Dall'Aglio[b]

[a] Department of Economics, Business and Law
University of Sassari and CRENoS, Italy
gdelchiappa@uniss.it

Stefano Dall'Aglio
[b] Econstat and Luiss University, Italy
stefano.dallaglio@luiss.it

Abstract

The Internet, and in particular User Generated Content (UGC), has provided new ways through which consumers dissatisfied with hotel services can air their complaints. This research applied content analysis to one of the numerous hotel review websites available on the Internet (booking.com) to study the overall ratings given by travellers and the words used in their comments, in order to identify the most frequent complaints. In particular, this paper aims at investigating how consistent customer ratings and their complaint behaviour are with their socio-demographic characteristics (i.e., gender and culture/country of origin), with whom they are travelling (e.g., alone, with family, etc.), with the time of year when their stay took place (winter, mid-season and summer) and with several hotels operational/business indicators (i.e., hotel class and size). Implications for hotel managers are discussed and suggestions for further research are given.

Keywords: User generated content, e-complaints, socio-demographic characteristics, hotel operation/business indicators, culture, Italy.

1 Introduction

UGC has introduced a new venue for both positive and negative word-of-mouth recommendations, eWOM, defined as "all informal communications directed at consumers through internet-based technology related to the usage or characteristics of particular goods and services, or their sellers" (Livtin et al., 2008). Prior research found that online reviews and recommendations have a higher credibility than traditional sources of tourist information (Akehurst, 2009; Gretzel & Yoo, 2008).

Several studies shed light on the significant influence exerted by UGC and Travel 2.0 applications in stimulating travel, over the actual planning process and during the post-travel phase (Gretzel & Yoo, 2008). Sometimes they can even induce tourists to alter their decisions after obtaining further information online (Del Chiappa, 2011b).

In the post-travel phase, tourists can use UGC to voice failures in service and make complaints (Au et al., 2010; Sparks & Browning, 2010). Customers who complain give businesses the opportunity to solve a problem, leading, sometimes, to a higher overall satisfaction than before the failure in service occurred (Maurer & Schaich, 2010). Consequently, Web 2.0 technologies have become an important tool for

hospitality marketers to manage complaints (Del Chiappa, 2011a; Maurer & Schaich, 2011), although, unfortunately, it would seem that several hospitality companies take no notice of e-complaints. For example, Del Chiappa (2011a) found that Italian hotel staff rarely replies to online reviews (only 23.5% do so). It is important to remember that UGC is just a new way for consumers to voice their opinions, and that these should be handled by hotels in the same way as traditional complaints were in the past (Tyrrel & Woods, 2004).

Research on e-rating and e-complaints within the hospitality sector is still in its early stages, and the aim of this study is to broaden our understanding of this subject. In particular, it investigates whether any significant differences exist relating to tourists' socio-demographic characteristics (i.e., gender and culture/country of origin), with whom they are travelling (alone, with family, etc.), the time of year when their stay took place (winter, mid-season and summer) and the characteristics of the hotels they are assessing (i.e., hotel class and size). To our best knowledge, no published study has been conducted in Italy on the topic of e-complaints and because of this, we decided to carry out our study by focusing on 731 online reviews in which tourists expressed their experience of 18 hotels in Taormina, a famous tourism destination in Sicily (Italy).

2 E-complaining Management: Literature Review

Jacoby and Jaccard (1981, p. 6) defined complaining behaviour as "an action taken by an individual that involves communicating something negative regarding a product or service". According to Singh (1988), consumer complaint behaviour (CCB) fits into two categories, that is, behavioural (i.e., action) and non-behavioural (i.e., no action). The different types of action can be public (i.e., seeking a refund from the company) or private (word-of-mouth to friends and relatives). Albrecht and Zemke (1996) found that in general only 5% of disgruntled customers voice their complaints.

UGC has certainly provided consumers with a new and easy way to air a grievance. Travel 2.0 applications allow tourists to express their dissatisfaction to the entire world without any time and/or psychological constraints, therefore reducing the likelihood that other people will choose the hotels they complained about. As the internet is increasingly used to look up information and to book accommodation, it becomes important for hotel managers to analyze and manage the online brand reputation of their hotels. "A negative online review is an expression of dissatisfaction that is posted with the possible aims of venting emotions, engaging with others, and sharing information" (Sparks & Browing, 2010, p. 799). Previous research has also shown that reviews including comments about negative experiences do not necessarily imply an overall negative assessment (Jiang et al., 2010).

Online consumer feedback is essential for recovering and improving services (Bronner & de Hoog, 2011; Sparks & Browning, 2010; Stringam & Gerdes, 2010), so that hospitality marketers can achieve higher customer satisfaction/commitment (Tax et al., 1998), thereby preventing customers from switching and, in turn, increasing customer repurchase intention, customer retention and loyalty (Larivet & Brouard, 2010; Stauss & Schoeler, 2004; Zheng et al., 2009), and hotel profitability. Prior research has shown that it is more cost effective to retain existing customers than to

invest financially in marketing actions to find new ones (Harrison-Walker, 2001; Shea et al., 2004). As suggested by the outcome of complaints voiced in traditional ways (Yavas et al., 2004), it is essential for hospitality managers to pay attention to e-complaints and handle them as soon as possible. For this reason, hotels must monitor online comments regularly and respond to each of them (Lee & Lu, 2004). By monitoring e-complaints, hotels can discover the main weakness in the service they provide and can effectively focus their limited resources on these (Crotts et al., 2009).

Previous studies on complaint management have highlighted that CCB is influenced by the socio-demographic characteristics of customers, such as gender, age and level of education (Han et al., 1995; Jacoby & Jaccard, 1981; Richins, 1983). Maurer and Schaich (2011) found significant gender differences in complaint behaviour, with men complaining more than women, while they did not find significant differences based on age. Other researchers found significant differences in CCB according to the reasons for travelling. For example, Lewis and McCann (2004) found that business travellers are more likely than leisure travellers to mention slow restaurant service, access to business facilities, inefficient staff and a poor variety of food. As in other sectors (Liu et al., 2001), several studies found that cultural background also influences CCB in the hospitality sector (Au et al., 2010; Ngai et al., 2007). Park and Letho (2008) reported that CCB is influenced by the travel party (spouse, children, spouse and children, etc.). Further complaint behaviour can be influenced by psychographic characteristics (Morganosky & Bucklev, 1987), situation-specific attributes (Stauss & Seidel, 2004) and operational/business indicators (hotel class, average daily rate, popularity index, etc.) (Jeong & Jeon, 2008).

Stringam and Gerdes (2010) found the hotel overall cleanliness to be a frequent concern in guests' online comments. Maurer and Schaich (2011) found that features relating to the room (room cleanliness, bathroom cleanliness, room size, etc.) have the highest number of failures, followed by complaints about the overall hotel appearance and guest services (competence and friendliness). Lee and Hu (2004) found that the most frequent complaints included "service provided is not as agreed" (18.02%), "service declined in quality" (16.67%), "rude customer service representatives (14.41%), "service never provided" (13.06%) and "over-charging" (12.16%). Lewis and McCann (2004) have shown that the most common problem mentioned by guests was poor restaurant service (61.7%), followed by inefficient staff for business guests (60.3%), failures in service for food and drink, and quality of rooms.

3 Methodology

For the purpose of this study, 731 online reviews were sourced during 2010 from an online travel agency (OTA) with booking and rating/review functions (booking.com). A selection was made of 18 hotels of different sizes located in a famous Italian tourism destination (Taormina – Sicily). In order to reduce any location-related bias, all the hotels were chosen at a similar distance from the seaside (less than 1.5 Km). Booking.com only accepts hotel reviews from travellers who have made a reservation for that hotel and paid for it on their website. Therefore, compared to other research (Jeong & Jeon, 2008; Maurer & Schaich, 2011; Zheng et al., 2009), we were able to reduce/overcome the risk of building our analysis on online reviews posted by competitors.

On booking.com, users can assess their overall satisfaction with the hotel services by using a 10-point Likert Scale applied to a list of six attributes (cleanliness, comfort, location, services, staff and value for money). However, for each review, booking.com only allows viewers to see the overall assessment, calculated as the medium of the rating that consumers gave for each hotel attribute. Additionally, people are allowed to post both positive and negative comments on the site.

Each review was read and basic open-coding was carried out (Charmaz, 2006), using grounded theory approach and keyword analysis. It emerged that all the negative comments were trigged by, or targeted at, specific features internal or external to the hotel. In particular we considered the following: "rating and promises" (i.e., the coherence with the classification rating, expected services not delivered), "location and surroundings" (i.e., position of the hotel, view, external environment, etc.), "access and parking" (i.e., accessibility, parking facilities, etc.), "design, look and feel" (i.e., shape, decoration, atmosphere, general maintenance, etc.), "staff" (i.e., empathy, competence, responsiveness, attitude, etc.), "organization" (i.e., general organization, room assignment, opening hours of guest services, etc.), "room experience" (i.e., comfort, cleanliness, space, equipment, air conditioning, etc.), "bathroom" (shower, towels, courtesy products, etc.), "breakfast" (i.e., quality, quantity, variety, etc.), "food and drink" (i.e., quality/quantity of food and drink in the restaurant and/or bar), "facilities and services" (i.e., common spaces, garden, swimming pool, beach, shuttle services, internet, etc.) and, lastly, "price and quality-price ratio".

The reviews were coded manually to extract the negative feelings expressed in each of them. It was then possible to calculate the number of negative comments and corresponding percentage for each of the 12 features above. If a guest complained about more than one aspect for the same feature, only one complaint was counted. Overall, 451 e-complaints were identified. To enhance external validity and ensure the reliability of human coding, the study used a team approach to analyze the data obtained from the reviews. Initially, one researcher read the reviews and carried out basic open-coding. After this, the initial codes were reviewed by the research team, and an independent person was brought in to revise the coding and decide whether he/she agreed with the codes. Whenever the research team and the independent reviewer disagreed about the coding adopted, they discussed the point until an agreement was reached, and the final coding was then undertaken.

Data was coded and analyzed using SPSS (version 17.0), then a series of ANOVA and Chi-square tests were applied, where appropriate, to identify whether there were any significant differences in evaluations based on the guests' socio-demographic characteristics (i.e., gender and culture/country of origin), with whom they were travelling (alone, with family, etc.), the time of year of when their stay took place (winter, mid-season and summer) and on the hotel's operational/business indicators (i.e., hotel class and size).

4 Results

4.1 How Do Guests Assess the Overall Satisfaction of Hotel Services?

As Table 1 shows, the overall assessment tourists gave about hotel services differs significantly when considering gender (male, female and "unknown) ($F.=5.791$; $p<0.01$). In particular, it could be argued that people who reveal their gender give higher scores when making complaints because of their fear of being identified.

Table 1. Overall assessment and gender (with "unknown"): ANOVA test

Gender	N	Mean	Standard deviation	F.	Sig.
Male	362	7.555	1.6965	5.791	.003
Female	186	7.782	1.7495		
Unknown	183	7.179	1.7496		
Total	731	7.519	1.7347		

Table 2 shows a series of ANOVA tests that we ran to determine if significant differences are to be found in the assessment that tourists give, based on their cultural background, the party with whom they are travelling, the time of year of their stay, as well as the class and size of hotels where they stayed.

Domestic and international guests differ, although slightly, in their assessment ($F.=3.920$, $p<0.05$). In particular, foreign travellers give better ratings (mean=7.368; S.D.=1.8278) than Italian guests (Mean=7.625; S.D.=1.6601).

When examining the travel party, organized groups/groups of friends give the best overall assessment (mean: 7.999; S.D: 1.577) among the segments considered, while families with older children/teenagers give the worst (mean: 7.243; S.D: 1.675). However, there are no significant differences in the way the different segments make their assessment ($F.= 1.785$; $p>0.05$).

The overall assessment differs significantly based on the time of year when the hotel stay took place ($F.=6.291$; $p<0.01$), with mid-season (April-May, October-November) scoring higher than summer. It could be argued that this occurs because during the mid-season there are fewer guests staying in the hotel resulting in less "organization stress" and better customer service. Hospitality services are both seasonal and perishable, which raises the marketing stress level for hospitality managers and staff (Rao & Singhapakdi, 1997). Surprisingly, the overall assessment is higher in winter than in summer. This could be explained by arguing that in winter people enjoy relaxing in uncrowded hotels.

The overall assessment is higher for 3 star hotels (mean=7.595; S.D.=1.7807) than for 4 star hotels (mean=7.477; S.D.=1.7095), which does not confirm previous research, where 4 star hotels were given a significantly more positive assessment than 3 star hotels (Jeong & Jeon, 2008). This could be because of unmet promises in higher hotel categories and/or because of the positive rating given by guests to the personal touch and organization offered by smaller, often family-run, hotels. However, the difference between the way guests assess 3 and 4 star hotels does not differ significantly ($F.=.771$; $p>0.05$).

Table 2. Country of origin, travel party, time of year of stay, hotel category and size, and their influence on overall assessment: ANOVA test

	Overall assessment				
	N=731	Mean	Standard deviation	F.	Sig.
Country of origin					
Italian	302	7.368	1.8278	3.920	.048
International	429	7.625	1.6601		
Travel party					
Organized groups/groups of friends	73	7.999	1.577	1.785	.113
Single travellers	84	7.364	1.826		
Young couples	271	7.456	1.880		
Family with babies/small children	66	7.395	1.714		
Family with older children/teenagers	54	7.243	1.675		
Mature couples	183	7.616	1.520		
Time of year of stay					
Summer	439	7.348	1.8249	6.291	.002
Mid-season	234	7.842	1.5405		
Winter	58	7.509	1.6104		
Total	731	7.519	1.7347		
Hotel class					
3 stars	259	7.595	1.7807	.771	.380
4 stars	472	7.477	1.7095		
Hotel size (no. of rooms)					
< 50	291	7.611	1.7761	1.753	.174
50 – 150	250	7.352	1.7414		
> 150	190	7.596	1.6528		

Lastly, mid-size hotels (those with 50 to 150 rooms) seem more exposed to low ratings and more complaints. This could be explained in several ways. For example, it could be argued that mid-size hotels lose the "human touch" (found in small hotels), while not benefitting from the economies of scale and/or scope of large hotels, resulting in a comparably worse standard in the delivery of services. However, also in this case, the way guests assess hotels based on their size does not differ significantly ($F.=1.753$; $p>0.05$).

4.2 How Do Guests Behave when Complaining Online about Hotels?

Out of the total of the 731 reviews we analyzed, 415 included negative elements. Tables 3 to 6 record the most frequently mentioned complaints about hotel services (grey box) made by travellers, based on their gender, cultural background, the party with whom they were travelling, the hotel class/size and time of year when they stayed at the hotel. They also show if there are any significant differences in complaining behaviour for these variables. Table 3 shows that women complain more than men, a finding that does not confirm prior research (Maurer & Schaich, 2011). In particular, women complain significantly more than men for the "access & parking"

service (χ^2=4.315, p<0.05). The following services/aspects, "breakfast", "facilities and services", "location & surroundings", "staff", "bathroom", "price/quality ratio", receive the most complaints from international guests, while Italian travellers complain more about the other aspects. In particular, international guests complain about "breakfast" and "location & surroundings" significantly more than Italian travellers (breakfast: χ^2=14.314, p<0.01 – location & surroundings: χ^2=22.222, p<0.01). On the whole, these findings confirm prior research, which found that culture influences complaining behaviour (Au et al., 2010; Ngai et al., 2007).

Table 3. E-complaints, gender and cultural background: chi-square test
(* significant at 0.05 level, ** significant at 0.01 level)

Attributes	Gender			Country of origin of guests		
	Male %	Female %	χ^2	It.	Int.	χ^2
Room experience	15.2	19.9	1.942	17.9	16.6	.221
Organization	14.1	15.1	.093	14.9	13.1	.508
Breakfast	9.4	13.4	2.096	5.6	14.5	14.314**
Access & parking	7.5	12.9	4.315*	9.3	9.3	.001
Facilities & services	7.7	8.1	.018	7	9.1	1.075
Location & surroundings	7.2	10.2	1.499	2.3	11.9	22.222**
Staff	6.9	7.5	.072	7	8.2	.364
Bathroom	6.9	7.0	.001	6.6	7.2	.099
Design, look & feel	5.2	4.8	.043	6.3	6.1	.016
Rating & promises	5	5.4	.041	5.3	4.9	.060
Price/quality ratio	3.6	3.8	.010	3.3	4.4	.581
Food & drink	3.9	3.2	.144	4.3	3.5	.314
It.=Italian guests Int.=International guests						

According to Table 4, tourists complain differently depending on whom they are travelling with. For example, "facilities and services" is the aspect of most concern for families with babies and small children (19.7%), while families with older children and teenagers complain more frequently than the other groups about "access and parking" (16.7). However, there are significant differences only for "facilities and services", with families with babies/small children making complaints more frequently than the other categories (χ^2=13.686, p<0.05).

Table 5 shows that significant differences in complaining behaviour can be found for both hotel class and size. For example, 4 star hotels receive complaints more frequently than 3 star hotels for "organization" (χ^2=21.695, p<0.01) and "staff" (χ^2=6.608, p<0.05). Moreover, tourists complain significantly more for 3 star hotels than for 4 star hotels for "access and parking" (χ^2=13.307, p<0.01) and "bathroom" (χ^2=20.538, p<0.01).

Table 4. E-complaints and travel party: chi-square test
(* significant at 0.05 level, ** significant at 0.01 level)

Attributes	Travel party						
	G %	YC %	FBK %	FCT %	S %	MC %	χ^2
Room experience	9.6	19.2	16.7	18.5	22.6	14.2	6.709
Organization	15.1	14	15.2	11.1	16.7	12	1.604
Breakfast	1.4	11.8	7.6	13	11.9	13.1	9.118
Access & parking	12.3	8.5	7.6	16.7	9.5	7.7	5.307
Facilities & services	8.2	7.4	19.7	3.7	7.1	7.1	13.686*
Location & surroundings	8.2	10.7	4.5	5.6	4.8	7.1	5.634
Staff	6.8	8.5	6.1	9.3	6	7.7	1.110
Bathroom	2.7	8.9	4.5	3.7	6	8.2	5.542
Design, look & feel	6.8	8.9	6.1	7.4	3.6	2.7	8.313
Rating & promises	1.4	6.6	6.1	3.7	3.6	4.9	4.219
Price/quality ratio	-	4.4	6.1	7.4	3.6	3.3	5.866
Food & drink	4.1	4.4	1.5	1.9	4.8	3.8	2.010
G=organized groups/groups of friends S=single travellers YC=young couples			FBK=families with babies/small children FCT=families with older children/teenagers MC=mature couples				

Table 5. E-complaints, hotel class and hotel size: chi-square test
(* significant at 0.05 level, ** significant at 0.01 level)

Attributes	Hotel class			Hotel size (no. of rooms)			
	3 stars %	4 stars %	χ^2	< 50 %	50-150 %	>150 %	χ^2
Room experience	15.1	18.2	1.180	15.1	22.8	12.6	9.211**
Organization	5.8	18.2	21.695**	6.5	17.2	20.5	22.565**
Breakfast	11.6	10.4	.251	11.7	8.8	12.1	1.609
Access & parking	14.7	6.4	13.307**	13.4	9.6	2.6	15.845**
Facilities & services	6.9	8.9	.843	6.9	9.6	8.4	1.343
Location & surroundings	6.6	8.7	1.032	7.6	6	11.1	3.886
Staff	4.2	9.5	6.608*	4.5	10.8	8.4	7.833*
Bathroom	12.7	3.8	20.538**	11.7	3.2	4.7	16.898**
Design, look & feel	5.4	6.6	.391	5.2	7.2	6.3	.985
Rating & promises	3.5	5.9	2.101	4.1	7.2	3.7	3.662
Price/quality ratio	3.1	4.4	.812	3.1	2	7.9	10.816**
Food & drink	1.9	4.9	3.930*	2.4	3.6	6.3	4.826

Small hotels (those with less than 50 rooms) receive significantly more complaints for "access & parking" (13.4%) and "bathroom" (11.7%) than medium-sized and large hotels (access & parking: χ^2=15.845, p<0.01 – bathroom: χ^2=16.898, p<0.01). Large hotels receive significantly more complaints that the others for "organization" (χ^2=22.565, p<0.01), "price, quality/price ratio" (χ^2=10.816, p<0.01). Interestingly, medium-sized hotels (50 to 150 rooms) receive the most frequent complaints for "staff" (χ^2=7.833, p<0.05). This could be explained by their difficulty in reaching the "minimum efficient organizational dimension".

Finally, Table 6 shows that hotel stays during mid-season always receive fewer complaints than those visited in winter or summer. Significant differences in complaining behaviour can be found based on the time of year of the stay.

Table 6. E-complaints and time of year of stay: chi-square test
(* significant at 0.05 level, ** significant at 0.01 level)

Attributes	Time of year of stay			
	Winter %	Mid-season %	Summer %	χ^2
Room experience	24.1	12	18.9	7.389*
Organization	13.8	9.4	16.2	5.878
Breakfast	15.5	9	11.2	2.208
Access & parking	15.5	7.7	9.3	3.375
Facilities & services	1.7	6	10.3	7.205*
Location & surroundings	3.4	8.1	8.4	1.756
Staff	3.4	3.8	10.3	10.341**
Bathroom	8.6	7.7	6.4	.699
Design, look & feel	3.4	4.7	7.3	2.570
Rating & promises	6.9	2.1	6.4	6.156*
Price/quality ratio	1.7	1.7	5.5	6.489*
Food & drink	5.2	2.1	4.6	2.733

For example, hotels receive significantly more complaints in summer than in any other period for "facilities & services" (χ^2=7.205, p<0.05), "price, quality/price ratio" (χ^2=6.489, p<0.05) and "staff" (χ^2=10.341, p<0.01). This could be explained by the higher stress experienced by staff during the summer, due to the greater number of guests, therefore resulting in more complaints. On the contrary, in the winter, hotels receive more complaints than in any other period for "room experience" (χ^2=7.389, p<0.05) and "rating & promises" (χ^2=6.156, p<0.05).

5 Conclusion

Research on e-complaints management has not been studied extensively despite its significant importance to the hospitality industry. In particular, to the authors' best knowledge, no paper has been published giving these insights into the Italian hospitality sector. The present study was therefore carried out to analyze ratings and complaints posted in 731 online reviews on 18 hotels in Taormina, a famous tourism destination in Sicily (Italy). Of the 731 reviews we examined, 451 reported negative elements.

Our findings provide an interesting insight into how the rating and complaining behaviour of tourists over the internet can be differentiated, according to the following six variables, that is, the tourists' gender, cultural background, whom they travel with, time of year of their stay, as well as hotel class and size. When rating behaviour was considered, our study found that there were significant differences depending on "cultural background" and "time of year of stay". Beyond that, the study provided an insight into the most frequently made complaints for the 12 hotel features we investigated, based on the aforementioned variables. Results revealed that CCB shows significant differentiation for all the six variables considered. When gender is taken into account, women were seen to complain significantly more than men for "access & parking" services. International guests complain significantly more than Italian travellers about "breakfast" and "location & surroundings". Regarding the travel party, families with babies/small children were found to complain about "facilities and services" significantly more frequently than the other categories. Hotels receive significantly more complaints in summer than in any other period about "facilities & services", "price, quality/price ratio" and "staff". When the hotel category is considered, 4 star hotels receive more frequent complaints about "organization" than 3 star hotels. On top of this, tourists complain significantly more about "access and parking" and "bathroom" for 3 star hotels than for 4 star hotels. Large hotels receive significantly more complaints than medium-sized or small hotels for "organization" and "price, quality/price ratio". Interestingly, medium-sized hotels receive the most frequent complaints for "staff".

These conclusions are relevant both for researchers and hospitality managers. From a theoretical perspective, the study adds to the growing literature in the field of e-rating and e-complaints. In particular, our findings confirm previous research showing that consumers are often dissatisfied with basic services, including the failure to deliver a service or inappropriate behaviour of personnel (Lee & Hu, 2004). Further, they highlight that CCB can be significantly influenced by the socio-demographic background of guests, whom they travel with and time of year when their stay takes place, as well as by several business/operation indicators. Interestingly, our results confirm pioneering studies analyzing the influence that culture exerts on the complaining behaviour of tourists (Au et al., 2010). Regarding the latter, our study contributed to the scientific debate by providing insights into Italian culture, thereby addressing an area that, to our best knowledge, so far has never been investigated. For marketing practices, these results will help hotel managers to direct their limited resources on the features of the hotel and the services they provide that are most frequently mentioned in online reviews and which reflect the characteristics of their guests and/or of the hotel businesses they manage.

Although this study helps fill a gap in existing knowledge in literature and does offer some implications for practitioners, its scope was restricted. Firstly, our study was based on a single online travel site and was limited to 18 hotels in one tourism destination. Therefore, our findings cannot be generalized to other hotels operating in other tourism destinations. It could be argued that further research in different destinations is needed to establish if these findings can be confirmed and/or whether the study should take a location factor into consideration.

Secondly, when analyzing CCB, we only considered how often complaints were made for each hotel feature, whilst not considering at all the extent to which each feature can influence customers' overall satisfaction. According to importance-performance modelling (Martilla & James, 1977), it could be argued that a relatively high number of complaints for a relatively non-significant hotel feature would not be a priority for hotel managers (Zainol et al., 2010). Further research is needed to recognize the most important factors that exert most influence over the overall satisfaction of guests.

References

Akehurst, G. (2009). User generated content: the use of blogs for tourism organizations and tourism consumers. *Service Business* 3(1): 51-61.

Albrecht, K. & Zemke, R. (1996). *Service America! Doing Business in the New Economy*. Homewood, IL: Dow Jones-Irwin.

Au, N., Law, R. & Buhalis D. (2010). The impact of Culture on eComplaints: Evidence from Chinese Consumers in Hospitality Organizations. In U. Gretzel, R. Law and M Fuchs (Eds.), *Information and Communication Technologies in Tourism 2010*, 285-296. Vienna, Austria: Springer.

Bronner, F. & de Hoog, R. (2011). Vacationers and eWOM: Who Posts, and Why, Where, and What? *Journal of Travel Research* 50(1): 15-26.

Charmaz, K. (2006). *Constructing grounded theory*. London: Sage.

Crotts, J. C., Mason, P. R. & Davis, B. (2011). Measuring Guest Satisfaction and Competitive Position in The Hospitality and Tourism Industry. *Journal of Travel Research* 48(2): 139-151.

Del Chiappa, G. (2011a). Awareness, use and attitude of the tourism *enterprises* towards the web 2.0: and empirical analysis on the Italian Hospitality sector. In U. Collesei and J.C. Andreani (Eds), Proceedings of the *X International Conference, "Marketing Trends*", 21-23 January 2011, *ESCP-EAP*, Paris.

Del Chiappa, G. (2011b). Trustworthiness of Travel 2.0 applications and their influence on tourist behaviour: an empirical investigation in Italy. In R. Law, M. Fuchs and F. Ricci (Eds.), *Information and Communication Technologies in Tourism 2011,* 499-511. Vienna, Austria: Springer.

Gretzel, U. & Yoo K.H. (2008). Use and Impact of Online travel Reviews. In P. O'Connor, W. Höpken and U. Gretzel (Eds.), *Information and Communication Technologies in Tourism 2008*, 35-46. Vienna, Austria: Springer.

Han, S., Keng, K.A. & Richmond, D. (1995). Determinants of Consumer Complaint Behavior: A study of Singapore Consumers. *Journal of International Consumer Marketing* 15(5): 283-289.

Harrison-Walker, L.J. (2001). E-complaining. A content analysis of an Internet complaint forum. *Journal of Service Marketing* 15(5): 397-412.

Jacoby, J. & Jaccard, J.J. (1981). The sources, meaning and validity of customer comment behaviour: A psychological analysis. *Journal of Retailing* 57(3): 4-24.

Jeong, M. & Jeon, M.M. (2008). Customer Reviews of Hotel Experiences through Consumer Generated Media (CGM). *Journal of Hospitality & Leisure Marketing* 17(1-2): 121-138.

Jiang, J., Gretzel, U. & Law, R. (2010). Do negative Experiences Always Lead to Dissatisfaction? – Testing attribution Theory in the Context of Online Travel Reviews. In U. Gretzel, R. Law and M. Fuchs (Eds.), *Information and Communication Technologies in Tourism 2010*, 285-296. Vienna, Austria: Springer.

Larivet, S. & Brouard, F. (2010). Complaints are a firm's best friend. *Journal of Strategic Marketing* 18(7): 537-551.

Lee, C. C. & Hu, C. (2004). Analyzing Hotel Customers' E-complaints from an Internet Complaint Forum. *Journal of Travel and Tourism Marketing* 17(2-3): 167-181.

Lewis, B. R. & McCann, P. (2004). Service failure and recovery: evidence from the hotel industry. *International Journal of Contemporary Hospitality Management* 6(1): 6-17.

Liu, T., Warden, C. A. & Huang, C. (2001). Fatal Service Failures Across Cultures. *Journal of Hospitality & Leisure Marketing* 8(1-2): 93-111.

Livtin, S.W, Goldsmith, R.E. & Pan B. (2008). Electronic word-of-mouth in hospitality and tourism management. *Tourism Management* 29(3): 458-468.

Martilla, J. & James, J. (1977). Importance performance analysis. *Journal of Marketing* 41(1): 77-79.

Maurer, C. & Schaich, S. (2011). Online Customer reviews Used as Complaint Management Tool. In R. Law, M. Fuchs and F. Ricci (Eds.), *Information and Communication Technologies in Tourism 2011,* 499-511. Vienna, Austria: Springer.

Mohsin, A. & Lockyer, T. (2010). Customer perceptions of service quality in luxury hotels in New Delhi, India: an exploratory study. *International Journal of Contemporary Hospitality Management* 22(2): 160-173.

Morganosky, M.A. & Buckle, H.M. (1987). Complaint Behaviour: Analysis by Demographics, Lifestyle, and Consumer Values. *Advances in Consumer Research* 14: 223-226.

Ngai, E. W. T., Heung, V. C. S. Wong, Y. H. & Chan, K. Y. (2007). Consumer complaint behaviour of Asians and non-Asians about hotel services: An Empirical Analysis. *European Journal of Marketing* 41(11-12): 1375-1391.

Park, O., Lehto X. & Park. J. (2006). Service failures and complaints in the family travel market. a justice approach. *Journal of Services Marketing* 22(7): 520-532.

Rao, C. P. & Singhapakdi, A. (1997). Marketing ethics: A comparison between services and other marketing professionals. *The Journal of Service Marketing* 11(6): 409-426.

Richins, M. L. (1983). Negative Word-of-Mouth by Dissatisfied Consumers: A Pilot Study. *Journal of Marketing* 47: 68-78.

Shea, L., Enghagen, L. & Khullar, A. (2004). Internet Diffusion of an E-Complaint: A Content Analysis of Unsolicited Responses. *Journal of Travel and Tourism Marketing* 17(2-3): 145-165.

Singh, J. (1988). Consumer Complaint Intentions and Behavior: Definitional and Taxonomical Issues. *Journal of Marketing* 52(1): 93-107.

Sparks, B.A. & Browning, V. (2010). Complaining in Cyberspace: The Motives and Forms of Hotel Guests' Complaints Online. *Journal of Hospitality Marketing & Management* 19(7): 797-818.

Stauss, B. & Schoeler, A. (2004). Complaint management profitability: what do complaint managers know? *Managing Service Quality* 14(2-3): 147-156.

Stauss, B. & Seidel, W. (2004). *Complaint Management. The Heart of CRM.* Ohio: Thomson.

Stringam, B.B. & Gerdes J. (2010). An Analysis of Word-of-Mouse Ratings and Guest Comments of Online Hotel Distribution Sites. *Journal of Hospitality Marketing & Management* 19(7): 773-796.

Tax, S. S., Brown, S. W. & Chandrashekaran, M. (1998). Customer Evaluations of Service Complaint Experiences: Implications for Relationship Marketing. *Journal of Marketing* 62: 60-76.

Tyrrell, B. & Woods R. (2004). E-Complaints: Lessons to Be Learned from the Service Recovery Literature. *Journal of Travel and Tourism Marketing* 17(2-3): 183-190.

Yavas, U., Karatepe, O. M. & Avci, T. (2004). Customer Complaints and Organizational Responses: A study of Hotel Guests in Northern Cyprus. *Journal of Hospitality & Leisure Marketing* 11(2-3): 31-46.

Zainol, N.A., Lockwood, A. & Kutsch E. (2010). Relating the zone of tolerance to service failure in the hospitality industry. *Journal of Travel and Tourism Marketing* 27(3): 324-333.

Zheng, T. Youn, H. & Kincaid, C.S. (2009). An Analysis of Customers' E-Complaints for Luxury Resort Properties, *Journal of Hospitality Marketing & Management* 18(7): 718-729.

Classification of Customer Reviews based on Sentiment Analysis

Dietmar Gräbner[a],
Markus Zanker[b],
Günther Fliedl[b], and
Matthias Fuchs[c]

[a] econob Informationsdienstleistungs GmbH, Austria
dietmar.graebner@econob.com

[b] Institute of Applied Informatics
Alpen-Adria-Universität Klagenfurt, Austria
{markus.zanker@uni-klu.ac.at, guenther.fliedl}@uni-klu.ac.at

[c] The European Tourism Research Institute (ETOUR)
Mid-Sweden University, Sweden
matthias.fuchs@miun.se

Abstract

In this paper we propose a system that performs the classification of customer reviews of hotels by means of a sentiment analysis. We elaborate on a process to extract a domain-specific lexicon of semantically relevant words based on a given corpus (Scharl et al., 2003; Pak & Paroubek, 2010). The resulting lexicon backs the sentiment analysis for generating a classification of the reviews. The evaluation of the classification on test data shows that the proposed system performs better compared to a predefined baseline: if a customer review is classified as *good* or *bad* the classification is correct with a probability of about 90%.

Keywords: Web 2.0, sentiment analysis, customer reviews, classification

1 Introduction

The degree of interactivity established by Web 2.0 applications shifted the priority of the Internet from an information source to an opinion source (Dippelreiter et al., 2007; Schmalegger & Carson, 2008). Every piece of information, whether it is a product offered in an online store or a post in your social network of choice, can be commented or rated in some way (Litvin et al., 2008; Xiang et al., 2010). Surveys show that the majority of Internet users do research on products they intend to buy (Pan et al., 2007; Vermeulen & Seegers, 2009). More precisely, 73% to 87% of consumers of product reviews within the tourism domain (e.g. hotel or restaurant reviews) denote that reviews influenced their purchase decision (Pang & Lee, 2008; Zehrer et al., 2011; Ye et al., 2011).

Indeed, the exploitation of available opinions is interesting for companies as well as users (Lin, & Huang, 2006; Carson, 2008). The former may wish to automatically extract customer feedback from online sources, or emails. By contrast, the latter request a more concise representation of opinions (Bosangit et al., 2009). Sentiment analysis, typically, quantifies the degree of positivity or negativity towards the main subject of a text. Thereby it captures the subjectivity in terms of the semantic orientation associated with the constituents of a text (Taboada et al., 2011). In

essence, sentiment analysis does what every user is required to do after writing a product review e.g. at amazon.com (http://amazon.com): to quantify the opinion represented by the text with stars.

The aim of this paper is to generate a reliable classification approach of customer reviews based on an existing domain-specific corpus by applying a lexicon-based sentiment analysis. The study comprises three steps: First, we build a lexicon of those text components with a semantic orientation. Second, we apply a sentiment analysis based on the lexicon in order to generate a classification of customer reviews. Finally, classification results are evaluated against a set of withheld reviews with quantitative ratings. We choose two different setups to demonstrate the flexibility of the proposed approach. A first analysis adopts the common classification scheme of the corpus and classifies reviews into five star-categories. The second analysis distinguishes between three categories only, thus, automatically identifying a positive, negative or neutral tendency, respectively.

The paper is structured as follows: after briefly discussing related work (section 2), section 3 introduces the corpus (3.1), describes the process of lexicon construction (3.2) as well as the sentiment computation algorithm and classification (3.3). Evaluation results are presented in section 4. Finally, section 5 concludes and gives an outlook on further research activities.

2 Related Work

A recent publication on lexicon-based sentiment analysis by Taboada et al (2011) shows the relevance of the research area for major text analysis tasks. By applying the Semantic Orientation CALculator (SO-CAL) the authors present a system performing sentiment analysis using manually created lexicons. They show that lexicon-based methods are superior to current state-of-the-art (i.e. statistically trained) classifiers. Moreover, extending lexicon content by linguistic information increases the robustness of a system, particularly, when texts stem from different domains. Both are important conclusions motivating our proposed lexicon-based approach. A survey on sentiment analysis is provided by Pang and Lee (2008). The authors focus on applications of sentiment analysis that go beyond extracting a sentiment value from a single text. Their applications range from sentiment computation towards identifying topics of a text, the visualization of sentiments as well as automatically defining the usefulness of a customer review.

Pak and Paroubek (2010) use Twitter's micro-blogging service (http://twitter.com) as opinion source to generate a corpus. Although their methodology is not lexicon-based, the corpus statistics indicate that linguistic analysis of a corpus is the key for generating lexicons of superior quality: the correlation between occurrences of a certain word category and the overall rating of a text is clearly pointed. Below, we can show from corpus statistics of our study that the same correlation exists on a semantic layer.

Finally, Taboada and Grieve (2004) defined an alternative corpus that might be reused for the evaluation of our sentiment analysis approach. 400 customer reviews extracted from epinions.com (http://epinions.com), each associated with a category, are

classified either positive or negative. Although the corpus is rather small and the classification is binary, the subset of hotel reviews may be reused. Boiy and Moens (2009) apply machine learning techniques to classify web texts into positive, negative and neutral. In contrast to our token-based sentiment calculation the authors use the sentence as sentiment unit in order to determine the semantic orientation of a document.

Gindl, Weichselbraun and Scharl (2010) focus on generating domain independent lexicons for sentiment analysis. The required disambiguation is achieved by considering the context of sentiment terms in contextualized dictionaries. The domain independent lexicon is filled with those terms identified as generic with respect to two different corpora of separate domains. Ohana and Tierney (2009) present an approach integrating data from SentiWordNet, a WordNet-based dictionary associating terms with positivity and negativity values, into the sentiment computation. Again the focus is the integration of general purpose lexicons to improve the classification performance. Contrasting their approaches we currently aim to build a domain-specific lexicon with a minimum amount of costly preprocessing steps.

3 A Lexicon-Based Classification

As the vocabulary of domain-specific documents is limited we suggest that the sentiment analysis of domain-specific documents is ideally achieved through a lexicon-based method (Taboada et al., 2011). This section summarizes our proposed approach.

3.1 The Corpus

The applied corpus comprises customer reviews of TripAdvisor (http://tripadvisor.com), the major web 2.0-platform with focus on travel and vacation services (O'Connor, 2008). Customers can book, rank and review hotels, flights and restaurants. The focus of the portal is to filter content based on rankings that are derived from user ratings. Thus, rankings are split into several categories, like value, rooms, location, cleanliness and sleep quality. Available rating categories are determined by the type of the reviewed object. A rating scale contains five values, ranging from *'terrible'* to '*excellent'*. These values are further referred to as *1star* to *5star*. A separate mandatory overall rating summarizes the total customer satisfaction. Finally, the natural language part of the review comprises a title and a text. The title is displayed in quotation marks and users are invited to use concise formulations, like *"We loved it and we'll be back!",* or *"There were things I hated."*. The text is of variable size.

Tab. 1. Number of reviews for hotels in New York

Class label	1star	2star	3star	4star	5star
#Reviews	495	613	1316	3695	4850

The used corpus from TripAdvisor is restricted to reviews of hotels written in English, Each record contains a hotel category, the overall rating, the title and the review text. Furthermore, the entries of the subcategories value, rooms, location,

cleanliness, check-in, service, and business are available. These subcategories may contain null values denoting that the user didn't care about that detail. In total, the corpus comprises over 80 000 reviews from various large tourism cities in different continents. Texts and ratings were automatically extracted from TripAdvisor.

Finally, a subset was chosen by restricting reviews to hotels located in New York. From Table 1 emerges that the number of available ratings increases with the positivity of the overall rating. For the conducted sentiment analysis a sample consisting of 200 reviews was randomly taken from each class. This restriction intends to ease the process of document analysis and lexicon construction as well as to provide the same amount of training data for each class label (Pang & Lee, 2008).

Sample sets were further split into a training set of 180 and a test set of 20 reviews. As the title tends to summarize the review, title and text were merged. Table 2 displays corpus statistics of the generated training set after using commercial natural language processing (NLP) applications of econob Informationsdienstleistungs GmbH (http://econob.com). Quality metrics for the components producing the results are not available. Nevertheless we assume that the results correspond to current state-of-commerce. The latter firm provides several NLP components that generate text annotations. An annotation, basically, is a markup of a text portion from the review, thus, identifying structural or semantic properties (Manning & Schuetze, 1999). The majority of annotations are structural annotations, like tokens or sentences, generated with a standard tokenizer and sentence detector. Semantic annotations are generated with several components responsible for named entity recognition. Amongst the most frequent semantic annotations are facilities, cities, position, and money amounts (Tab. 2).

Table 2. Corpus statistics of documents in the training set grouped by class label.

Annotation	1star	2star	3star	4star	5star	Overall
# Character	190156	215115	229302	175941	175695	986209
# Token	40891	46435	49083	37053	36683	210145
# Annotations	48829	54595	57898	44048	44747	250117
# Sentences	2669	2933	2892	2300	2126	12920
#Document	180	180	180	180	180	900
#Facility	100	102	170	240	282	894
#Money	172	186	258	74	20	710
#City	46	47	88	70	187	438
#Position	118	101	89	55	55	418

By splitting the statistics according to their class label interesting tendencies become apparent: the size of customer reviews is highest for texts labeled with *3star*, while texts with higher ratings tend to be shorter. The occurrences of money amounts indicate a negative rating, whereas occurrences of cities imply a higher rating. Hence, even from this small sample set, certain types of entities are distinctive for a specific rating - or at least indicate a positive or negative review tendency.

3.2 Lexicon Construction

As the overall sentiment value of the document is solely derived from the entries in the lexicon, its quality is the key issue in a lexicon-based sentiment analysis (Taboada et al., 2011). The lexicon in the present study was generated on the base of the vocabulary in the training set only. An entry in the lexicon is defined by a token and its part-of-speech (POS) tag. By considering no additional data sources during lexicon construction, we aim to demonstrate that generating a customized lexicon is straightforward and easy to automate. The lexicon is highly domain specific and the ensuing sentiment analysis reveals the usefulness of such a dictionary. The lexicon contains a list of tokens, each associated with a sentiment value. Values above zero denote positivity, values below zero denote negativity and zero indicates neutrality.

Let us consider the following example: in the context of hotel reviews the token 'rat' is typically associated with a strong negative value. By contrast, the token 'beautiful' is clearly representing positivity. The assignment of sentiment values to concrete lexicon entries is done prior to classification and is further described in the next section. For the construction of the lexicon the meta-data generated by NLP components is used to select the relevant subset of tokens. In detail, the processing steps include a tokenizer and part-of-speech tagger. The former identifies relevant lexical units of the text, while the POS tagger assigns a word category to each token. In order to capture the domain and to generate a lexicon of significant size all verbs and nouns are considered relevant (Taboada et al., 2011). Table 3 shows an excerpt of the most frequent tokens in the training set.

Table 3. The most frequent tokens from the training set.

Token	Frequency	POS
hotel	1851	noun
room	1842	noun
staff	618	noun
location	510	noun
stay	426	verb
breakfa	354	noun

For the sake of classification, it is important to consider only those tokens in the lexicon that are discriminating between the different class labels. Thus, for each class label a separate lexicon containing the characteristic tokens is constructed. The metric used is based on the relative token frequency with respect to POS tag and the class label. Accordingly, a token is relevant for a class label X, if the relative frequency of the token for the class label is higher than the relative frequency of the same token for all other class labels.

Moreover, an additional parameter α is used to control the size of the lexicon. That means, if $\alpha=0$ each token is assigned to those class labels it occurred at least once. For $\alpha=1$ each token is assigned to the class label with the highest relative frequency. Finally for $\alpha>1$ each token is assigned to a class label with a relative frequency that is α times higher than the relative frequency of the same token in all other classes. Thus,

for high values of α only those tokens will remain that occur only once among the five classes. This method guarantees that the lexicons are disjoint sets of tokens with a parameter $\alpha \geq 1$ (Taboada et al., 2011).

Table 4 shows the size of the lexicon per class with respect to the parameter α. Entries with α=100 show a problematic aspect of the lexicon: obviously, the majority of the entries are tokens that occurred only once in the training set. Thus, for future experiments it will be necessary to increase the number of samples in the training set. Nevertheless, the experiments described in subsequent sections are based on the lexicons generated with the parameter α=2. This value ensures that besides all tokens specific for a class, a minimum amount of overlapping tokens are included in the lexicon of one class.

Table 4. The size of the lexicon per category.

α	**1star**	**2star**	**3star**	**4star**	**5star**
0	3481	3513	3565	2919	2781
1	1558	1510	1534	1026	1209
2	1197	1098	1121	805	963
3	1141	1032	1048	774	890
100	1096	1007	1007	753	829

The second analysis is based on a restructured training set distinguishing only between three class labels. For a broad range of applications distinguishing between positive, negative and neutral texts seems sufficient, as the difference between ratings of finer grained rating schemes is difficult to define anyway. Nevertheless, it is important to state that our proposed approach aims to be flexible with respect to defined classes. More concretely, the class label *bad* contains all texts from the categories *1star* and *2star*, the class label *neutral* comprises the reviews from *3star* and the class label *good* includes the remaining texts from *4star* and *5star*. As we are interested in the best possible classification performance for the classes *good* and *bad*, we only generate two dictionaries. An additional lexicon for the *neutral* class label reduces the other lexicons by terms indicating slightly positive or negative sentiment and thus reduces the accuracy results for our approach. Table 5 shows the size of the lexicons according to the second analysis.

Table 5. The size of the lexicon per category for a second analysis.

α	**good**	**bad**
0	4327	5216
1	2809	3983
2	2384	3455
3	2240	3225
100	2058	2947

Obviously, decreasing the number of target labels affects the size of lexicons, as the individual training corpora are larger. Again, we choose the lexicons generated with the parameter α=2 throughout the remaining sections to ensure comparability.

3.3 Classification

In order to compute a sentiment value from reviews the lexicon entries are associated with a semantic orientation (Taboada et al., 2011). Each of the five distinct lexicons and, thus, all of its entries are assigned separate values. The basic assumption is that a document without any values identified is neutral and, thus, has a sentiment value of 0. Hence, no prior probabilities for a customer review that belongs to certain class labels are considered. Values are assigned straightforward: *1star* dictionary entries are weighted negative -2, *2star* are weighted -1, *3star* are associated with the neutral 0, *4star* are slightly positive 1 and, finally, *5star* are excellent +2.

The classification function is computed based on the sentiment analysis of the documents in the training set and the lexicons generated. In order to compute the sentiment value for one document, the sum of all identified sentiment values is generated (Pang & Lee, 2008). More sophisticated algorithms (Taboada et al., 2011) are not necessary, as we currently do not account for negation or intensification.

For the classification functions the average sentiment value of all documents per class label is used. Figure 1 shows the average sentiment values for all class labels. These functions can subsequently be used to classify new customer reviews by first computing the sentiment score and then calculating the distance of the score to the classes' average. The class with minimal distance to the documents' sentiment value is used as class label.

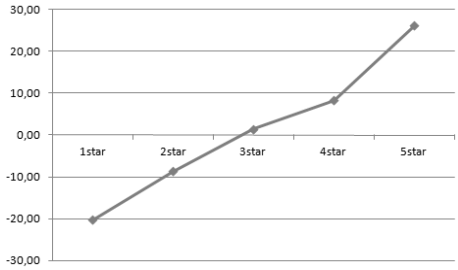


Fig. 1 Average sentiment values for five classes

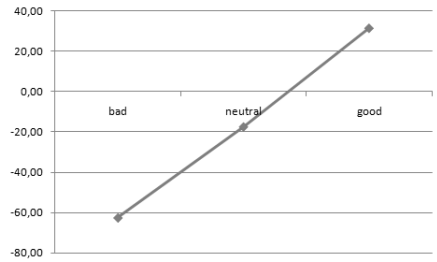


Fig. 2 Average sentiment values for three classes

Functions for different labels are well separated and should, thus, well reflect the characteristics of the different classes. The second classification function limited to three class labels is comparably computed. The entries of the lexicon with target label *bad* are associated with the value -2 and *good* are assigned with +2. Figure 2 shows the resulting classification function.

The high distance between the different class labels is mainly due to the higher distance between the sentiment values of the dictionaries as well as the larger amount of entries. In the next section the computed classification functions are evaluated.

4 Evaluation

Preceding the computation of the classification the data set was split into a training set containing 90% and a test set containing 10% of all reviews of a given class label. Thus, five distinct test sets are evaluated separately, each containing about 20 documents. As evaluation metrics precision and recall measures were chosen (Manning & Schuetze, 1999). Precision defines the proportion of reviews the system classified correctly to all reviews classified. Recall describes the proportion of reviews selected correctly to all reviews selected. An additional F-measure combines both precision and recall into a single measure by computing the harmonic mean. The F-measure uses a parameter controlling the influence of precision and recall. As we have no concrete application in mind, we assume precision and recall equally important and, thus, set the weighting parameter to 0.5. The performance of the proposed system is compared to a baseline computed from randomly assigning class labels. Figure 3 summarizes the evaluation results.

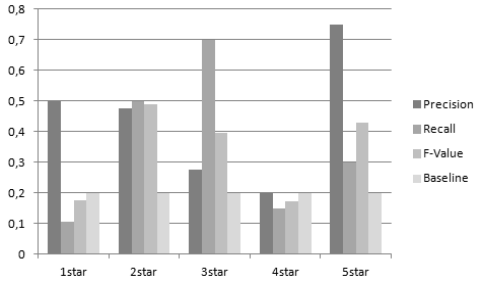


Fig. 3 P,R and F1 for five classes

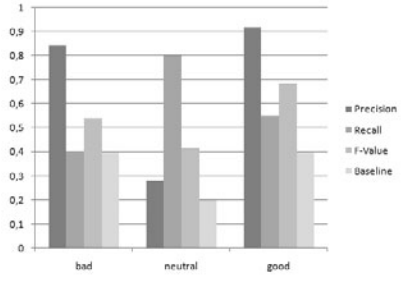


Fig. 4 P,R and F1 for three classes

The precision, recall and F-measure values over all class labels (i.e. *1star* to *5star*) are considerably higher than the baseline values. A closer look at individual performance values of the classes uncovers several interesting issues. The recall for the classes *1star* and *5star* is very low. This, clearly, is due to the high (respectively low) sentiment values needed for a document to be labeled correctly. A high (resp. low) score is achieved with a lot of either positive or negative sentiment bearing tokens identified in the document.

Hence, the lexicon is much too small or not distinctive enough to encounter reviews that are not part of the training set. On the other hand, precision is highest for *5star* and *1star* because very few reviews from other classes are scored with a high value that fall into this category. Class *4star* is a negative outlier: 70% of the documents belonging to this class were incorrectly classified as *3star* documents. Again, this indicates that a very small amount of semantically relevant tokens were found, leading to the conclusion that the quality of the lexicon should be improved. Sentiment scores near 0 indicate either that the semantic orientation is neutral (positive and negative values neutralize each other) or that no sentiment indicators were found. This again reveals the weakness of the first used lexicon.

Figure 4 shows the results of the evaluation of the second classification function covering only three distinct classes. As in the dictionary generation process the reviews of multiple training sets were merged, the size of the test set for both class labels *bad* and *good* comprises 40 documents each. The obtained results for the classes *bad* and *good* are nearly similar: a precision of 84% respectively 92% signals the high quality of a positive classification. By contrast, the recall values of 40% for *bad* and 55% for *good* indicate that many examples were falsely classified as *neutral*. However, the false classification of a review belonging to the class *bad* into *good* (or the other way around) was observed only once in the test set, leading to the conclusion that those classes are very well separated. The high recall of 80% for the class label *neutral* paired with a precision of 28% approves the observation from the analysis using five classes: for false positive reviews too few tokens with semantic orientation were identified or the positive and negative tokens neutralize each other. Thus, by decreasing the distance in the classification function between class labels, the recall increases at expense of precision. Nevertheless, the average F-value of 55% is significantly higher than the computed baseline.

To sum up, comparing both presented evaluations shows that decreasing the number of class labels accompanied with a larger training set and a larger lexicon would significantly increase the overall classification performance. The system was also evaluated with one lexicon per target label, but the performance was inferior to the current solution using two lexicons for *bad* and *good*.

Gindl et al. (2010) also present an evaluation based on reviews from TripAdvisor. Their evaluation only considers the classification of positive and negative reviews, ignoring neutral texts for training and testing. The classification of positive reviews has a precision of 66% and a recall of 97%. The classification of negative reviews has a precision of 95% and a recall of 46%. Our precision values are constantly high (84% and 92%) for both positive and negative reviews. Our recall values are lower because a lot of reviews are classified as neutral. Directly comparing the performance of both systems is difficult, because Gindl et al. do not consider neutral reviews. But keeping an application in mind, it is seems crucial to also consider neutral customer reviews.

5 Conclusion and Future Work

In our study on sentiment analysis we proposed a lexicon-based approach to classify customer reviews in the tourism domain. With precision and recall values significantly exceeding the given baseline our proposed methodology for constructing a domain specific lexicon paired with the algorithm for sentiment analysis proved to be successful. Especially the analysis using only three target labels (i.e. *good, neutral* and *bad*) may be used in real world applications to extract sentiment from text resources, as the precision for the class labels representing positive and negative sentiment showed to be remarkably high. Put simply, if a customer review is classified as *good* or *bad* the classification is correct with a probability of about 90%. Furthermore, since literature is still scarce, the outlined system performance could be used as a new baseline for future evaluations.

Finally, the study supports the definition of future research designs and optimization goals to improve the overall performance of customer classifiers based on sentiment values. In more detail, such improvements are sketched below:

First, the samples taken from the corpus determine the quality of the lexicon. The analysis of classification results and the size of the lexicons show that increasing the sample size affects the specific vocabulary used for customer reviews. This is corroborated by the fact that customer reviews are rather short documents. As a consequence, sample size should be much larger. Second, the system is backed by five, respectively three, lexicons each containing disjoint semantic indicators for corresponding class labels. However, the set of lexicons has to be extended by a domain specific lexicon relaxing the disjoint criteria combined with lexicons introducing intensifiers, downgraders and negators (Pang & Lee, 2008).

Further domain independent lexica of commonly used positive or negative tokens may be integrated. These efforts will be positive for the quality of the classification at the expense of higher costs for lexicon construction. Additionally, the information gained from corpus analysis (section 3.1; e.g. the occurrence of an entity representing a city is a positive indicator), may be used to generate additional lexicons encoding encyclopedic knowledge. Experiments with the value of the parameter α should also help in optimizing the quality of the lexicon currently used. Third, the further refinement of sentiment values, currently ranging from -2 to +2, might also help to optimize the lexicon. Fourth, for the computation of the classification performance a priori probabilities for customer reviews belonging to certain categories may be considered as the corpus shows that positive reviews are much more likely than negative ones. Additional corpus statistics, such as text length, should be incorporated.

Finally, the corpus grounding the analysis comprises customer reviews from hotels around the world. However, for the presented analysis only reviews for hotels located in New York were considered. Thus, further experiments should be conducted in the future to evaluate the impact of various domain and region specific parameters. As described in section 3.1 the corpus contains, besides the overall rating, a subcategory rating for each review. Extending sentiment analysis to these subcategories may similarly lead to interesting new results.

References

Bosangit, C., McCabe, S. & Hibbert, S. (2009). What is Told in Travel Blogs? Exploring Travel Blogs for Consumer Narrative Analysis. In W. Höpken, U. Gretzel & R. Law (Eds.). *Information and Communication Technologies in Tourism 2009*. Wien New York: Springer, pp.61-71.

Boiy, E. & Moens, M. (2009). A machine learning approach to sentiment analysis in multilingual Web texts. *Information Retrieval* 12(5): 526-558.

Carson, D. (2008). The Blogosphere as a Market Research Tool for Tourism Destinations. *Journal of Vacation Marketing* 14(2): 111-119.

Dippelreiter, B., Grün, Chr., Pöttler, M., Seidel, I., Berger, H., Dittenbach, M. & Pesenhofer, A. (2007). Online Tourism Communities on the Path to Web 2.0 - An Evaluation, Virtual Communities in Travel and Tourism. *Information Technology & Tourism* 10(4): 329-353.

Gindl, S., Weichselbraun, A. & Scharl, A. (2010). Cross-Domain Contextualisation of Sentiment Lexicons. In *Proceeding of the 2010 conference on ECAI 2010: 19th European Conference on Artificial Intelligence*, pp. 771-776.

Lin, Y.S. & Huang, J.Y., (2006). Internet blogs as a tourism marketing medium: A case study. *Journal of Business Research* 59(10-11): 1201-1205.

Litvin, S.W., Goldsmith, R.E. & Pan, B. (2008). Electronic Word-Of-Mouth in Hospitality and Tourism Management. *Tourism Management* 29(3): 458-68.

Manning, C.D. & Schuetze, H. (1999). *Foundations of Statistical Natural Language Processing* (1st edition). MIT Press,.

O'Connor, P. (2008). User-Generated Content and Travel - A Case Study on Tripadvisor.Com. In O'Connor, P., Höpken, W. and Gretzel, U (Eds.), *Information and Communication Technologies in Tourism 2008*. New York: Springer, pp. 47-58.

Ohana, B. & Tierney, B. (2009). Sentiment Classification of Reviews Using SentiWordNet. In *Proceedings of the 9th IT&T Conference.* Dublin, Ireland: Dublin Institute of Technology.

Pak, A. & Paroubek, P. (2010). Twitter as a Corpus for Sentiment Analysis and Opinion Mining. In *Proceedings of the Seventh conference on International Language Resources and Evaluation* (LREC'10), Valletta, Malta.

Pan, B., MacLaurin, T. & Crotts, J. C. (2007). Travel Blogs and the Implications for Destination Marketing. *Journal of Travel Research* 46(4): 35-45.

Pang, B. & Lee, L. (2008). Opinion Mining and Sentiment Analysis, *Foundations and Trends in Information Retrieval* 2(1-2): 1–135.

Scharl, A., Pollach, I. & Bauer, C. (2003). Determining the Semantic Orientation of Web-based Corpora. In J. Liu, Y. Cheung & H. Yin (Eds.), *Intelligent Data Engineering and Automated Learning*, 4th Int. Conference, Ideal, 2003, Hong Kong (Lecture Notes in Computer Science Vol. 2690). Berlin: Springer, pp. 840-849.

Schmallegger, D. & Carson, D. (2008). Blogs in Tourism: Changing Approaches to Information Exchange. *Journal of Vacation Marketing* 14(2): 99-110.

Taboada, M., Brooke, J., Tofiloski, M., Voll, K.D. & Stede, M., (2011). Lexicon-Based Methods for Sentiment Analysis. *Computational Linguistics* 37(2): 267–307.

Taboada, M. & Grieve, J. (2004). Analyzing Appraisal Automatically, *Proceedings of the AAAI Spring Symposium on Exploring Attitude and Affect in Text*.

Vermeulen, I.E. & Seegers, D. (2009). Tried and Tested: The Impact of Online Hotel Reviews on Consumer Consideration. *Tourism Management* 30(2): 123-127.

Xiang, Z. & Gretzel, U. (2010). Role of Social Media in Online Travel Information Search. *Tourism Management* 31(2): 179-188

Ye, Q., Law, R., Gu, B. & Chen, W. (2011). The Influence of User-Generated Content on Traveler Behavior: An Empirical Investigation on the Effects of e-Word-Of-Mouth to Hotel Online Bookings. *Computers in Human Behavior* 27(2): 634-39.

Zehrer, A., Crotts, J.C. & Magnini, V.P. (2011). The Perceived Usefulness of Blog Postings: An Extension of the Expectancy-Disconfirmation Paradigm. *Tourism Management* 32(1): 106-13.

Acknowledgements

Special thanks goes to econob (http://www.econob.com) for providing infrastructure, interesting ideas and the natural language processing suite.

Monitoring and Summarization of Hotel Reviews

Walter Kasper, and
Mihaela Vela

Language Technology Lab
DFKI GmbH
Stuhlsatzenhausweg 3
66123 Saarbrücken, Germany
{Walter.Kasper,Mihaela.Vela}@dfki.de

Abstract

User reviews and comments on hotels on the Web are an important information source in travel planning. Knowledge of such comments therefore can be important to the hotel management for quality control. But often it is difficult to find and to follow such information on the Web. We present a system that automatically monitors user comments on the Web from various sites and provides classified summaries of positive and negative features of a hotel.

Keywords: Opinion Mining, Sentiment Analysis, Summarization, Hotel Reviews

1 Introduction

Travel planning and booking on the Web has become one of its most important commercial uses. With the rise of the Web-2.0 user-generated reviews, comments and reports about their travel experiences play an increasing role as information source. Especially for hotel booking, such user reviews are relevant since they are more actual and detailed than reviews found in traditional printed hotel guides, etc., they are not biased by marketing considerations as e.g. the hotels' home pages or catalogue descriptions, and, finally, they reflect actual experiences of guests.

Though nearly every Internet travel agency and hotel booking service nowadays offers also ratings and/or reviews of hotels, it is not that easy for hoteliers who want to know what is published about their hotels on the Web to gather the user-generated information. A standard search engine like Google will give thousands of hits for a hotel. But, though there seems to be a huge number of sites providing user reviews, often these are just the same because many sites use the same source, such as *openholidayguide.com*. In other cases, the links lead only to some general page from which one can access reviews besides other information and lacking transparent navigation structure. Also, the links might point to some individual review but leaving it open whether there are other reviews on the site. An additional problem is that the Web-2.0 provides a large number of publication types: besides travel agencies and hotel booking services there are numerous blogs, fora, newsgroups, social networks etc. related to travelling.

Another problem concerns the kind of information: travel agencies and hotel booking services often only publish scalar ratings, e.g. scores between 1 and 5. Such scores are not very helpful for hotel managers as the numeric value does not provide information of what guests actually considered positive or objectionable. Also, the numeric scores are not comparable: when a 3-star hotel receives a higher score than a 4-star hotel, that does not imply that the one is better than the other. For hotel managers the textual

user comments would be much more significant than the numeric scores since they would be interested to know *what* the users exactly commented on and *how* they thought of it.

Another problem for hotel managers is that of following updates and new reviews. Hotel booking services and travel agencies collect and publish user reviews systematically, e.g. by asking their customers for comments or ratings. So, new reviews appear quite frequently on their pages but it would be difficult to follow these by just using general search.

For the user who is accessing reviews on the Web for planning his travel, many of these considerations are not relevant, as he will be content with a momentary snapshot of reviews. But for hoteliers interested in user comments on the Web a service that automatically and systematically collects and summarizes the relevant information from the Web would be advantageous and perhaps even more useful than the paper forms many hotels use for gathering feedback from their guests.

The BESOHOT service presented in this papers aims at providing such a service for hotel managers that collects user reviews for hotels from various sites on the Web, analyzes and classifies the textual content of the review and presents the results in a concise manner. We will give an overview of the system in Section 2 and discuss the major components in more detail, the data acquisition from the Web (Section 2.2), the statistical polarity classification (Section 2.3) and the linguistic information extraction (IE) components (Section 2.4). The user interface will be presented in Section 3. In Section 4 evaluation results for the analysis system will be presented. In Section 5 we will relate our work to other work in opinion mining.

2 Overview of the System

The BESOHOT system is an interactive Web application. The core system on the server-side handles *data acquisition*, *analysis* and *storage* as shown in Fig . The user interface provides various types of summaries of the analyzed data, allows direct access to the information sources on the Web as well as free text search.

New data retrieved from the Web by the acquisition system are passed to the analysis system. The analysis system first does a *language check* (*LangID*) to filter out reviews in other languages than German because even German hotel review sites occasionally contain reviews in other languages. The review texts then get segmented into segments ("sentences").[1] These segments are then subjected to further analysis by the *statistical polarity classifier* and *linguistic information extraction* (IE) components for finer grained analysis of the polarity and the topic of the review. Polarity values are always assigned to text segments, not to complete reviews. The polarity values from the statistical and the linguistic classification get combined into a joint global polarity value. Finally, the analysis results are stored together with the review segments in a special *ResultDB* optimized for the retrieval and interaction requirements of the user interface.

[1] We prefer the term "segments" to "sentence" because the segments are not always sentences in a linguistic sense but just phrases.

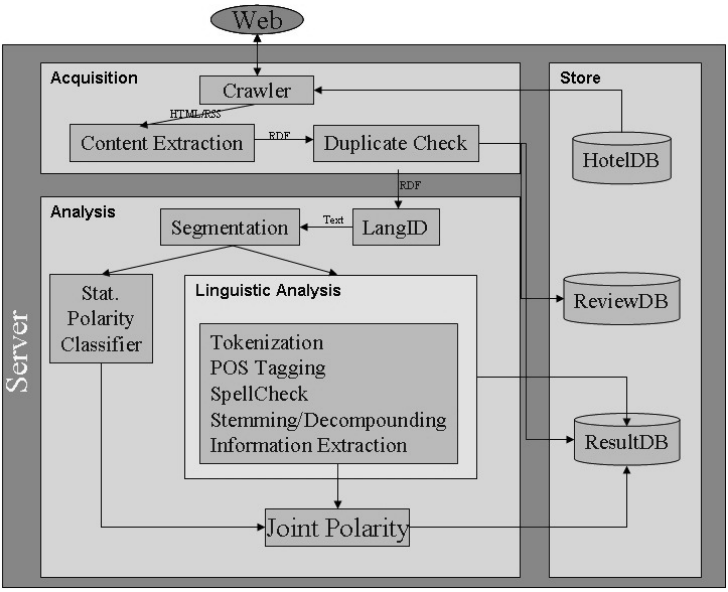


Fig 1. BESOHOT System Overview

2.1 Polarities

In general, we distinguish three possible polarity values for text segments: the segment can express a *positive* opinion, a *negative* one or a *neutral* one. By neutral segments, we mean purely descriptive ones that do not carry an evaluation, such as *We spent three days at the hotel*. The delimitation of neutral/descriptive and evaluative text is not always easy, not even to humans. A remark like *no minibar* on the one hand just describes a fact but on the other hand the remark probably is meant as a negative comment describing a deficiency.

Another problem for a polarity classification on text segments is that a segment might address more than one topic. For example, *clean rooms and friendly personnel* addresses the two topics *room* and *personnel* both rated as positive. But for a comment like *Room ok, but poor breakfast* it would be unclear what the overall polarity value of the comment should be, as there are actually two ratings on two different topics. Similar issues arise with respect to multiple ratings on the same topic as in *clean, but tiny room*.

The BESOHOT IE system is able to detect such multiple topics and ratings on a text segment. Nevertheless, as we have not yet found a good solution for handling these cases in the user interface, at present we prefer to disregard them in favour of a global polarity assignment, even if that sometimes might be a bit random. This will be further discussed in Sections 2.5 and 4.

2.2 The Acquisition System

The acquisition of reviews from the Web is handled by a Web crawler. The *HotelDB* defines for each hotel a set of crawl configurations for the sites to visit. Currently,

BESOHOT covers 7 of the major hotel review and hotel booking sites, such as *Tripadvisor.de*, *Holidaycheck.de*, *Trivago.de*, *hotel.de*, etc. *Holidaycheck* is treated as the best representative for the many sites that use *openholidayguide.com* as data source.

When a target page is retrieved a *content extraction module* is applied that extracts the relevant textual content of the review but also other metadata such as scores and information about the reviewer/guests. If a page contains several reviews, for each of them a separate review instance is created. Since the content of the crawled Web pages is dynamic the system needs to determine whether it has seen a review before or whether it is a new review. The *duplicate check* uses review fingerprints created from the textual content without any formatting. This provides reliable and efficient tests independent of text size and formatting. Review instances that survive the duplicate check are stored in the *ReviewDB* and passed to the analysis system. The review texts there first are split into text segments that become the units of further analysis.

2.3 Statistical Polarity Classification

The statistical polarity classifier assigns to each text segment a polarity value. As a basis for statistical polarity classification the classification engine of Steffen (2004) is used. This engine is based on *character n-grams* instead of terms as in standard methods. For instance, the word *hotel* gets decomposed into the 2 4-grams *{hote, otel}*, into 3 3-grams, 4 2-grams and 5 1-grams (the letters). The classifier then computes for each segment and class the probability of the "best" matching n-gram sequence. We use 4-gram models, using the lower n-gram ranks if some 4-gram is not found, e.g. in words not seen in the training data. Higher values of *n* do not improve the performance. This approach has several advantages: robustness against the frequent orthographic errors, robustness against unknown terms e.g. from word compounding in German, it diminishes the sparse data problem for training and it can be applied to short texts.

For getting training data for the statistical classifier we exploited the fact that on some hotel review sites, such as *hotel.de,* users themselves classify their comments into positive and negative text items. A corpus of such hotel review texts from these sites was collected and used for training the classifier with 2 polarity classes (positive/negative), roughly about 7200 text segments for each class. Cross-validation benchmarks using randomly selected 50% and 90% of the training corpus for training and the rest for testing demonstrated a good performance with precision/recall rates of 90% and 93%, respectively. We use only two polarity values for the statistical classifier. An experiment to add a neutral category from manually classified data showed clear performance degradation. Therefore, we preferred to leave the detection of neutral segments to the IE. In Section 4 we will further discuss the performance with respect to manually annotated data and the problem of multi-topic and neutral text segments.

2.4 Information Extraction

The main task of the linguistic analysis components in the BESOHOT system is to identify from a text segment its *topics* (what is talked about) and how these get rated

within the segment. The core of that analysis is a rule-based information extraction (IE) component (Drozdzynski et al., 2004). The IE system is designed to supply answers to the following questions: topic of the review (what is evaluated), dimension of the evaluation (what properties are evaluated), dimension value (what is the value on that dimension) and polarity of the evaluation (is it positive or negative or neutral).

Some preprocessing and normalization steps are applied before submitting the text to the IE system. Part of Speech (POS) tagging is used to reduce the search space for lexically ambiguous words and word forms. A frequent problem in processing German is word composition by which several terms are combined into a single word. Such compounds often are missing in the usual dictionaries and therefore difficult to process. German morphology allows to decompose unknown word compounds into multi-word expressions. Also, to improve input quality a spell check is applied that automatically can correct frequent types of spelling errors like transposed and left-out characters.

The IE component uses a dictionary of *domain-specific terms* relevant for the hotel domain as well as a *sentiment dictionary* that associates basic polarity values with terms[1]. The IE analysis rules allow for fine-grained analysis not limited to some word classes such as adjectives and nouns (e.g. Hu and Liu (2004) or Blair-Goldensohn et al. (2008)) and it can take into account linguistic contexts and markers that can affect the expressed polarity such as negation, contrastive or adversative markers and clauses etc.

The IE system distinguishes several types of possible roles for a polarity value that influence in different ways what actual polarity is expressed in a segment.

- evaluative speech act indicators, such as *regrettably*. These can override any other polarity expressed.
- negation particles, e.g. *not* that will turn polarities in their scope to the opposite.
- polarity modifiers, e.g. the *too* in a phrase like *too small* that can override the default polarity at phrasal level.
- "missing features" indicators, such as *without*
- a default lexical polarity, e.g. that *nice* expresses a positive rating
- idiomatic polarity expressions

Fig depicts the semantic representation of *kein kostenloses schnelles WLAN (no free fast WLAN)* from IE as a feature structure. It can be read as: *WLAN* is the topic belonging to the *telecommunication* category. There are two properties attached that by default denote positive properties (*free, fast*), shown as values of the LEXPOL feature. But these occurrences are in the scope of a negation polarity, the `NEGPOL` value that will invert these default values. This is handled by an IE postprocessor. So in the end we will have two negative ratings for the WLAN topic as being neither free

[1] A generic German sentiment lexicon was developed by Waltinger (2010) that we used in initializing our sentiment lexicon from corpus data.

nor fast. The POLMOD and SAPOL features are used to mark polarity modifiers and speech act polarities, respectively.[1]

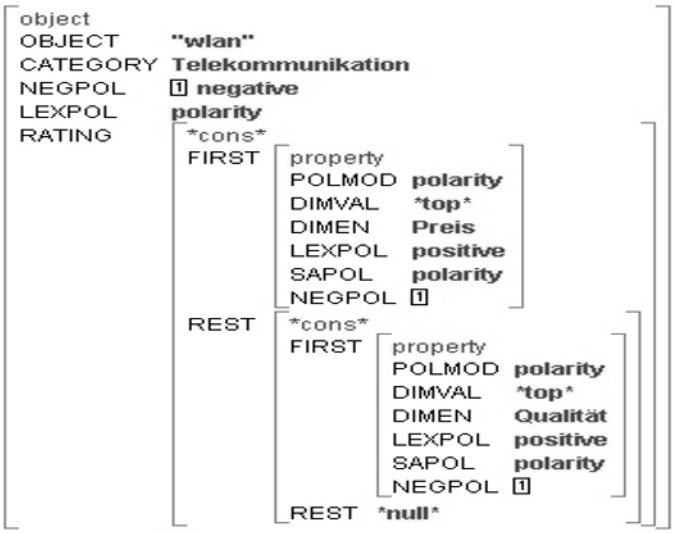


Fig 2. Semantic representation for *kein kostenloses schnelles WLAN*

The resulting annotation structures from IE are passed to a postprocessor that evaluates them and computes the final rating values for a segment taking into account the different types of polarities and their scopes. This postprocessor would recognize that the positive lexical default polarity values of the adjectives in Fig occur in the scope of a negative polarity marker and therefore would invert them, resulting finally in two negative ratings instead of two positive and some negative polarity. Also, isolated annotations that cannot be related to ratings get eliminated here.

It is obvious that for the IE system the representation of multiple topics and multiple ratings in a text segment is not a problem. The absence of rating annotations in a segment can be regarded as evidence that the segment belongs to the neutral polarity category.

2.5 Combining Statistical and IE Polarities

For each segment the statistical polarity classifier yields a positive or negative polarity value. From IE, more fine-grained polarity values are available even for parts of the segments (sub-segments). We developed an experimental system that would use the IE to create finer phrases as sub-segments of the text segments according to the recognized topic changes. Unfortunately, in many cases that resulted in text fragments that are incomprehensible without their syntactic context and so cannot be

[1] The value *polarity* on any of the *POL features that correspond to the different roles of the polarity values just designates a neutral/unmarked value.

presented to users. Therefore we kept to the approach to assign a global polarity value to the whole text segment, but the assignment of that global value would take into account both classification sources, the statistical value and the IE values. The statistical value is regarded as baseline value and the ratings from IE are used to possibly correct that value. As an approach that would give the IE ratings preference to the statistical value proved unsatisfactory, we developed a method for using the IE ratings as length-normalized weights on the statistical values: for each polarity, the IE weight is defined as the number of ratings of that polarity divided by the token length of the segments. On short segments, the IE ratings thus will have larger weight than on longer segments. The global polarity values then are computed by combining the scores of the statistical classifier with these weights according to (1).

$$pol = \arg\max_{p \in \{pos, neg\}} \frac{sp(p)}{1 + ie(p)/sl} \tag{1}$$

where p is a polarity, $sp(p)$ its statistical score, sl the segment length and $ie(p)$ the number of the IE ratings with that polarity. This approach reconciles the confidence of the statistical classifier with the IE results better than a preference based approach. A side effect of the formula is that the statistical polarity value will be kept, if the IE does not yield ratings. The main motivation is that the statistical classifier has larger coverage than the current IE. Therefore we keep the statistical polarity value and treat the absence of IE ratings as meaning "IE does not know" rather than "This is neutral polarity". This provides more flexibility for the user interface that can decide how to handle this case.

3 The User Interface

The BESOHOT system is a tool to support hotel managers in quality control. So it should provide them with fast and comprehensive overviews and summaries of how their hotel is rated on the Web and how it is commented on by guests and visitors on the Web.

Fig shows the main result overview that the user will see when accessing the BESOHOT service after selecting a hotel. The top panel displays some statistics about scores from source sites, normalized to a scale between 1 and 10, and about guest types, as far as this information could be extracted from the source Web pages. Also, the time range can be restricted to show only recent reviews. The *Aktualisieren* button allows to start the crawler to search for new reviews on the Web for the selected hotel.[1]

The main panel provides a summary of the reviews by displaying text snippets from the reviews according to their polarity and category. A click on a segment opens a popup panel that displays the full review text highlighting the displayed segment in context. This allows users to check the text in context and also makes it unnecessary to visit the source page, though this would be easy by just following the provided link to the source page. Additionally, the popup displays information about the guest that

[1] This Actualize button exists only in the demonstration system. In the final system the server automatically updates the databases periodically.

provided the rating. For this display we exploit the IE's capability to identify neutral text segments: text segments that do not receive an IE rating are omitted here from the view. An open issue in designing the user interface is the handling of text segments belonging to more than one category. Adding these segments to each category tends to result in rather crowded and redundant category fields, impairing the usefulness. So, presently such ambiguous segments currently are displayed only in one category, preferably a dominant one.

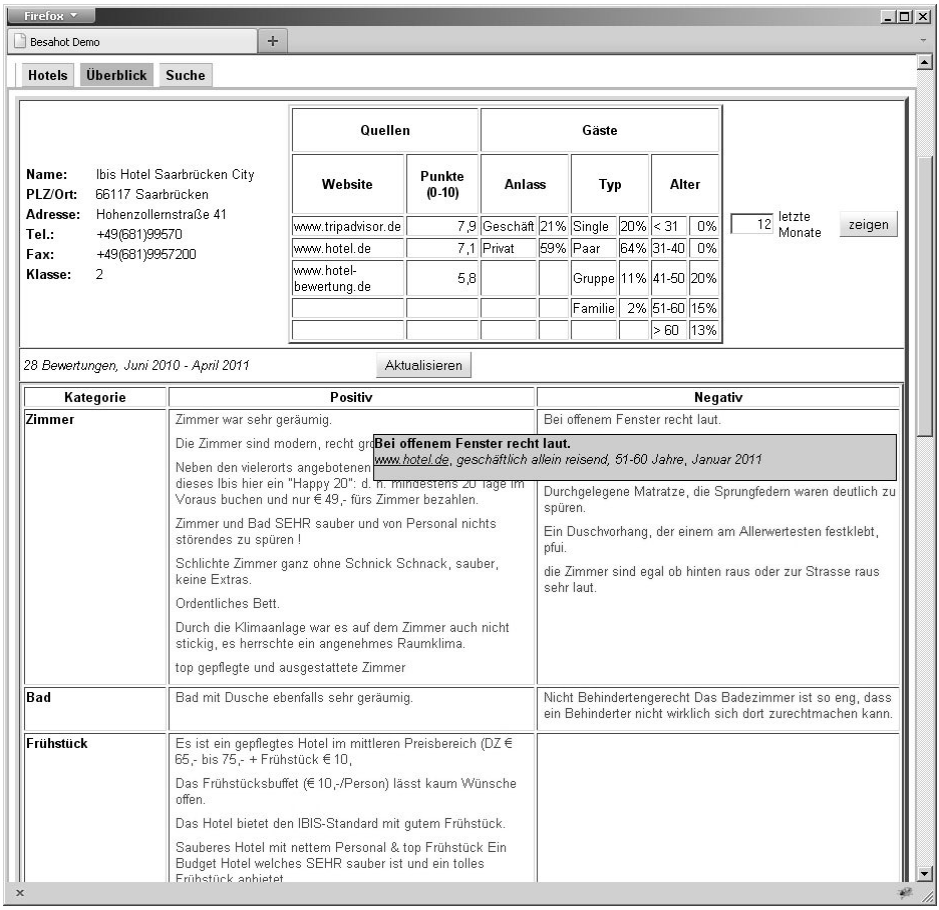


Fig 3. Classified summary for a given hotel

In addition to the overview presentation, a free text search function allows users to search the review database for text segments by freely chosen keywords, independent of the predefined categories and polarity values.

4 Evaluation

We evaluated the analysis system on a corpus of 1,559 hotel reviews crawled from the Web. These reviews contained 4,792 text segments. For the evaluation, these

segments were manually classified with respect to their polarity, including the neutral polarity besides positive and negative ones. Also, we annotated the segments whether they cover more than one topic. The distribution from this manual classification is shown in Table 1.

Table 1. Manual Corpus Classification

Segments	positive	negative	neutral	multi-topic
4792	2240	1183	938	431

We evaluated the performance of the statistical classifier alone, the IE system alone and the hybrid system combining the polarity classifications from the statistical classifier and the IE system. Evaluation results on all segments and the non-neutral segments only are shown in Table 2.

Table 2. Accuracy of Polarity Classification

	Correct	False	Accuracy total	Accuracy without neutrals
Stat	3145	705	0.66	0.80
IE	2604	486	0.54	0.66
Stat+IE	3208	646	0.67	0.81

It shows that the IE system currently covers less data than the statistical classification, but that it improves the overall classification accuracy. The insufficient coverage is due mainly to gaps in the domain and sentiment lexicon, not the rule base. The overall accuracy on the set of all segments is impaired significantly by the high rate of neutral and multi-topic segments in our corpus as is apparent from the difference to the accuracy values when these cases are discarded. The difference of about 14% shows that it would be beneficial to be able to identify neutral and multi-topic/multi-polarity ratings. The pure statistical classifier did not look promising in that respect. An evaluation of how well the IE system would recognize the neutral and the multi-topic segments showed good accuracies of 72% and 75%, respectively. We expect that improving the coverage of the IE system will also improve these figures. That will also provide a strong motivation for changing the interpretation of the absence of a polarity rating from IE as “don't know” to “classify that as *neutral*”.

5 Related Work

The development of the WWW and the possibility for customers/users to express their opinion online made the online available reviews interesting for both the vendor as well as for the potential customer. Therefore, the interest in opinions and sentiments of customers published on the Web has increased tremendously. In parallel, the development boosted research in opinion mining and sentiment analysis

in recent years. Good overviews on existing opinion mining techniques and methods are given by Pang and Lee (2008) and Liu (2010). Most research in this area concentrates on opinions about products. Also, domains such as movie reviews or news found considerable interest especially in research, since large datasets and corpora are publicly available. The goals of opinion mining can vary considerably. In many cases, one is only interested in a global overview: how many users/reviews rate a product positive or negative. For these, a global polarity classification of the complete document or sentences can be sufficient without having to take into account details of a product.

While our IE system for the categorization and extraction of product features relies on manually created rules, there are a number of approaches that use machine learning techniques to achieve that, such as the work of Popescu and Etzioni (2005) or Hummel and Zucker (1983). A more linguistically inspired approach that resembles ours is described in Ku et al. (2009).

The tourism domain is not one of the mainstream domains for opinion mining research. Aciar (2009) uses a corpus of English reviews from *tripadvisor.com* in order to present a rule-based method for classifying opinions with respect to the differing needs of different guest types. A larger English corpus also from *tripadvisor.com* is used in the study of Baccianella et al. (2009) that uses linguistic preprocessing with the SENTIWordnet (Esuli & Sebastiani, 2006)) but machine learning techniques for feature assignment. Haruechaiyasak et al. (2010) describe a framework for constructing Thai language resources for feature-based opinion mining for hotel reviews.

Their approach for extracting features and polarity words from opinionated texts is based on syntactic pattern analysis as in our IE system. In general, these approaches focus on research on specific technologies but there is little indication about what the results are used for in an application, who the users of the results are and how results can be used by them. In many cases the research is related to building recommendation systems so that the results are not directly used by humans but just by machines. The BESOHOT system, on the hand, targets explicitly human users, not machines.

Closely related to BESOHOT is the work on review summarization such as Hu and Liu (2004) and Blair-Goldensohn et al. (2008). Summarization there means extracting relevant sentences classified according to their polarity and some category, called *features* or *aspects* in these papers. They focus on adjectives as carriers of polarity and nouns/noun groups as designators for features, ignoring other word classes. Negation seems to be recognized only if adjacent to an opinion term. Irrelevance/neutrality is defined by thresholds on scores. The methods of feature extraction based on nouns in the context of opinion terms tend to yield high numbers of features. Blair-Goldensohn et al. (2008) therefore introduce a second level of manually created static high-level aspects that resemble more the high-level categories used in BESOHOT. The BESOHOT-IE approach looks more flexible as it is not restricted to few word classes and it can handle larger contexts and relevant linguistic phenomena better than these approaches. Also, resources for the IE are easy to extend and to adapt for new data and phenomena.

A few other services for tracking hotel reviews have appeared recently. *Raveable.com* uses review analysis for ranking hotels according to some categories and also provides text extracts from reviews, but unclear arranged and without polarity indication. It is more a recommendation service for travellers. Commercial services for hotel management difficult to evaluate and compare technologically with BESOHOT from outside are *ReviewAnalyst.com* and *TrustYou.com*. Only *TrustYou.com* handles German.

6 Conclusion

We presented a Web based opinion mining system for hotel reviews and user comments that supports the hotel management in monitoring what is published on the Web about their houses. The system is capable of detecting and retrieving reviews on the Web, to classify and analyze them, as well as to generate comprehensive overviews of these comments. The system provides good performance for the analysis and the classification tasks. Further research will be necessary especially with respect to the demarcation of evaluative and neutral text as well as to the handling of multi-topic segments, especially for the user interface. Also, we are preparing a pilot test of the BESOHOT service with members of the Saarland hotel association to improve the information value and usability of the system. We are aware of legal issues with respect to the review data when the system is deployed in a commercial context.

Acknowledgments

The reported work was supported by a grant from the Saarland Ministry of Economics and by the European Commission under contract number FP7-231527 (IKS).

References

Aciar, S. (2009). Mining context information from consumer's reviews. In *Proceedings of the Context-Aware Recommender Systems (CARS) Workshop*.

Baccianella, S., Esuli, A. & Sebastiani, F. (2009). Multi-facet rating of product reviews. In *Proceedings of the 31th European Conference on IR Research on Advances in Information Retrieval*, ECIR '09, 461-472. Berlin ,Heidelberg: Springer-Verlag.

Blair-Goldensohn, S., Hannan, K., McDonald, R., Neylon, T., Reis, G. A. & Reynar, J. (2008). Building a sentiment summarizer for local service reviews. In *NLPIX2008*, Beijing.

Drozdzynski, W., Krieger, H.-U., Piskorski, J., Schäfer, U. & Xu, F. (2004). Shallow Processing with Unification and Typed Feature Structures - Foundations and Applications. *Künstliche Intelligenz* 1: 17-23.

Esuli, A. & Sebastiani, F. (2006). Sentiwordnet: A publicly available lexical resource for opinion mining. In *Proceedings of the 5th Conference on Language Resources and Evaluation (LREC-06)*.

Haruechaiyasak, C., Kongthon, A., Palingoon, P. & Sangkeettrakarn, C. (2010). Constructing thai opinion mining resource: A case study on hotel reviews. In *Proceedings of the Eighth Workshop on Asian Language Resources*.

Hu, M. & Liu, B. (2004). Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*.

Hummel, R. A. & Zucker, S. W. (1983). On the foundations of relaxation labelling processes. In *PAMI*.

Ku, L.-W., Huang, T.-H. & Chen, H.-H. (2009). Using morphological and syntactic structures for chinese opinion analysis. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*. Volume 3.

Liu, B. (2010). Opinion mining and sentiment analysis. *Handbook of Natural Language Processing*.

Pang, B. & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval* 2(1-2): 1-135.

Popescu, A.-M. & Etzioni, O. (2005). Extracting product features and opinions from reviews. In *Proceedings of the Human Language Technology Conference and the Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP)*.

Steffen, J. (2004). N-Gram Language Modeling for Robust Multi-Lingual Document Classification. In *Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC-2004)*, 731-734, Lisboa. ELRA.

Waltinger, U. (2010). GermanPolarityClues: A lexical resource for German sentiment analysis. *Proceedings of the 7th International Conference on Language Resources and Evaluation (LREC-2010),* Malta.

Destinations Similarity Based on User Generated Pictures' Tags

Alessandro Inversini,
Davide Eynard,
Elena Marchiori, and
Leonardo Gentile

webatelier.net
Università della Svizzera italiana, Switzerland
alessandro.inversini@usi.ch
davide.eynard@usi.ch
elena.marchiori@usi.ch
leonard.gentile@gmail.com

Abstract

Pictures about tourism destinations are part of the contents shared online through social media by travellers. Additional pictures information, such as geo-tags and user description of place, can be used to create groups of similar destinations. This paper investigates the possibility of defining destination similarities based on implicit information already shared on the web. Flickr.com was used as a case study as it represents the most popular picture sharing website. Results show the possibility to group similar destinations based on visual components, represented by the contents of the pictures, and the related tag descriptions.

Keywords: destination similarity, folksonomies, geo-tagging, social media.

1 Introduction

Tourism has always been recognized as an information intensive domain (Gretzel et al., 2000; Buhalis, 2003) where information gathering, processing and distribution is essential for day to day operations (Poon, 1993). Recently researchers (e.g. Gretzel, 2006) demonstrated that Web2.0 and social media are assuming more and more importance within the tourism online promotion. Destination managers are aware that beside website communication different web sources, mostly informal, are spreading the same messages with different strategies (Inversini & Buhalis, 2009).

These sources can be analysed, monitored and exploited with a well-defined strategy to take marketing and selling advantages from them (Inversini et al., 2009). Wise destination managers are already integrating and exploiting social media within their online communications as separated mean or incorporated them within their websites.

Within this scenario user generated pictures are gaining more and more importance (Yoo & Gretzel., 2009) as they informally represent and describe a destination. Further, user generated pictures carry a great amount of information because they are often described by sets of small terms called "tags" and sometimes represent places within a map. These map places are termed geo-located tags or geo-tagged.

2 Literature Review

2.1 Tourism Destinations and Technologies

The continuous development of Information Communication Technologies (ICT) during the last few decades has had profound implications for the tourism industry as a whole (Buhalis, 2000). The increasing importance of technology has influenced not only the way transactions and purchase processes (Werthner & Klein, 1999) have evolved, but also the way communication and promotion of tourism goods have develop on the internet (Buhalis, 2003).

Recently the advent of, the so called, Social Media (Blackshaw, 2006) enabled tourists to share information on the internet within the "read/write web", where the end user has become both information consumer, player (Nicholas, et al., 2007) and provider. Marketing managers and researchers are exploiting new ways to adopt social media in the marketing and promotion arenas in order to take advantage of this "electronic word-of-mouth" (Litvin et al., 2008). Recent studies demonstrated that social media contributes to spreading to web information about destinations using different channels and different strategies (Inversini and Buhalis, 2009) as internet users are in need of communicating their touristic experiences (Inversini & Cantoni, 2009).

The web, especially social media, offers a variety of different platforms to share experiences, facts and even rumours (Blackshaw & Nazzaro, 2006). This information published on popular social media is contributing significantly to the massive growth of information on the web, be it relevant for the end user or not. Furthermore, one important role within information spreading by social media is played out by user generated pictures. Following Yoo and Gretzel (2009) one in two tourists view destination photos via UGC in different web communities. The relevance of pictures in travel both to understand culture (Pengiran-Kaha et al., 2010) and to recommend a place to visit (Linanza et al., 2011) has also been investigated by recent studies. Contents of pictures shared online can be fragmented into different topics, such as nature, products, and facilities (Govers & Go, 2005).

The amount of these experiential-type images can contribute to the online representation of a tourism destination. Pictures shared online can act as a mediated source of information for a prospective consumer, which may influence his/her decision to visit a destination. As in Govers and Go (2005), DMOs can take advantage from online pictures by learning the meanings of pictures shared about a destination and better improve the rich tourism experience that the tourists are looking for. Within social media, user generated travel pictures carry a lot of information. They are often described by sets of small terms called "tags." Once collected the tags build a *folksonomy*. Tags, also, often represent places within a map. These are called geo-located tags or geo-tags.

2.2 The Concept of Folksonomies

The term folksonomy was introduced by Vander Wal (2004), by mixing the terms "folk" and "taxonomy". In practice, users assign a set of terms called tags to an individual piece of content in order to group or classify it for retrieval (Sturtz, 2004).

The result is an informal social network of terms based on users' informal classification of content. The collection of all assigned terms for a piece of content of a single user is called personomy, while the collection of personomies is called folksonomy (Hotho et al., 2006). Some examples of successful folksonomies are delicious.com (formerly known as del.icio.us), Steve (steve.museum), and Flickr (flickr.com). In folksonomies users are not forced to use the same tags; however, users with similar interests tend to converge onto a shared vocabulary with their tags.

One of the factors of success for folksonomies is the fact that no specific skills are needed to participate (Hotho et al., 2006). It is therefore possible to argue that folksonomies invite deliberate and idiosyncratic tagging, also called meta noise, which decreases system utility (Wu et al., 2006). For the general purpose of this research folksonomies have been classified into broad and narrow folksonomies: (i) *Broad folksonomies* emerge in systems whose users can tag any resource, and where a resource can be annotated with many identical tags, one for each user tagging it; in (ii) *Narrow folksonomies* where tags are singular in nature for each object and users are allowed to tag only a limited set of resources, typically the ones they, or their restricted circle of friends, have provided.

2.3 Knowledge from Geotags

Geotagging is the process of annotating objects and online resources with geospatial context information, ranging from specific point locations to arbitrarily shaped regions. These annotations can be explicitly provided by users or extracted automatically (i.e. by analyzing where the user is connecting from, getting GPS data from a mobile device, or extracting geographic metadata from a photo. In online photo sharing communities, user text-based annotation (tags) and location metadata (geotags) often co-exist. Geotag information is typically embedded within picture metadata (stored in EXIF format, see http://www.exif.org/specifications.html). In this report the term "geotag" refers to the geographic annotations (e.g.where a photo was taken) while "tag" is always intended as the textual annotation related to a photo (e.g. "cat").

Several studies have been conducted on extracting knowledge from Flickr georeferenced metadata. Clements et al. (2010) introduces a method to predict similar locations, wormholes, based on human travel behaviour. A wormhole is defined as a similar, but not necessarily spatially close, location on the planet. There are two hypotheses for this:

(i) users have specific travel preferences and therefore visit locations that are similar to some extent; and

(ii) taking a photo in a specific location is an indication that the user likes that location.

From a given target location (L) the algorithm aims to find similar locations around the world. For each user (u), a weight (w) is computed based on the distance of the nearest geo-tagged photo of the user to the target location. Wormholes are then found by aggregating the geotags of all users with W_u as weight per user, and selecting the most relevant positions on Earth according to this metric.

Ahern et al. (2007) show how to analyze tags associated with geo-referenced Flickr images to generate aggregate knowledge in the form of "representative tags" for arbitrary areas of the world. Tags are used to create a visualization tool, "World Explorer" (http://tagmaps.research.yahoo.com/worldexplorer.php), which displays derived tags and original photo items on a world map. Data analysis algorithms are based on multi-level clustering and the scoring of tags is based on TF-IDF (term frequency, inverse document frequency). A user interface shows, for each map region and zoom level, the best-scoring tags for the generated clusters; these tags are shown as text over the map area where each cluster occurs.

3 Research Design

The aim of this research is to describe destination similarity starting from user generated picture tags from the popular photo sharing social media website Flickr.com. In other words this study tries to define similarities amongst a given number of destinations based on unrestricted user descriptions of the destination itself (i.e. the tags associated to the pictures taken at a specific location). Similarities could be used to compile a destinations recommendations list based on pictures shared on social networks by a given users.

Furthermore, the research attempts to understand if additional information related to users (i.e. those who uploaded a given photo) and pictures (i.e. pictures sharing the same tags) is driving the process of recognizing similar destinations. Given the goal, and the above mentioned literature (i.e. Clements et al., 2010; Ahern et al., 2007) the two main research questions are:

RQ1: while finding similar destinations thanks to tags, does picture-related information matter to provide better results?

RQ2: while finding similar destinations, thanks to tags, does users-related information matter to provide better results?

In order to pursue this general goal a holistic methodology, based on the related works, has been designed. Research objectives are similar to the ones exposed in Clements et al. (2010), but they are using a TF-IDF approach similar to the one described in Ahern et al. (2007) in order to extract representative tags describing each place. The designed procedure can be summarized as follows: (i) given a dataset of destination-related tags from Flickr a representative description for each city was extracted; then (ii) the similarities between cities were calculated according to the generated representations; and finally (iii) an interface was provided that, given a city as a query term, showed the cities that were more similar according to the previously calculated similarities. The next section describes the research approach of this study.

4 Research Approach

4.1 A Tag-Based City Representation

The first step to define similarity between cities was to find an appropriate representation for them. In order to develop the model presented below the "document-term" matrix concept -typically used by document indexing software and

search engines- was exploited. According to this representation, documents are considered in terms of the words that appear within them. Similarly, each city was represented as the collection of all the tags assigned to its pictures. For this experiment the tag distribution of Flickr geo-tagged photos belonging to the union of two sets (i.e. the ones related to the top 150 tourism city destinations for the years 2007 and 2008, according to "Euromonitor International": http://www.euromonitor.com/Top_150_City_Destinations_London_Leads_the_Way) was analysed for a total count of 233 cities.

Flickr.com easily allows users to get the top tags for a given location (specified as a WOEID) through its flickr.places.tagsForPlace API. However, this API only returns the top 100 unique tags, without any information about the photos taken or the users who uploaded them. For this reason, two distinct datasets were built: The former, called *Top100*, contains the tags retrieved using the aforementioned API; the latter, called *Random*, contains a random sampling of photo metadata obtained by querying Flickr APIs with YQL (Yahoo Query Language: an expressive SQL-like language that lets developers query, filter, and join data across Web services.) Selecting, for each city, 10 photos from 300 random days, taken at random hours, avoids bias due to day- or time-related events. An important advantage of this second approach is that user- and photo-related information is available, providing new dimensions across which tag analysis can be performed.

The resulting collections also contained tags which were not useful or not related to the destinations (e.g. very common terms like `day`, `dog`, and `friends`, or photography-related terms such as `canon`, `nikon`, and `black&white`). So, a blacklist was heuristically created based on the analysis of the top 1000 tags in Flickr. Tags occurring in only one city were also pruned, because in the VSM they represent terms that will never match when computing the similarity scores. After cleaning the tag vocabulary the *Top100* and *Random* datasets counted, respectively, 23'300 and 55'000 distinct tags. Figure 1 shows the distribution of the top 1,000 keywords. Categories were chosen iteratively during the analysis.

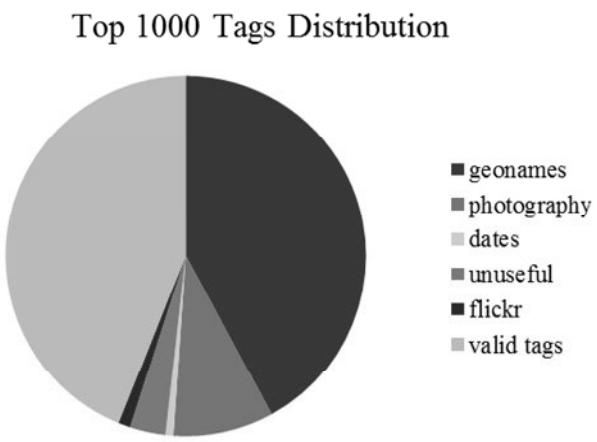


Fig. 1. Tag Distribution Top1000

This shows the top 1,000 tag distribution to be: valid tags – 43,9%; geotags – 42,1%; photography related tags – 9%; other tags (dates, Flickr related and non-useful tags) – 5%. For the purpose of this study photography related and other tags (dates, Flickr related and non-useful tags) were not considered.

4.2 Calculation of Similarity between Cities

The VSM (Vector Space Model) was used to represent cities in terms of their related tags. In the VSM each city is represented by a vector in an *n*-dimensional space, where *n* is the number of distinct tags, whose components are weighted according to tag frequency. Of course, as tags which are too popular tend to be more widespread and thus less informative, frequency is normalized using the TF-IDF approach, which normalizes a TF (Term Frequency) with an IDF (Inverse Document Frequency) factor. This takes into account the number of documents containing that term (in our case, the number of cities for whose photos a given tag has been used). In the *Top100* dataset the standard IDF has been calculated for normalization:

$$IDF_i = \log_2 \frac{|D|}{|D_i|} \tag{1}$$

where $|D|$ is the number of destinations used in the analysis (233) and $|D_i|$ is the number of destinations containing the tag t_i. For the *Random* dataset two more variants of IDF have been calculated, taking advantage of the additional information we were able to gather. The former variant, called IDFP, exploits the extra information coming from the tagged photos:

$$IDFP_i = \frac{|P|}{|P_i|} \tag{2}$$

where $|P|$ is the total amount of photos in the *Random* dataset and $|P_i|$ is the number of photos tagged with t_i. The latter variant, called IDFU, exploits information about the users who have tagged photos about each specific destination:

$$IDFU_{i,j} = \frac{|P_j|}{|P_{i,j}|} \frac{|U_{i,j}|}{|U_j|} \tag{3}$$

where $|P_j|$ is the number of photos for the destination j, $|P_{i,j}|$ is the number of photos for the destination j tagged with t_i, $|U_{i,j}|$ is the number of users tagging destination j with the tag t_i, and $|U_j|$ is the number of distinct users who tagged at least one photo in j. We calculated tag weights for each city by normalizing tag frequencies in four different ways: (System A) Top100 dataset, standard IDF; (System B) Random dataset, standard IDF; (System C) Random dataset, IDFP, (System D) Random dataset, IDFU. For each different normalization, we then calculated similarities between cities in terms of the cosine distance between their matching vectors:

$$\text{similarity}(a, b) = \cos(\alpha) = \frac{a \cdot b}{|a||b|} \tag{4}$$

4.3 Front-End Implementation

The interface was based on an existing open source project, Geoplanet Explorer (http://isithackday.com/geoplanet-explorer/). It was implemented as a PHP Web based application accessing a MySQLdatabase.

4.4 System Evaluation

To evaluate the system, an online survey was created in order to ask users which of the four similarity measures (i.e. System A - Top100 dataset, standard IDF; System B -Random dataset, standard IDF; System C - Random dataset, IDFP; System D - Random dataset, IDFU) was the best one according to them. The survey was composed by a demographic section and a group of cities to be evaluated. Users were chosen randomly by posting the survey link on popular social networks. After the demographic section the survey proposed to users clear instructions explaining that the similarities between places are not necessarily geographic and are based on place descriptions (i.e. on harvested information – Figure 2).

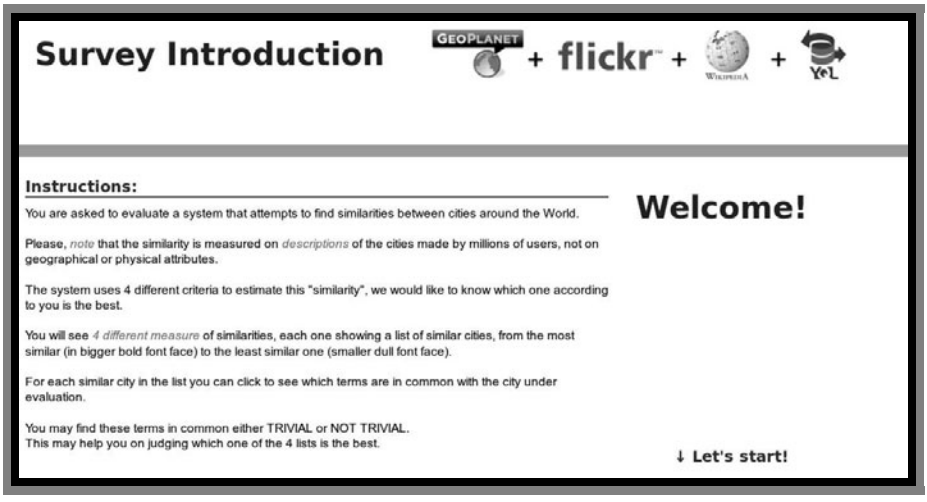


Fig. 2. Welcome Page and Survey Instructions

A small group of cities has been selected for the survey with a high percentage of European tourism destinations. Additionally a small group of cities from Asia, South America and the USA was inserted in the test. Each user judged one (random) city at a time, for a total of five distinct cities. For each of them the user was provided results of the four scoring systems, in the form of four lists containing the top-five related cities.

5 Results

The survey user interface was designed to give basic information about the city (i.e. country, administrative regions, map, etc.), in order to let users easily identify the tested city (in this case Rome, Figure 3). In the front end the sections are clearly defined: (1) information about the place taken from Yahoo! GeoPlanet

(http://developer.yahoo.com/geo/geoplanet/), which is used to disambiguate the city from its homonyms (e.g. Rome in Georgia, US); (2) main city pictures from Flickr.com; (3) Map from Yahoo! maps; (4) the four systems to be rated; (5) the most popular 100 tags from Flickr.com.

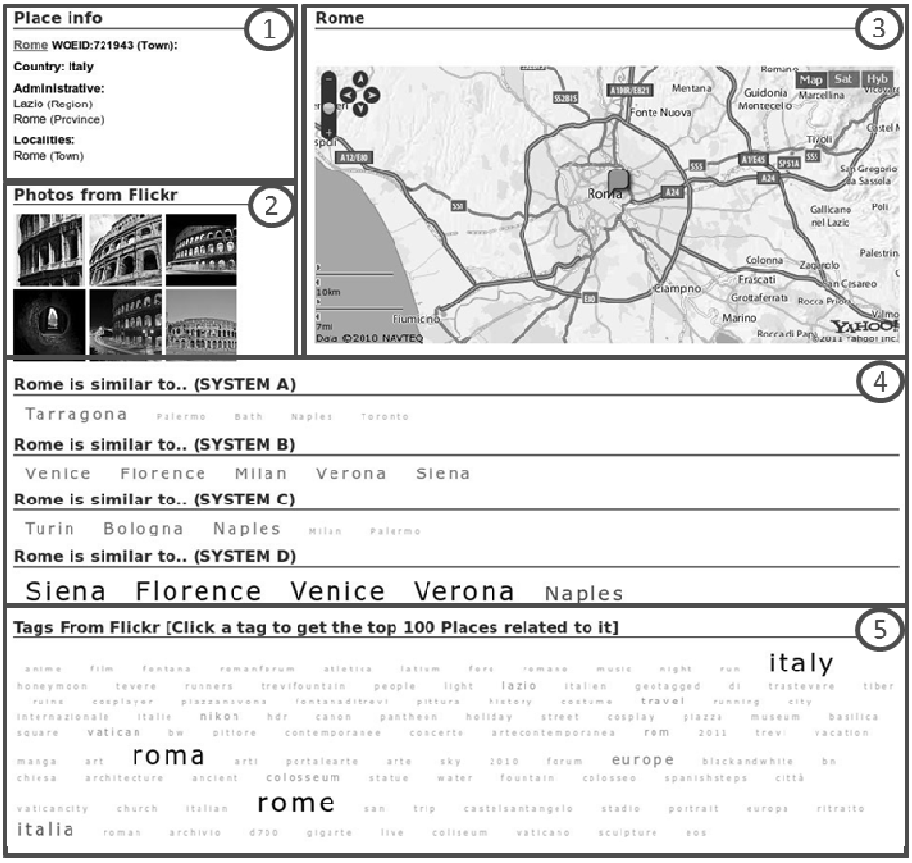


Fig. 3. Graphic User Interface

In section 4, users could easily understand not only the similarity ranking, but also (to some extent), the *level* of similarity thanks to the differing font sizes (the bigger the font the higher the similarity explained on the survey starting page). Since the similarity between two places is based on the similarity of their descriptions, users had the possibility to check the tags in common between two cities. This is clearly a simplification, as the adopted similarity metrics were not based on a simple term match. It was considered useful to provide a rough idea of why two cities had been considered similar.

5.1 Evaluation

The survey was filled out by 113 users, mostly master students from Università della Svizzera italiana and Politecnico di Milano. Users produced 516 valid answers. The final result (Figure 4) yielded a great deal of useful information.

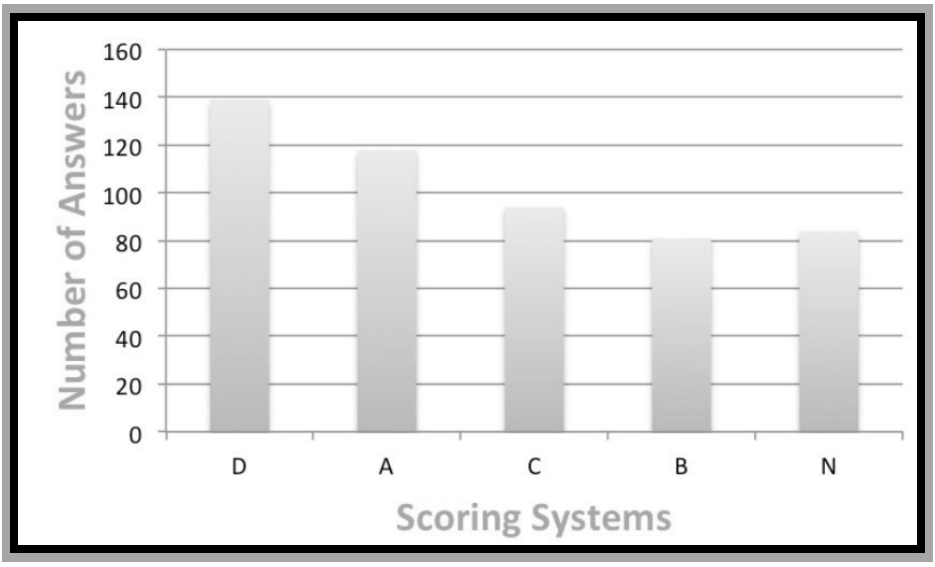


Fig. 4. Survey Results

The best system according to users is System D (n=139 preferences). System D is based on the weighting scheme in which a greater importance was given to the user factor.

System A represents the second most popular answer (n=118 preferences) and it shows that we can extract valuable knowledge gathering information even from the *Top100* dataset. In other words, using only the most 100 popular tags for a city and the classical IDF we can still have good performance, even better than other methods applied to the bigger dataset.

The random dataset based on picture information (System C) ranked third according to users' preferences in the survey. Finally, system B, differing from A only for the dataset used, ranked last in users' preferences. "N" represents the "no answer" group (n=84). Returning to the above mentioned research questions, it is thus possible to claim that (RQ2) the additional information coming from user tagging activities (i.e. who tagged a photo with a given tag) is much more relevant, while defining the similarity of two places or destinations, than (RQ1) picture-related information (i.e. which pictures have been tagged with a given tag).

6 Discussion and Conclusion

Surprisingly System A, relying on the much more limited *Top100* dataset, ranked better than other two systems (i.e. Systems B and C) that were using an extended

sample based on the random harvesting of picture tags obtained with Yahoo Query Language. This means that the top 100 tags Flickr.com allows to be downloaded with its API are representative enough of the destinations and can be used to compare them with discrete results. Furthermore, the relevance of user-related information seems to be a driver to define similar cities/destinations.

This confirms the validity of other approaches (such as the "Wormholes" one) and provides a useful insight for the development of new tag- and user-based recommender systems. Actually, with the current model and technology this system is already able to suggest and/or recommend to users, which own a collection of pictures of a given place/destination in a popular social network such as Flickr.com, to visit other places/destinations without asking them any additional information about their preferences.

7 Limitations and Future Work

Limitation includes three issues. First, it was assumed that random sample was more precise in shaping destination similarity, but from the results of this research it does not seem to provide more information. Its real advantage, instead, is the additional information about users and photos that can be exploited to provide a better similarity measure. Second, demographic data was asked within the survey but not used in the final evaluation due to a shortfall of the system. Finally, as the current research was a starting point for a more in-depth research a light evaluation methodology with a snowball sampling was used.

Future work will cover these limitations. It will study, in depth, the difference between top 100 samples and random samples using both the case information about users and information about pictures to define if these preliminary findings can be confirmed. The demographic data will be added within the evaluation phase to try and leverage users' travel preferences and history. Finally a better sampling methodology will be used.

References

Ahern, S., Eckles, D., Good, NS. & King, S. (2007). Over-exposed?: privacy patterns andconsiderations in online and mobile photo sharing. In *Human Factors in Computing Systems*.

Blackshaw, P. (2006). *The consumer-generated surveillance culture*. Retrieved October 13, 2008, from http://www.clickz.com/showPage.html?page=3576076.

Blackshaw, P. & Nazzaro, M. (2006). *Consumer-generated media (cgm) 101: Word-of-mouth in the age of the web-fortified consumer*. Retrieved August, 1, 2007.

Buhalis, D. (2000). Marketing the competitive destination of the future, *Tourism Management* 21(1): 97-116.

Buhalis, D. (2003). *eTourism: Information technology for strategic tourism management*. Harlow: Prentice Hall.

Clements, M., Serdyukov, P. & De Vries, A. (2010). Finding Wormholes with Flickr Geotags. *Advances in Information* 5993: 658–661.

Govers, R. & Go, F. (2005). Projected Destination Image Online: Website Content Analysis of Pictures and Text. *Information Technology and Tourism* 7: 73-89.

Gretzel, U. (2006). Consumer generated content - trends and implications for branding. e-*Review of Tourism Research* 4(3): 9-11.

Gretzel, U., Fesenmaier, D. R. & O'Leary, J. T. (2006). The transformation of consumer behaviour. In Buhalis, D. and Costa, C. (Eds.), *Tourism Business Frontiers*. Oxford: Butterworth-Heinemann.

Gretzel, U.,Yu-Lan, Y. & Fesenmaier, D. (2000). Preparing for the New Economy: Advertising Strategies and Change in Destination Marketing Organizations. *Journal of Travel Research* 39(2): 146-156.

Hotho, A., Jaschke,R., Schmitz, C. & Stumme, G. (2006). Information retrieval in folksonomies: Search and ranking. In *Proceedings of the 3rd European Semantic Web Conference*, Lecture Notes in Computer Science. Springer.

Inversini, A. & Buhalis, D. (2009). Information Convergence in the Long Tail. The Case of Tourism Destination Information. In W. Hopken, U. Gretzel and R. Law (Eds.), *Information and Communication Technologies in Tourism 2009* - Proceedings of the International Conference in Amsterdam, Netherland (pp. 381-392). Wien: Springer.

Inversini, A., Cantoni, L. & Buhalis, D. (2009). Destinations' Information Competition and Web Reputation. *Journal of Information Technology & Tourism* 11: 221-234.

Inversini, A. & Cantoni, L. (2009). Cultural Destination Usability: The Case of Visit Bath. In W. Hopken, U. Gretzel and R. Law (Eds.), *Information and Communication Technologies in Tourism 2009* – Proceedings of the International Conference in Amsterdam, Netherland (pp. 319-331). Wien: Springer.

Linaza, A., Agirregoikoa, A. Garcia, Torres, J. & Aranburu, K. (2011*)*. Image-based Travel Recommender System for small tourist destinations. In R. Law, M Fuchs, F. Ricci (Eds.), *Information and Communication Technologies in Tourism 2011* - Proceedings of the International Conference in Innsbruck, Austria (pp. 1-12). Wien: Springer.

Litvin, S. W., Goldsmith, R. E. & Pan, B. (2008). Electronic word-of-mouth in hospitality and tourism management. *Tourism Management* 29: 458-468.

Nicholas, D., Huntington, P., Jamali, H.J. & Dobrowolski, T. (2007). Characterizing and evaluating information seeking behavior in digital environment: spotlight on the bouncer. *Information processing and Management* 43(4): 1085-1102.

Pengiran-Kahar, D.I.N.F., Syed-Ahmad, S.F., Ismail S.H.S. & Murphy, J.(2010). Shared Arabian Muslim Travel Photos. In Gretzel, U., R. Law, M Fuchs (Eds.), *Information and Communication Technologies in Tourism 2010* - Proceedings of the International Conference in Lugano, Switzerland. Wien: Springer, 591-603.

Poon, A. (1993). *Tourism, Technology and Competitive Strategies*. Wallingford, CT: CAB International, Oxford.

Sturtz, D.N. (2004). *Communal Categorization: the Folksonomy*. INFO622: Content Representation, December.

Vander Wal, T. (2004). *You down with Folksonomy?* - Reteived March 28, 2008, from http://www.vanderwal.net/random/entrysel.php?blog=1529

Werthner, H. & Klein, S. (1999). *Information Technology and Tourism – A Challenging Relationship*. Wien -New York: Springer Verlag.

Wu, H. Zubair, M. & Kurt, M. (2006). *Harvesting social knowledge from Folksonomies*. Odense, Denmark: HT.

Yoo, K.H. & Gretzel, U. (2009). Trust in travel-related consumer generated media. In W. Höpken, U. Gretzel and R. Law (Eds.), *Information and Communication Technologies in Tourism*, Wien: Springer, 49-59.

What's in a Travel Review Title?

Silvia De Ascaniis [a], and
Ulrike Gretzel [b]

[a] Webatelier.net
Faculty of Communication Sciences
University of Lugano - USI, Switzerland
silvia.de.ascaniis@usi.ch

[b] Institute for Innovation in Business and Social Research
University of Wollongong, Australia
ugretzel@uow.edu.au

Abstract

When looking for online travel reviews (OTRs), users have to make a quick first choice to select reviews which seem to be most relevant. Some users might only browse through the titles without ever looking at the complete review text. OTR titles thus constitute important metadata, but no research on OTRs has specifically looked at titles. This paper seeks to provide a first exploration of the issue, considering the average length of OTR titles, their level of informativeness, the indication they provide on the review orientation, the word diversity and their communicative function. Understanding OTR titles is also of practical relevance as it can help with developing algorithms for the automated selection and classification of OTRs.

Keywords: online travel reviews; communicative functions; text analysis; UAM Corpus Tool.

1 Introduction

Information search involves time and effort and is limited by human cognitive capacity to process incoming information. Therefore, when looking online for information about a certain hotel, attraction, tourism service or, generally, a destination, people do not go through everything they retrieve from their search but select online resources they think can meet their needs. In order to do so, they have to make judgments regarding the search results. These judgments have to be quick, given the amount of information to be processed and, thus, are typically made based on a first impression of the results (Marchionini, 1995; Wöber, 2006). To support such quick judgments, information retrieval systems display metadata such as title, URL, date, etc. These metadata serve the function of overview and preview (Balatsoukas, Morris, & O'Brien, 2009) and contribute to the creation of a first impression of an online resource (Pan & Fesenmaier, 2006; Xiang & Fesenmaier, 2006). The first impression is so crucial for online information search that a number of studies have investigated the decisive factors contributing to its creation (Kim & Fesenmaier, 2007, 2008).

The same scenario holds for online travel reviews (OTR). Since there is usually a huge number of reviews available for the same product, service or destination, the user has to make a quick first choice to select those reviews which seem to be most relevant. Some users might only browse through the titles without ever looking at the complete review text. Thus, just like the titles in bibliographic searches, OTR titles constitute important metadata. The general function of titles in whatever type of text

is to give anticipation about the text itself, like introducing the topic, taking a position, expressing a judgment. The title of an OTR, therefore, should give an indication of either the author's experience with the tourism service/product being reviewed, or a description of the service or product itself. However, OTR titles might have additional communicative functions and specific structures that distinguish them from other types of titles. No research on OTRs has specifically looked at titles and this paper seeks to provide a first exploration of the issue. In addition to providing theoretical contributions to the fields of communication, linguistics, and social media research in tourism, understanding OTR titles is of practical relevance for at least two reasons: first, OTR titles are important to consider in social media monitoring. They provide insights into how customers summarize experiences and show what first impressions other customers will likely get of an establishment. Second, identifying semantic and linguistic characteristics of OTR titles is a first step in developing algorithms for the automated selection and classification of OTRs. Accordingly, the paper intends to investigate some basic features of OTR titles, specifically answering the following questions:

a) How long are OTR titles and can systematic differences in length be discovered based on certain OTR characteristics?
b) Are OTR titles informative, i.e. do they help the reader in figuring out what the review is about?
c) Do the titles forecast the general orientation of the review, i.e. do titles with a positive/negative connotation correspond to reviews with a positive/negative rating?
d) How diverse is the information contained in OTR titles for the same product or service, and which are the words and constructs most frequently used?
e) Which communicative functions can be identified?

2 Theoretical Foundations

2.1 Background

A review is a specific textual genre having its proper characteristics, which are determined by its intended function (Rigotti & Cigada, 2004; Rocci, 2005). An Internet user expects that a travel review provides him/her with information and gives recommendations about a place or service. The online platform directly influences the shape of the text, making explicit what its function should be. *TripAdvisor*'s motto, for instance, states that it "provides unbiased reviews, articles, recommendations and opinions on 'destination X' vacations", and then specifies in the guidelines what the review should be about and what it should avoid to say. However, while it gives elaborated indications on what the review text should or should not contain, the instructions for titles are rather vague. The instructions for hotel reviews include an example of a recommendation. For attractions, the title tip "If you could say it in one sentence, what would you say?", asks for a one sentence synthesis of the attraction experience; it is then up to the author to formulate the synthesis as a recommendation ("Go there!"), as a direct evaluation of the place ("It is not worth a visit"), as a list of the main features of the attraction ("[The Colosseum has] Roman ruins, ..."), as a short description of the personal experience at the place ("A boring walk under a burning sun"), or using other kinds of formulations. The indications given by the

website to write reviews put certain constraints on the author's freedom of speech, but they also assure that the review meets minimum standards of acceptability. Besides the intended function suggested by the platform, authors have their own motivations when writing a review. The motivations of writers have been studied in the context of travel-related virtual communities (Wang & Fesenmaier, 2003) and specifically for travel reviews on TripAdvisor (Yoo & Gretzel, 2008). The main motivations include altruism (helping companies and other consumers), reciprocity (giving back to a company that provided a good experience as well as giving back to the platform), and hedonism (sharing enjoyment).

2.2 Communicative Function of Travel Review Titles

The general principle of human communication is that every communicative act is produced to address an interlocutor (Grice, 1989). It means that the first main goal of every communication is to reach its addressee: I write because I want or I expect someone to read, I speak because I want or I expect someone to listen. Thus, different strategies are employed to reach the goal of catching the interlocutor's attention. In the same vein, it can be said that the overall reason for writing a review is for it to be read, on the assumption that what is reported may be of interest to the reader. It is therefore meaningful to ask which communicative strategies are used by the authors of OTRs to make their texts attractive for an unknown audience.

The role of titles in OTRs can be compared to the role of headlines in newspapers or the role of taglines or slogans for an advertisement. In the literature, newspaper headlines have been frequently characterized as short summaries of the respective news items (Van Dijk, 1988; Bernstein & Garst, 1982), and they have been analysed according to their semantic-referential function (Bell, 1991). Headlines were classified, therefore, according to categories like: 'quotation headlines' when they contained a quotation, 'highlight headlines', when they focused on a single detail of the story, 'summarizing headlines', when they summarized the story. This kind of classification, however, does not take into consideration the pragmatic function of headlines, which indeed guides readers to select the stories which are worth for them to be fully read. In a recent study about the communicative function of newspaper headlines Dor (2003, p. 720) argues that they should to be considered "*communicative devices* whose function is to produce the optimal level of affinity between the content of the story and reader's context of interpretation, in order to render the story optimally relevant for the reader". He builds upon Sperber and Wilson's technical notion of *relevance*, and develops the notion of *relevance optimization*. A newspaper headline reaches the optimal level of relevance if it allows the reader to achieve the greatest possible contextual/cognitive effects (that is a change in the individual's set of prior assumptions) with the smallest processing effort. A headline which is long and full of details or is ambiguous, for instance, requires a huge reading effort, while a headline which does not contain new information or is not interesting for the target readers, requires a low reading effort but does not carry valuable contextual effects. According to Dor, the function of newspaper headlines is not – at least is not primarily – that of attracting the attention of the reader and provoking him to go on reading the whole story, as claimed by Bell (1991), but to help the reader to receive the best deal of information with the least cognitive/procedural effort. He argues that this explanation better fits the modern behavioural patterns manifested by newspaper

readers, who "spend most of their reading time *scanning* the headlines rather than reading the stories" (2003, p. 697). His approach goes beyond the usual definition of newspaper headlines as summaries of the news item and encompasses the usual strategy adopted in the literature to analyse them based on their semantic function but also considers their pragmatic function.

It may be argued that news headlines should also include some elements of persuasion beyond the tension for a perfect balance between informativeness (i.e. contextual effects) and readability (i.e. processing effort). They should be able to catch the reader's curiosity to convince him/her to read the correspondent news item. Since the first impression is one of the factors which influences people's behaviour, the appeal of news headlines and, first of all, of the headlines of front-page news, plays an important role. The significance of this persuasion factor has been mainly studied in conjunction with taglines and slogans in advertisements. If a tagline is able to catch the consumer's attention, the probability that the product or service being advertised will be remembered increases. The tag phrase creates the first impression and is often remembered even before the product or service itself (Freeman, 2005). Considering the effects on the market, to write persuasive taglines has become almost a science which continues to receive attention in marketing and advertising studies. Some studies have investigated the influence of culture in brand and slogan design (Aaker & Maheswaran, 1997; Burgoon et al., 1982; Glenn, Witmeyer & Stevenson, 1977), others have focused on the executional elements (such as type of appeal, colour, illustration, layout) (Ernst, 1980; Stewart & Furse, 1986; Laskey, Fox & Crask, 1994), and a few others on commercial message strategies (such as claim of uniqueness, claim of superiority, focus on user, focus on brand personality) (Yssel & Gustafson, 1998; Laskey, Fox & Crask, 1994), to analyse and evaluate their appeal.

OTR titles share with news headlines the fact that they should anticipate something about what is reported in the text. Speaking with Dor (2003), well written titles should allow the reader to achieve a valuable contextual effect without an excessive processing effort. They share with taglines the need for persuasion in that an OTR titles aims to attract the reader's curiosity in order to convince him/her that what is written is of interest. Differently from taglines, however, OTR titles do not need to be remembered because they are not associated with a brand, so their persuasion goal may not be as explicit.

3 Methodology

3.1 Data Collection

In order to investigate the characteristics of OTR titles, a corpus was built, based on reviews published on *TripAdvisor* (www.tripadvisor.com). Three city destinations were chosen assuming that they represent somewhat comparable tourism products, following the only one requirement of being quite big and well-known to have the opportunity to collect a high number of review titles for each one of them. They are: Rome, Paris and Sydney. Destination reviews were retrieved from the *Travel Guide* section of the *TripAdvisor* destination page (http://www.tripadvisor.com/AllReviews-g187791-Rome_Lazio.html), where all the reviews related to that destination are published. Only the reviews about the destination itself were considered, leaving out

the reviews that talked exclusively about a specific type of accommodation, attraction, etc. Reviews on the destination are listed by *TripAdvisor* as the very last group (here referred to as the "destination group") of reviews in the *Travel Guide* section. Only reviews in English, German, French and Italian were included. The coders identified and excluded those reviews that seemed to have been posted in the wrong place, e.g. those not focusing on the destination itself but on other elements of the tourism experience (e.g. hotels or transportation means). The data collection phase was completed September 3rd, 2011. The corpus is described in Table 1. The coding was done by two coders, who discussed the transparency and pertinence of categories, as well as ambiguous or borderline cases.

Table 1. Description of the Corpus

	Rome	**Sydney**	**Paris**	**Total**	
Total N° of reviews in the *Travel Guide* section	9.734	2.231	10.084	22.049	
N° of reviews in the "destination group"	1.012	240	887	2.139	
N° of reviews after the filtering	**701**	**173**	**600**	**1.474**	**% of tot N°**
Reviews in English	470	109	345	924	62,7%
Reviews in German	134	52	119	305	20,7%
Reviews in French	33	0	46	79	5,35%
Reviews in Italian	64	12	90	166	11,25%
N° of reviews with rating 4 (very good) or 5 (excellent)	589	169	504	1262	85,6%
N° of reviews with rating 3 (average)	34	3	46	83	5,6%
N° of reviews with rating 2 (poor) or 1 (terrible)	78	1	50	129	8,8%

3.2 Coding Procedure and Analysis

The OTR titles were classified according to three categories: *language*, *rating*, *connotation*. *Rating* refers to the overall rating provided by the review author for the destination (1= terrible, 2= poor, 3= average, 4= very good, 5= excellent). *Connotation* refers to the overall orientation of the title, if isolated from the rating and the review main text; the coder had four choices: positive = the title had an overall positive connotation, negative = the tile had an overall negative connotation; not-applicable = the title was not univocally polarized or was ambiguously formulated (ex. "This is a Roman holiday?", "Dirty city, wonderful sights", "Rome is great but full of thieves"); neutral = the title had not any recognizable evaluative connotation (ex. "Our weekend in Rome", "Tips on Sydney", "Paris with kids"). The corpus was coded manually, for the main reason that the category *connotation* required the understanding of the semantic nuances of titles, which is hardly precisely caught by an automated tool (O'Donnel, 2008; Colombetti, 2001). The meaning of a text is, in fact, linguistically built through the interconnection of the logical-semantic level and the pragmatic level, that is through the words used and the syntax as well as the context in which it is communicated (Rigotti & Rocci, 2001). In addition, since we

are interested in the role of titles to create first impressions, manual coding is justified as the coders were representative of actual Internet users who are able to grasp titles' connotation thanks to their proficiency in a language.

The corpus was coded and analysed using *UAM Corpus Tool* version 2.8.3, a software for human and semi-automatic annotation of texts and images, developed by Mick O'Donnell[1]. UAM allows the user to explore linguistic patterns and linguistic features in a text which cannot be explored with simple concordances, and which cannot be automatically tagged because they pertain to the semantic or pragmatic level. The central concept of UAM is the "project", which consists of a corpus of text files, which can be annotated at a number of linguistic layers: at the document layer, if one wants to assign features to the document as a whole (e.g. date, author, text type), at semantic-pragmatic or syntactic layers, for features pertaining to segments within the text, or at the lexical layer, if the feature characterizes single words. For each layer, the user can provide a hierarchically-organized tagging scheme. Our corpus was annotated at the semantic-pragmatic level, using the same scheme for the three different destinations. Within each layer each review title was annotated three times, this way representing three different segments, each corresponding to one of the features we were interested in, i.e. language, rating, connotation. The *UAM Corpus Tool* also provides other functionalities to support the user, including basic statistics. Two kinds of analyses can be performed on the corpus: *general text statistics*, which offers basic statistics of the different text files in the corpus, such as the total number of segments, the number of words per segment, the average segment length, lexical density; and *feature usage*, which allows to specify a feature in a layer and describe its usage in the corpus at that layer in terms of count, mean and standard deviation (e.g. the feature *language* for the layer *Rome*).

To answer questions regarding the informativeness of OTR titles, the current study used *lexical density* (LD) as a measure. The Dictionary of Language Teaching and Applied Linguistics (Schmidt, Platt & Platt, 1995, p. 163) defines lexical density as "a measure of the ratio of different words to the total number of words in a text". It is a measure of the density of information in any passage of text, according to how tightly the lexical items (content words) have been packed into the grammatical structure (Halliday, 1994). LD can be measured as the number of lexical items (*Nlex*) divided by the total number of words per clause (*N*): $Ld = (Nlex / N) \times 100$. Lexical items include nouns, regular/main verbs, and most adjectives and adverbs. In any piece of text or discourse there are variations in LD from clause to clause but, generally, the LD of spoken language tends to be lower than written language, since writing is more planned and formal. It varies also from language to language, according to lexical and grammatical specificities. In this paper, the assessment of LD will be limited to English titles, because at present UAM can only describe LD for English texts. Since English titles in the corpus amount to 62.7% of the total (i.e. 924 out of 1,474 titles), results can still be considered relevant for a general discussion. In contrasting written

[1] The software can be freely downloaded from the website: http://www.wagsoft.com/CorpusTool/. The description of the software is based on the Version 2.7 User Manual (February 2011) and on the following paper by Mick O'Donnel: *The UAM Corpus Tool: software for corpus annotation and exploration*. Proceedings of the XXVI Congreso de AESLA, Almeria, Spain, 3-5 April 2008.

and spoken versions of the same text in English, Eggins (1994) found that on average the spoken text was 33% lexical, while the written version was 42% lexical. Since the main function of titles in a text is to anticipate something about the text in a concise but effective way, it may be expected that the average LD for OTR titles is higher than the average LD found by Eggins for written text.

The *Corpus Search* functionality of UAM Corpus Tool allows the user to search for instances in the annotated corpus matching some criteria (e.g. search for the feature *positive* within the segments tagged with the feature *rating 4* or *rating 5*), for segments containing strings (search for the string "*eternal*" in all the segments of the layer *Rome*), and for lexical patterns (e.g. search for *second person pronouns*). It is also possible to make searches across layers. This function allowed us to hypothesize the representativeness of certain lexical patterns for specific communicative functions of review titles. One last functionally of the tool that we used to investigate the corpus is called *Explore*. It allows compiling the absolute frequency of words, the frequency of keywords and the frequency of key-phrases in the corpus. The absolute frequency of words lists usually place on top words such as "the", "of" or "and", while the keywords and key-phrases frequency orders words or phrases (called *n-grams*) in terms of their "specialness" for the corpus, that is in terms of how important each word is for a specific corpus when compared with other corpora.

4 Results

The results of the study will be presented in the form of answers to the questions listed in the Introduction section.

4.1 Title Length

The average length of OTR titles is 4.4 words per title, with a maximum length of 17 words. The average length of titles with a positive connotation (tot. 887/1474) is 4.4 words while for titles with a negative connotation (tot. 143/1474) it is 4.9, for titles with a neutral connotation (tot. 143/1474) it is 3.8 (350/1474), and the average length of titles with a not-applicable connotation (94/1474) is 5.4. The average length of titles was 4.5 for Rome, 4.3 for Sydney and 4.4 for Paris.

The average number of words used for titles of positive (rating = 4 or 5) reviews (1262/1474) is 4.3, while that for titles of negative (rating = 1 or 2) reviews (129) is 4.6 and that for titles with average (rating = 3) reviews (83) is 4.6 as well. These results show that, on average, people use less words for introducing a positive experience at a destination than for depicting a negative or average impression.

4.2 Informativeness

As mentioned above, results regarding the lexical density are limited to English titles, because at present UAM can only describe LD for English texts. The average number of lexical items per English title is 3.0 out of 4.6 average number of words per English title. This means that nearly 3 out of 5 words are content items (versus grammatical items), that is, words which communicate univocal information. The average lexical density is 65%; as expected, it is higher than Eggins' (1994) results for general written English, which was found to be 42% lexical. However, lexical

density might not be a fully representative measure for very short texts like OTR titles; thus, rather than expressing the level of informativeness of titles, it indicates their readability, that is how easy it is for a reader to understand their content.

4.3 Indication of Review Orientation

Table 2 shows that there is an overall concordance between titles with a positive or a negative connotation and, respectively, titles of positive or negative reviews. Within the titles with a negative connotation, most of the reviews have an average rating (3) rather than a negative rating (1 or 2). Titles with a neutral connotation, thus giving a non-evaluative indication about the content of the review, pertain mostly to positive reviews. The significance of the relationship between rating and title connotation has been calculated using Chi-square statistics comparing connotations for OTRs with very good or excellent ratings (set 1) and OTRs with poor or terrible rating (set 2) (Table 3). A calculation at the city level is necessary because of the way the data layers can be coded in the software. Sydney had to be excluded from this analysis as 98% of its reviews were positive. The results show that titles with positive connotations are significantly linked to reviews with positive ratings and titles with negative connotations significantly correspond to reviews with negative ratings.

Table 2. Comparison between OTR Title Connotation and OTR Rating

Review rating	**Title Connotation**			
	Positive	Neutral	Not-applicable	Negative
4 (very good) or 5 (excellent)	**98.0%**	**87.7%**	**68.1%**	14.7%
3 (average)	1.2%	7.4%	16.0%	**63.0%**
1 (terrible) or 2 (poor)	0.8%	4.9%	16.0%	**22.4%**

Table 3. Chi-Square Results for Title Connotation by OTR Rating

Connotation	Chi-Square Results	
	Rome	Paris
Positive	126.1**	66.2**
Negative	377.9**	282.0**

**p<.02

4.4 Word Diversity

Table 4 shows that OTR titles make strong use of superlatives (wonderful, schönste, favorite, great city, the best in the world), and slogans (Rome the eternal city, Paris the city of love) and that positive words are much more frequent than negative ones, which indeed do not appear at all in the top keywords list. There might be a persuasive reason for that: superlatives describe something that exceeds normality, slogans sketch images that catch the attention and words with a positive connotation recall positive experiences that everyone desires for themselves. Besides that, many titles try to characterize the destination highlighting one of its features, like for Rome the fact of being full of history or for Paris the romantic atmosphere. Among the frequent keywords in OTR titles are also *meta-words* pertaining to the experience of

travelling, such as “tour”, “visit” or “time”. In general, however, the frequency counts are very low, suggesting a great diversity of words being used in the titles.

Table 4. Word Frequencies

Token	Total frequency	Rome	Sydney	Paris
Rome/Rom/Roma	316	316	0	0
Paris/Parigi	288	0	0	288
Sydney	62	0	62	0
City/Stadt/città/ville	324	147	49	128
Amazing/great/wonderful/toll	93	50	16	27
Beautiful/schöne/bella/belle	64	44	10	20
Sydney	62	0	62	0
World/welt/mondo/monde	57	22	22	13
Favorite/favourite	55	27	7	21
Place	46	20	12	14
Eternal/ewige/eternelle/eterna	43	43	0	0
Love/liebe/amour/amore	41	11	0	30
Trip	37	17	2	18
Italy/Italien/Italie/Italia	33	33	0	0
Visit	32	19	4	9
Tour	32	10	3	19
Romantic/romatisch/romantica	27	8	0	19
France/Frankreich/Francia	25	0	0	25
Days	23	12	4	7
Time	22	12	5	5
History/Geschichte/histoire/storia	21	18	0	3

Table 5. Most Frequent Key-Phrases

2-grams/3-grams/4-grams	Frequency
(Rome) the eternal city/die ewige Stadt/città eterna	45
My favorite city/place	33
(Paris) city of love/Stadt der Liebe/città dell’amore	23
The most beautiful city/die schönste Stadt	20
immer eine Reise wert	20 (out of 305 German r.)

4.5 Communicative Function

Titles may give *temporal indications* to relate the experience to a specific period during the year (e.g. “Rome in Winter”); describe the time-frame of the travel experience (e.g. “Paris in three days”); express a judgment about the experience (e.g. “Fun times in Rome”). Cues for temporal indications in the text are *temporal nouns* and *temporal adverbs*. In addition, the author may wish to highlight the *evaluation* of the experience or of the destination. *Evaluative adjectives*, especially superlatives, are frequent in the titles and usually reveal the polarization of the review, as in: “Great city”, “Rude citizens”, or “The best kept secret”. Some evaluative adjectives also *describe* an aspect of the destination: “Romantic Paris”, for instance, not only gives a positive judgment about the city, but also characterizes it. Comparative adjectives, on the other hand, tell something about *expectations*; the emblematic example is “better”,

in titles like "Better than expected". The role of indicators for the experience or destination evaluation may also be played by some verbs; among them are *judgmental verbs*, like "Loved the Roman ruins". If these verbs occur together with *first person-pronouns*, the evaluation takes the nature of a *personal judgment*, as in "We loved Rome!". One last group of indicators which deserves a mention includes words pointing to *direct advice/recommendation*. The advice function of a title may be indicated by *second-person pronouns*, *verbs oriented to the addressee* and verbs in the *infinitive form*, e.g. "Go to Paris, or you won't know what it is really like!".

Table 6 reports some examples of linguistic indicators and the correspondent codification using the UAM Corpus Tool. At the moment the tool only allows searches for lexical patterns or features in English. Also, results for some lexical features (e.g. verbs in the infinitive form) need to be double-checked by the researcher, because the tool does not distinguish between their use as verbs and their use as nouns (e.g. "fun" vs. "to have fun"). Frequencies represent the total number of occurrences of the lexical feature in the corpus. The results indicate that descriptive words are much more common than comparative words or words that would suggest a direct advice addressed to the reader. Similarly, first person pronouns are more common than second person pronouns, indicating that most titles represent a summary statement of a personal experience or a general destination description.

Table 6. Linguistic Indicators of Title Functions

Linguistic indicator	Lexical feature for concordance searching in UAM Corpus Tool	Examples	Total frequency
Temporal nouns	@temporal-noun @time-name	Day, night, holiday October, December	141
Temporal adverbs	@temporal-descriptive	First, again, late	32
Evaluative adjectives	@absolute-adjective @superlative-adjective	Huge, great, lovely Best, greatest, worst	184
Evaluative-descriptive adjectives	@comparable-adjective	Romantic, crowded, expensive	342
Comparative adjectives	@comparative-adjective	Better	6
Mental verbs	@mental-verb	Love, enjoy, wonder	104
First person pronouns or adjectives	@1p-pronoun	I, me, my, we, us	118
Second person pronouns or adjectives	@2p-pronoun	You, your	22
Verbs oriented to the addressee	@addressee-oriented-verb	Warn, advice, tell	1
Verbs in the infinitive form	@infinitive	Every verb in the infinitive form	530

5 Conclusion

This paper presents a first exploration of OTR titles and the results reveal interesting specificities. There is an overall concordance between titles with a positive or a negative connotation and positive or negative review texts. This means that titles are

quite representative of the review orientation and accomplish the general function of text titles in helping readers anticipate what follows in the text. The measurement of lexical density, that is an indicator of text informativeness and readability, showed that nearly 3 out of 5 words in OTR titles communicate univocal information. Since the longer a text is the more representative the value of lexical density becomes, this result should be better interpreted as an indicator of the readability than of the informativeness: OTR titles give basic information about the review and are easy to read. Further, they make strong use of superlatives and slogans, and positive words are much more frequent than negative ones, which corresponds to the overall dominance of positive reviews in TripAdvisor. The frequency counts for keywords and key-phrases are very low, suggesting a great diversity of words being used. In addition, travel reviewers use less words for introducing a positive experience at a destination than for depicting a negative or average impression.

An explanation for this finding may be that negative judgments need to be convincingly supported, since negative critiques are usually less accepted then positive ones. This issue deserves further investigation in future research, specifically considering, for instance, negative arguments about a travel experience both from a logical and from a dialectical point of view. As far as the communicative functions of OTR titles are concerned, the results indicate that they mostly summarize and preview the personal experience descriptions provided in the reviews. Titles often refer to the destination (e.g. “Rome”) or the kind of experience being reviewed (e.g. “My trip to Paris”), as suggested by the frequent use of temporal indications and meta-words. Such titles do not add any relevant information for the reader, and platform managers should clearly indicate that they should be avoided. A case to be further investigated is that of titles constituted by concessive sentences, where the value of an experience or a destination is subordinated to some conditions. They are a kind of warning or limit the applicability of the statements in the review, and as such represent important information for a prospective tourist. Concessive sentences are usually introduced by subordinate conjunctions such as “but” or “if”, so that an automated search might be undergone to identify them. A comparison between the strategies adopted by tourists to write the title of their reviews and the strategies used by Destination Management Organization to advertise the destination should be included in future research. Overall, the exploratory research shows that titles should be included in the analysis of OTRs and could indeed be used for classifying reviews.

References

Aaker, J. L. & Maheswaran, D. (1997), The effect of cultural orientation on persuasion. *Journal of Consumer Research* 24: 315-28.

Balatsoukas, P., Morris, A. & O’Brien, A. (2009). An evaluation framework of user interaction with metadata surrogates. *Journal of Information Science* 35(3): 321-339.

Bell, J. (1991). *The Language of News Media*. Oxford: Blackwell.

Bernstein, M. & Garst, R.E. (1982). *Headlines are Deadlines*. New York: Columbia University Press.

Burgoon, M., Dillard, J.P., Doran, N.E. & Miller, M.D. (1982). Cultural and situational influences on the process of persuasive strategy selection. *International Journal of Intercultural Relations* 6: 85-100.

Colombetti, M. (2001). A Language for Artificial Agents. *Studies in Communication Sciences* 1: 1-32.

Dor, D. (2003). On newspaper headlines as relevance optimizers. *Journal of Pragmatics* 35: 695-721.

Eggins, S. (1994). *An Introduction to Systemic Functional Linguistics*. London: Continuum International Publishing Group.

Ernst, S. B. (1980). A feature analysis of Clio winning ads. *Journalism Quarterly* 57: 321-324.

Freeman, K. (2005). Creating strategic taglines. *Strategic Direction* 21(10): 3-4.

Glenn, E. S., Witmeyer, D. & Stevenson, K. A. (1977). Cultural styles of persuasion. *International Journal of Intercultural Relations* 3: 52-65.

Grice, P. (1989). *Studies in the Way of Words*. Cambridge MA: Harvard University Press.

Halliday, M. (1994). *An Introduction to Functional Grammar*. London: Edward Arnold.

Kim, H. & Fesenmaier. D. R. (2007). Persuasive Design of Tourism Web sites in the United States. In *Proceedings of the Annual Conference of the Travel and Tourism Research Association*, Ljubljana, Slovenia, July.

Kim, H. & Fesenmaier. D. R. (2008). Persuasive Design of Destination Web Sites: an Analysis of First Impression. *Journal of Travel research* 47 (1): 3-13.

Laskey, H. A., Fox, R. J. & Crask, M. R. (1994). Investigating the impact of executional style on television commercial effectiveness. *Journal of Advertising Research* Nov/Dec: 9-16.

Marchionini, G. (1995). *Information seeking in Electronic Environments*. Cambridge, UK: Cambridge University Press.

O'Donnell, M. (2008). The UAM CorpusTool: Software for corpus annotation and exploration. In *Proceedings of the XXVI Congreso de AESLA*, Almeria, Spain, 3-5 April 2008.

Pan, B. & Fesenmaier, D. R. (2006). Travel Information Search and Navigation on the Internet. In D. R. Fesenmaier, H. Werthner and K. W. Wöber (Eds.), *Travel Destination Recommendation Systems: Behavioural Foundations and Applications*. Cambridge, MA: CAB International, 30-44.

Rigotti E. & Rocci A. (2001). Sens - non-sens - contresens. *Studies in Communication Sciences* 2: 45-80.

Rigotti, E. & Cigada, S. (2004). *La Comunicazione Verbale*. Milano: Apogeo.

Rocci, A. (2005). Connective Predicates in Monologic and Dialogic Interaction. *Studies in Communication Sciences*, *special issue on Argumentation in Dialogic Interaction*: 97-118.

Schmidt, R., Platt, J. & Platt, H. (1995). *Longman Dictionary of Language Teaching and Applied Linguistics*. 4th edition.

Stewart, D. & Furse, D. (1986). *Effective television advertising: a study of 1 000 commercials*. Mass: Lexington.

van Dijk, T. A. (1988). *News as Discourse*. Lawrence Erlbaum Associates, Publishers.

Wang, Y. & Fesenmaier, D. R. (2003). Understanding the motivations to contribute to online communities: an empirical study of an online travel community. *Electronic Markets* 13(1): 33-45.

Wöber, K. (2006). Domain Specific Search Engines. In D. R. Fesenmaier, H. Werthner & K. Wöber (eds.), *Travel Destination Recommendation Systems: Behavioural Foundations and Applications*. Cambridge, MA: CAB International, 205-226.

Xiang, Z. & Fesenmaier, D. R. (2006). Assessing the Initial Step in the Persuasion Process: Meta Tags on Destination Marketing Web sites. *Information Technology & Tourism* 8(2): 91-104.

Yoo, K. H. & Gretzel, U. (2008). What motivates consumers to write online travel reviews? *Information Technology & Tourism* 10(4): 283-295.

Yssel, J. C. & Gustafson, R. L. (1998). A strategic analysis of the best television commercials ever produced (worldwide). *Communicatio* 24(2): 2-10.

An Empirical Study on the Relationship between Twitter Sentiment and Influence in the Tourism Domain

Donato Barbagallo,
Leonardo Bruni,
Chiara Francalanci, and
Paolo Giacomazzi

[a] Department of Electronics and Information
Politecnico di Milano, Italy
{barbagallo, bruni, francala, giacomaz}@elet.polimi.it

Abstract

Social media have a strong impact on the way users interact and share information. Several previous studies have highlighted how the structure of a social network can affect the dynamics of user interaction and information sharing. The majority of these studies have focused on the role of influencers, i.e. nodes with a central position in the network. Our claim is that while the information shared by influencers has a broader reach, the content of messages plays a critical role and can be a determinant of the social influence of the message irrespective of the centrality of the message's author. In this paper, we put forward four hypotheses supporting this claim by focusing on the sentiment of posts to characterize content and test them on a data set of 500,000 messages from Twitter in the tourism domain. Overall, our hypotheses posit that negative posts are more influential than positive ones. Results show how negative tweets are retweeted more than positive tweets. However, the time dynamics of retweeting seem independent of the sentiment of tweets.

Keywords: social media; influence; influencer; micro-blogging.

1 Introduction

Social media have a strong impact on the way users interact and share information. To handle the huge quantity of information from social media, there is a growing need for semantic technologies for social media intelligence. Companies are now aware of the benefits from the integration of social media intelligence with their business intelligence processes and related tools. This integration represents the first step of innovative marketing research, brand reputation management, and customer care.

In this context, the process through which users create and share opinions on brands, products, and services, i.e. the electronic word-of-mouth (*eWOM*) is gaining increasing attention. In the online context, the eWOM has been transformed from a communication act that takes place in a private one-to-one context to a *one-to-many* complex interaction. This represents the most powerful aspect of the eWOM. The reach of information sharing through eWOM can be both broad and fast. Companies know that controlling the dynamics of information sharing is very difficult. This need for improving control is one of the reasons why there is a growing interest in understanding how the structure of a social network can affect the dynamics of user interaction and information sharing.

Several previous studies have focused on the role of influencers, i.e. nodes with a central position in the network. In particular, micro-blogging platforms such as

Twitter are the focus of a wide range of studies that aim at understanding how messages spread inside the social network and how the role of the message author impacts on message reach. Micro-blogging networks are more and more used by companies as a communication medium for the promotion and engagement of customers. An emerging paradigm for the study of social networks as a communication medium is the *attention economy* (Beck et al., 2001). This paradigm starts from the observation that brands are involved in a competition for gaining the attention of possible customers. While on traditional media attention can focus not only on content, but also on the way a message is conveyed, on social media content plays a more central role. Content is even more central with micro-blogging, as the shortness of messages compels users to focus on the core of the information that they want to share. On Twitter, the standard size of a message limited to 140 characters is roughly the typical size of headlines and encourages users to produce contents that are easy to consume.

Our claim is that, while the information shared by influencers has a broader reach, the content of messages plays a critical role and can be a determinant of the social influence of the message irrespective of the centrality of the message's author. We make a distinction between *influence* and influencers. While nodes are influencers depending on their centrality in the social network, influence is the actual impact of messages, which depends not only on the structure of the network, but also on the ability of message content to raise attention. Studying how contents spread within social networks can be useful for explaining why some trends are followed more quickly and successfully than others, thus providing an invaluable input to business intelligence. The concept of influence can provide more accurate insights on how companies can leverage social media to strengthen their brand's reputation.

In this paper, we put forward three hypotheses supporting the claim that the content of messages plays a critical role and can be a determinant of the social influence of the message irrespective of the centrality of the message's author. We characterize the sentiment of posts as an important characteristic of message content. Overall, our hypotheses posit that negative posts are more influential than positive ones. Hypotheses are tested on a data set of 500,000 tweets in the tourism domain.

The remainder of this paper is structured as follows. Section 2 introduces the base concepts of graph theory that are needed to characterize influencers and influence. Section 3 presents related research and highlights the innovative aspects of this work. Section 4 presents our research hypotheses, while Section 5 reports testing results. Finally, Section 6 draws some conclusions and presents future research directions.

2 Base Concepts of Graph Theory

This section provides a discussion about basilar concepts of the graph theory useful to the contextualization of the work presented in this paper. The reader who is already familiar with the graph theory may skip to the next section.

The base idea of a social network is very simple. A network is a set of actors (also called elements, or nodes, or agents) that are related with each other in some way. Given a finite set U of elements:

$$U = \{X_1, X_2, \ldots, X_n\} \tag{1}$$

and a finite number of relations *R*:

$$R_t \subseteq UxU \tag{2}$$

where

$$t = 1,2,\ldots,r \tag{3}$$

a social network *N* is defined as the tuple composed by the finite set of elements *U* and by the *r* relations among them:

$$N = (U, R_1, R_2, \ldots, R_r). \tag{4}$$

Relations can represent any type of relationship, e.g., the father-son relationship, the cooperation of two employees within the same project, the friendship relation between persons.

In literature, the most frequently used types of social networks are the following:

- *undirected network*: all the relations R_t are symmetric, as in the case of the "married-to" relation;
- *directed network*: all the relations R_t are not symmetric, as in the case of the "father-of" relation;
- *mixed network*: in the same network coexist both directed arcs and non-directed arcs, as in the case of two different relations represented in a same graph such as "married-to" and "father-of".

The literature proposes several metrics to describe the position of a node inside a network (Brodka et al., 2009). The position of a node is key to understanding its role within the network. In particular, the *prestige position* and *centrality position* metrics measure the importance of the node. If the relations in the network represent communication channels between nodes, these metrics measure the potential impact of the node on the distribution of information across the network.

The *indegree centrality* and the *proximity prestige* metrics are widely used to measure the prestige position of a node. The former (IDC) is a "local" metric, as it considers adjacent nodes only by taking into account the number of direct inbound connections, while the latter (PP) is a global metric that defines the distance between the node, say a_i, and all other nodes. The following expression formally defines IDC:

$$IDC(a_i) = \frac{fanin(a_i)}{size(N) - 1} \tag{5}$$

where $fanin(a_i)$ is the number of nodes that have a connection towards node a_i and $size(N)$ is the total number of nodes in the network.

Various *centrality* metrics have been proposed in the literature to measure the centrality position of a node. The most common centrality metrics are the following:

- *Outdegree centrality*, measuring the breadth of reach of a node when forwarding messages through its outgoing arcs:

$$ODC(a_i) = \frac{fanout(a_i)}{size(N) - 1} \tag{6}$$

where *fanout(a_i)* is the number of outbound connections from node a_i and *size(N)* is the total number of nodes in the network.

- *Closeness centrality*, measuring the proximity of a node to the other nodes of the network:

$$CC(a_i) = \frac{size(N) - 1}{\sum_{j=1}^{N} d_{ij}}, \tag{7}$$

where d_{ij} is the distance between nodes a_i and a_j along the shortest path. As the distance between nodes in disconnected components of a network is infinite, this measure cannot be applied to networks with disconnected components.

- *Betweenness centrality*, measuring the number of times a node lies on the shortest path between two other nodes in the network:

$$BC(a_i) = \sum_{j<k} \frac{g_{jk}(a_i)}{g_{jk}}, \tag{8}$$

where $g_{jk}(a_i)$ is the number of times node a_i lies in a minimum path between nodes a_j and a_k, and g_{jk} is the total number of distinct paths between nodes a_j and a_k. Also this measure cannot be applied to networks with disconnected components.

3 Related Work

The study of social networks began in the 1930s with Simmel who has built the first theory that interprets social phenomena. In 1934, Moreno was the first to propose a formal representation of social networks as a combination of nodes and arcs. Then, Harary (Cartwright et al., 1965) and Cartwright (Cartwright, 1959) applied the concept of the graph theory to social networks that were called *sociograms*. With the introduction of directed arcs between nodes they were able to explain complex social patterns.

At the end of the 1930s, two different schools of thought emerged. The *sociocentric approach* (Chung et al., 2005) focused on identifying subgroups of people within the same network and understanding the relationships between subgroups. The *egocentric approach* was focused on the study of the whole community. This latter approach (Barres, 1954), (Granovetter, 1973), and (Milgram, 1967) emphasized the importance of social networks as a means to share knowledge and information. In paticular, Milgram introduced the concept of the *six degrees of separation* (Milgram, 1967), which attempted to demonstrate the idea of what he called "*small world phenomenom*," particularly interesting in understanding the power of eWOM.

Freeman focused on the definition of important nodes in a network and related metrics (Freeman, 1979). In this respect, micro-blogging has created new opportunities. One of the most interesting studies on micro-blogging is *The Million Follower Fallacy*, by (Benevenuto et al., 2010). With a very large Twitter data set consisting of about six millions users and considering the indegree (i.e., the number of followers) metric to measure the importance of users, this work analyses the correlation between indegree and mentions and retweets. The conclusion of this work is that user's popularity has a little effect on the actual attention from other users measured by retweets and mentions.

A recent study in the context of the Ecology Web Project (Leavitt, 2009) focuses on the influence of a set of 12 very popular users, based on an indepth analysis of their posts and corresponding responses. Users were divided into three clusters, celebrities, news, and social media analysts. The authors of this study found that celebrities have the largest number of followers and are able to produce significant volumes of responses with a very low effort (i.e., activity). Social media analysts can reach the highest values of influence if their responses are weighed by the number of followers; however, these high values are reached only with a very high effort. Finally, news have the greatest ability to have their contents forwarded by other users.

Other researches (Kwak et al., 2010; Asur et al., 2010) found that the propagation of a message on Twitter is greater if a *twitterer* is more influential, measuring influence by means of the PageRank algorithm (Brin et al., 1998). However, their approach has not been further verified and cannot be considered general. In particular, further work is needed to test the validity of the PageRank metric within the different clusters of users highlighted in (Leavitt, 2009).

4 Research Hypotheses

Our claim is that in social media, including Twitter, content plays a key role in determining the influence of information. In this section, we put forward four hypotheses supporting this claim. In our hypotheses, we focus on the sentiment of messages as an important characteristic of content of tweets. The literature provides evidence that social media users tend to be self-promoting by generating a higher number of messages with a positive sentiment (Sobel et al., 2009; Berger & Milkman, 2010). On the other hand, traditional media usually emphasize negative news (Galtung & Ruge, 1965). Our first hypothesis aims at verifying whether the sentiment of posts has the same impact on their influence as in traditional media:

H1: In Twitter, negative posts are more retweeted than positive posts.

The literature has studied the role of social media, especially Twitter, as a source of news (Kwak et al., 2010; Boyd et al., 2010). In particular, the literature has discussed the ability of social networks to quickly spread information and the relative volatility of information created and "consumed" by users. Kwak et al. (2010) show that most trending topics have an active period of one week, while half of retweets of a given tweet occurs within one hour and 75% within one day. In case of emergencies, Twitter has been seen as a reliable means of information (Starbird, 2010; Mendoza et al., 2010), with exceptionally high peaks of conversations on highly volatile issues. In

these occasions, 33% of tweets carry opinions and emotions (Qu et al., 2011), which is a significantly high percentage with respect to Twitter's usual 5% (Authors of this paper 2011). Building on previous results, our second hypothesis posits is that volatility is a general characteristic of posts, however, tweets carrying a positive sentiment are more volatile than tweets carrying a negative sentiment:

H2: In Twitter, on average negative tweets are retweeted more quickly than positive tweets.

Unlike Facebook, the Twitter mechanism of friendship is directed. Users select other users to follow and receive push notifications from them on their homepages, called timelines. This mechanism has been shown to favour more connected users (Zhou et al., 2010), since messages from non-followees must be searched manually, i.e. with a pull mechanism. In our third hypothesis, we posit that this role of followers is strong, since a shift from a push to a pull mechanism is known as an impactful mechanism in marketing literature. This leads to the following hypothesis:

H3: In Twitter, negative tweets are more retweeted by followers than by non-followers.

Analyses by (Bakshy et al., 2011; Benevenuto et al., 2010; Leavitt, 2009) have shown the relations among the number of followers and the propagation of messages. From their studies it is clear that it is important to have a (even small) set of focused followers that can be considered influencers, rather than many "normal" users. According to the literature, influencers can act as hubs and spontaneously generate snowball effects in the propagation of news, creation, and acceleration of innovations (Katz & Lazarsfeld, 1955; Rogers, 1995). The literature has not studied the relation between the propagation time of a message and the network size of the author. Anyway, Twitter has demonstrated to have different characteristics in influence dynamics as shown by (Benevenuto et al., 2010; Bakshy et al., 2011) and even people with not large networks are able to exploit their influence effectively. Thus we address this issue in our next hypothesis:

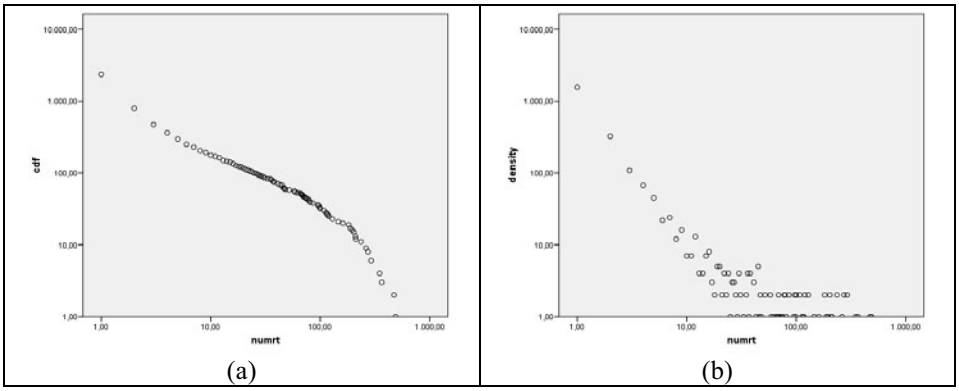
H4: In Twitter, the time (minimum, maximum, or average) necessary for a negative tweet to be retweeted the first time is not related to the number of followers of the author.

5 Empirical Testing

We have collected data on tweets and retweets using Twitter APIs. We have focused on the tourism domain, by collecting all the tweets containing the keyword "Milan" and used a semantic analysis tool to select only the tweets related to the city of Milan (Barbagallo et al., 2011) and evaluate their sentiment. More specifically, we have been able to classify tweets according to the Anholt model (Anholt, 2009) and created several categories of interest for the tourism domain, e.g., arts and cultures, services and transportations, food and drinks, hotels and accommodations, etc.

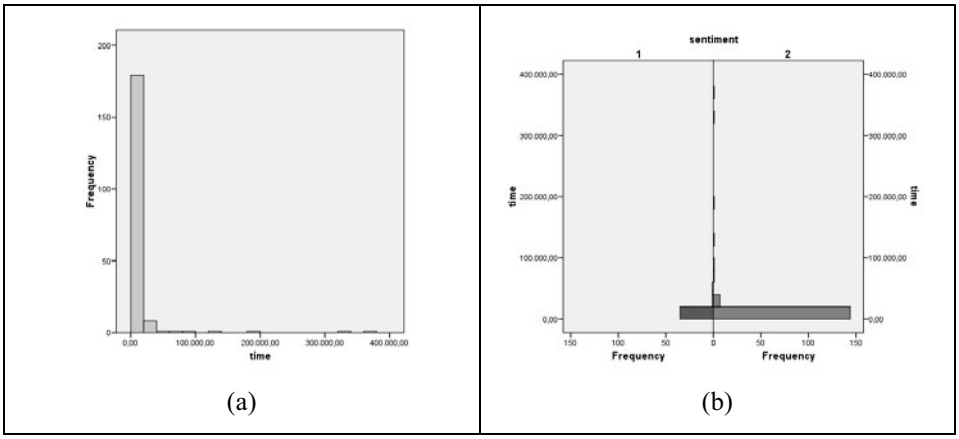
We have created two different datasets:

- Dataset A, used to test H1, contains all the tweets about the city of Milan, written between January 2011 and April 2011, containing 676,000 tweets (182,000 retweets).
- Dataset B, used for H2, H3, and H4, contains all the tweets about city of Milan, written between February 1st 2011 and February 28th 2011. This dataset contains 158,000 tweets (38,000 retweets).



(a) (b)

Fig. 1. Cumulative distribution function (a) and density (b) of number of retweets per each tweet with sentiment

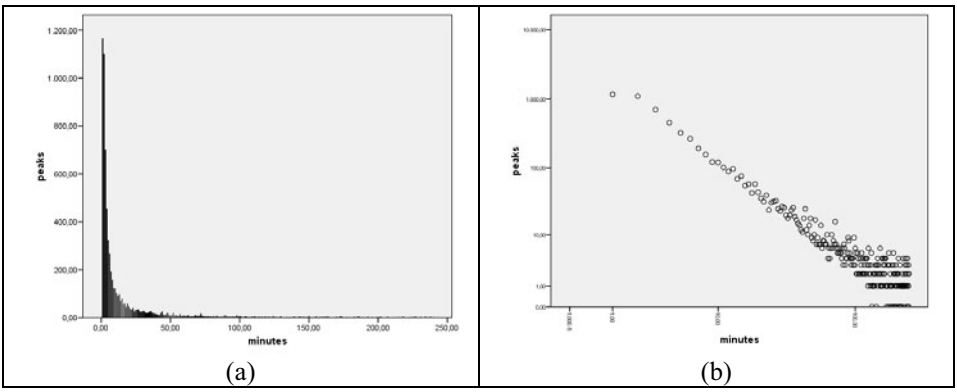


(a) (b)

Fig. 2. Time (expressed in seconds) distribution of the number of retweets with sentiment (a) and different retweeting time over sentiment (b)

In order to test hypothesis H1, we have used Dataset A and analyzed it using our sentiment analysis tool (authors of this paper 2011). The tool found 2355 tweets with sentiment, among which 387 have been classified as negative and 1968 as positive. We have then considered the number of retweets that each of these tweets has received. We have tested the normality of the distribution of retweets through a Kolmogorov-Smirnov test (sig. < 0.001). Fig. 1 confirms the results of the test showing the cumulative distribution function (a) and the density (b) of the number of

retweets of tweets with sentiment. Since the sample is not normally distributed, we have tested the difference of means using Mann-Whitney test (Z = -1.817, sig. = 0.069) showing a weak support for hypothesis H1.



(a) (b)

Fig. 3. Density of peaks of retweets in the first 4 hours with linear scale (a) and log-log scale (b)

In order to test hypothesis H2, we have analyzed the dynamics of retweeting. In particular, we have considered peaks and time distribution. More specifically, Fig. 2a shows that 92% of tweets carrying sentiment are retweeted within 5 hours, while Fig. 2b shows the difference of retweeting time of tweets with positive and negative sentiment. It is possible to observe that tweets with negative sentiment (column 1) are more concentrated in the first minutes, while positive tweets (column 2) have higher volumes but are retweeted later. Then, we define the time t of a peak of retweets as follows:

$$t = \arg\max_t (NRT(t)) \quad 0 \le t \le T, \tag{9}$$

where *NRT(t)* is the number of retweets of a given tweet at time *t* and T is the time of the last retweet. Time *t* is relative to the time instant in which the tweet is written. Fig. 3a reports the distribution of peaks over the first 4 hours from the time instant when each tweet is created and shows the cumulated number of tweets that have a peak of retweets in each time instant. It is clear that the large majority of peaks occur in the first minute. In our sample, 1168 tweets (17%, as can be seen fromTable 3) have their peak within the first minute. Fig. 3b reports the same data in a log-log scale. The graph that is obtained is very similar to a power-law distribution, since the decrease of the number of peaks over time is heavy-tailed. While Figures 2 and 3 are useful to obtain a qualitative understanding of retweeting dynamics, the Mann-Whitney test does not support H2 (Z = -0.11, sig. = 0.912).

A more indepth analysis can be conducted using data reported in Table 3: 80% of peaks happen within 31 minutes and almost one third happens after two minutes.

In order to test hypothesis H3, we have considered dataset B that contains 18206 retweets that have been posted by followers of the author and 13061 that have been posted by users who were not following the author of the message when the message

was posted. A Kolmogorov-Smirnov (sig. < 0.001) test has been run in order to verify whether the distribution of retweets in dataset B is normal. Since data are not normally distributed, we have run a Mann-Whitney test for the mean difference between the two groups (Z = -2.79, sig. = 0.005) supporting hypothesis H3.

Table 4. Cumulated percentage of number of peaks of retweets over time

	cumulated percentage
Percentage of retweets within 1 minute	17%
Percentage of retweets within 2 minutes	32%
Percentage of retweets within 5 minutes	53%
Percentage of retweets within 31 minutes	80%
Percentage of retweets before the average	90%

Hypothesis H4 has been tested through correlations. Since data are not normally distributed, we have used the Spearman correlation coefficient. Neither of the three correlations support hypothesis H4, neither for minimum time (ρ = -0.203, sig. < 0.001), nor for average time (ρ = 0.101, sig. < 0.001) and maximum time (ρ = 0.046, sig. < 0.001). Details are reported in Table 4.

Table 5. Spearman correlation coefficients and related significance values between the indegree centrality of each author and the minimum, average, and maximum time of retweet

	minimum time interval	**average time interval**	**maximum time interval**
indegree centrality	-0.203, sig. < 0.001	0.101, sig. < 0.001	0.046, sig. < 0.001

6 Conclusions and Future Work

While several studies underline the differences between traditional and social media, our results highlight interesting common features. Even if the absolute number of tweets with positive sentiment is greater than the number of negative ones, on average negative tweets seem to be more retweeted and attract more attention. However, the dynamics of retweeting seem to be independent of the sentiment carried by tweets (H2). Qualitative analyses have shown how tweets are volatile, as 80% of retweeting occurs within an hour from posting. This aspect is particularly critical for companies, as it shows that social media monitoring must be an almost real-time process. These results are based on a sample of tourism-related tweets on Milan city. It is interesting to note how (Chen et al., 2010) found that 4.5% of tweets discuss directly of tourism and travel tips, not considering those topics that can indirectly impact on tourism, such as arts and culture (31.8%), local news (54.5%), and sports (18.2%). Our results suggest how the influence of news in anyone of these domains could have an indirect impact on tourism that is significantly broader if news are negative.

The role of the number of followers is still unclear. Previous studies indicate that the relation between the centrality of a user and the ability of stimulating discussion is limited. This paper shows that centrality has no direct relation with the time of retweet (H4). On the other hand it is clear that having more followers increases the probability of being retweeted, especially for negative opinions, as posited in H3.

In the last years, cities are trying to become smarter, using technology to enhance their citizen life and to attract tourists by advertising their services. Twitter seems to be a very powerful tool to build a city brand reputation. (Scharl et al., 2008) have shown that social media coverage and sentiment influence a tourism destination image. Our results confirm these insights, showing that while the content of a post in terms of sentiment has an impact on the reach of the post, the time dynamics of user interactions seem to be a characteristic of the specific social medium that is independent of content.

We are currently working on the detection of trending topics in order to make a more precise measurement of the diffusion of news related to events. One of the limitations of this work is to consider only retweets containing the keyword "Milan", while studying the diffusion of events based on the Twitter hashtag mechanism would add to the generalizability of our results.

References

Anholt, S. (2009). *Competitive Identity: The New Brand Management for Nations, Cities and Regions.* Palgrave Macmillan, Eds.

Aus, S., Galuba, W., Huberman, B.A. & Romero, D.M. (2010). *Influence and passivity in social media*. ACM.

Bakshy, E., Hofman, J. M., Mason, W. A. & Watts, D. J. (2011). Everyone's an influencer: quantifying influence on twitter. In *Proceedings of the fourth ACM international conference on Web search and data mining (WSDM '11)*. New York, NY, USA: ACM, pp. 65-74.

Barbagallo, D., Cappiello, C., Francalanci, C. & Matera, M. (2011). Semantic sentiment analyses based on the reputation of Web information sources. In V. Sugumaran and J. A. Gulla (Eds), *Applied Semantic Web Technologies*. Taylor & Francis.

Barres, J. (1954). Class and committees in a Norwegian island parish. *Human relations* 7: 29-58.

Beck, T. & Davenport, J. (2001). *The attention economy: Understanding the new currency of business*. Cambridge: Harvard Business School press..

Benevenuto, F., Cha, M., Gummadi K.P. & Haddadi, H. (2010). Measuring user influence in Twitter: The million follower fallacy. In *Proceedings of the 4th International AAAI Conference on Weblogs and Social Media*. Association for the advancement of artificial intelligence.

Berger, J. & Milkman, K. (2010) *Social transmission, emotion, and the virality of online content*. Wharton Research Paper.

Boyd, D., Golder, S. & Lotan, G. (2010). Tweet Tweet Retweet: Conversational Aspects of Retweeting on Twitter. In *Proceedings of HICSS-42*, Persistent Conversation Track. Kauai, HI: IEEE Computer Society. January 5-8, pp. 1-10.

Brin, S. & Page, L. (1998). The anatomy of a large-scale hypertextual Web search engine. In *Proceedings of the 7th International World Wide Web Conference*. 30 (1-7), pp. 107-117.

Brodka, P., Kazienko, P. & Musial, K. (2009). User position measure in social networks. In *Proceedings of the 3rd Workshop On Social Network Mining and Analysis*. Paris, France: ACM.

Cartwright, D. (1959). *Studies in social power*. Vol. 6. University of Michigan, USA: Research Center for Group Dynamics.

Cartwright, D., Harary, F. & Norman, R. (1965). *Structural models: An introduction to the theory of directed graphs*. New York, NY, USA: John Wiley & Sons.

Chen, J., Nairn, R., Nelson, L., Bernstein, M. & Chi, E. (2010). Short and tweet: experiments on recommending content from information streams. In *Proceedings of the 28th international conference on Human factors in computing systems (CHI '10)*. New York, NY, USA: ACM, pp. 1185-1194.

Chung, K.K., Davies, J. & Hossain, L. (2005). Exploring sociocentric and egocentric approaches for social network analysis. In *Proceedings of International Conference on Knowledge Management*. Pp. 1-8, Wellington, New Zealand, Asia Pacific.

Freeman, L.C. (1979). Centrality in social networks: Conceptual clarification. *Social Networks* 1(3): 215-239.

Galtung, J. & Ruge, M. (1965) The structure of foreign news: The presentation of the congo, cuba and cyprus crises in four norwegian newspapers. *Journal of Peace Research* 2(1): 64-90.

Granovetter, M.S. (1973). The strength of weak ties. *The American Journal of Sociology* 78(6): 1360-1380.

Katz E. & Lazarsfeld P. F. (1955). *Personal influence; the part played by people in the flow of mass communications*. Glencoe: Free Press.

Kwak, H., Lee, C.,Park, H. & Moon, S. (2010). Finding influentials based on the temporal order of information adoption in Twitter. In *Proceedings of the 19th international World Wide Web conference*. Raleigh, North Carolina, USA: ACM, pp. 1137-1138.

Kwak, H., Lee, C., Park, H. & Moon, S. (2010). What is Twitter, a social network or a news media?. In *Proceedings of the 19th international conference on World wide web* (WWW '10). New York, NY, USA: ACM, pp. 591-600.

Leavitt, A. (2009). *The influential: New approaches for analyzing influence on Twitter*. Boston, USA: Web Ecology Project.

Mendoza, M., Poblete, B. & Castillo, C. (2010). Twitter under crisis: can we trust what we RT?. In *Proceedings of the First Workshop on Social Media Analytics (SOMA '10)*. New York, NY, USA: ACM, pp. 71-79.

Milgram, S. (1967). The small world problem. *Psychology Today* 1(1): 60-67.

Qu, Y., Huang, C., Zhang, P. & Zhang., J. (2011). Microblogging after a major disaster in China: a case study of the 2010 Yushu earthquake. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work (CSCW '11)*. New York, NY, USA: ACM, pp. 25-34.

Rogers, E. M. (1995). *Diffusion of innovations*, 4th edition. New York: Free Press.

Scharl, A., Dickinger, A. & Weichselbraun, A. (2008). Analyzing News Media Coverage to Acquire and Structure Tourism Knowledge. *Information Technology & Tourism* 10(1): 3-17.

Sobel, K., Jansen, B. J., Zhang, M. & Chowdury, A. (2009) Twitter power: Tweets as electronic word of mouth. *Journal of the American Society for Information Science and Technology* 60 (11): 2169-2188.

Starbird, K. (2010). Pass It On?: Retweeting in Mass Emergency. In *Proceedings of the 7th International ISCRAM Conference* – Seattle, USA, May 2010.

Zhou, Z., Bandari, R., Kong, J., Qian, H. & Roychowdhury, V. (2010). Information resonance on Twitter: watching Iran. In *Proceedings of the First Workshop on Social Media Analytics* (SOMA '10). New York, NY, USA: ACM, pp. 123-131.