Using online consumer reviews as a source for demographic recommendations: A case study using online travel reviews

Article *in* Expert Systems with Applications · October 2013

DOI: 10.1016/j.eswa.2013.03.046

2 authors:

Nikos Korfiatis
University of East Anglia

Marios Poulos
Ionian University

Using online consumer reviews as a source for demographic recommendations: A case study using online travel reviews

Nikolaos Korfiatis
(korfiatis@em.uni-frankfurt.de)
Chair for Databases and Information Systems
Department of Informatics
Goethe University Frankfurt
Frankfurt, Germany

Marios Poulos
(mpoulos@ionio.gr)
Department of Archives and Library Science
Ionian University, Corfu
Greece

Forthcoming 2013

Expert Systems with Applications (ESWA)

Abstract

Online consumer reviews play an important role in the decision to purchase services online, mainly due to the rich information source they provide to consumers in terms of evaluating "experience"-type products and services that can be booked using the Internet, with online travel services being a significant example. However, different types of travelers assess each quality indicator differently, depending on the type of travel they engage in, and not necessarily their cultural or age background (e.g. solo travelers, young couples with children etc.). In this study, we present architecture for a demographic recommendation system, based on a user-defined hierarchy of service quality indicator importance, and classification of traveler types. We use an algebraic approach to ascertain preferences from a large dataset that we obtained from the popular travel website Booking.com using a web crawler and compared with the customer-constructed preference matrix. Interestingly, the architecture of the evaluated recommendation system takes into account already defined demand characteristics of the hotels (such as the number of reviews of specific consumer types compared to the total number of reviews) in order to provide an example architecture for a recommendation system based on user-defined preference criteria.

Keywords: Online reviews, Experience Goods, User preferences, Recommender Systems, Demographic Recommendations

1. Introduction

The travel industry is an excellent example where uncertainty about the quality of service can negatively influence consumer demand (Ye, Law, & Gu, 2009). Online reviews are generally third-party evaluations by consumers of the product or services advertised on a website, and are displayed next to the service description in order to reduce customer uncertainty and improve the perceived communication characteristics of the medium (Kumar & Benbasat, 2001). Their availability has been proven to have a positive effect on the volume of online sales (Archak, Ghose, & Ipeirotis, 2011; Chen, Xie, & Hall, 2008; Chevalier & Mayzlin, 2006). Websites with an increasingly large number of user-supplied reviews tend to experience an increased volume of sales, especially in the case of intangible products, as occurs with more traditional bricks-and-mortar sales channels (Wolfinbarger & Gilly, 2001). Online reviews exert even more influence in the case of products with a utility that can only be evaluated upon consumption (e.g. travel services), and provide a useful input for consumer choice (Vermeulen & Seegers, 2009).

This category of products and services is known in economics and marketing literature as experience goods (Nelson, 1970), and are generally classified in that category when their evaluated utility for the buyer can only be evaluated after consumption. In other words, the buyer has no idea about the quality of the product he/she is going to purchase and the only way to assess its quality is after consumption. Furthermore, the assessment of the experience might be influenced by time-dependent parameters which are not directly related with the service provided, but have to do with the accumulation of experience from the consumption of the good (Dulleck & Kerschbamer, 2006). That category is a subclass of experience goods which in the literature are termed “credence goods” (Dulleck, Kerschbamer, & Sutter, 2011). This can be particularly important in cases where online consumer behavior is sensitive on other users generated content reflecting past experiences, such as online reviews (Ye, Law, Gu, & Chen, 2010).

On the other hand, consumer adoption of online travel services has been boosted significantly by the sharing of past traveler experiences, mainly due to the fact that past experiences from other travelers reduce uncertainty (Gretzel & Yoo, 2008). Online travel websites provide an interface for consumers to make their booking and after the registered period of stay ends, an invitation is sent to the consumer by e-mail to evaluate his/her stay at that particular hotel and the quality of the amenities provided. This has led to a massive amount of online reviews on travel websites, that make the development of filtering mechanisms necessary to avoid first, the information overload to the intended buyer, and second, the unnecessary reviews that do not fit the characteristics of the consumer. This highlights the basic intuition behind this study, which is the use of user-generated content as a foundation for developing a recommender system architecture that relies on inputs from *service quality indicators* and *implicit traveler types* in order to suggest the travel service (in this case the accommodation) closest to his/her self-disclosed preferences.

In particular in this paper, we outline a basic design for a demographic recommender system that builds on two main inputs: implicit traveler type categorization and service quality indicator prioritization. We present a basic architecture for a system based on this mechanism, the calculation procedure and an experimental evaluation based on real data from the popular online travel website Booking.com. Based on this objective, this paper is structured as follows: Section 2 provides the theoretical background and the intuition behind this study. Section 3 describes the dataset that we used in order to illustrate our case and the analytical formulation of the filtering problem. The pre-processing and processing stage are presented with an experimental evaluation in Section 4. The paper concludes in Section 5 with discussions of the applicability of such a system, its limitations and possible extensions.

2. Background and Motivation

Plog (Plog, 1974) pioneered the idea that there is a difference between travelers' personality types and their expectations of the holiday value they receive by vacationing in a specific destination. Plog went even further, by defining two main types of travelers: *Dependables* and *Venturers*. In that classification, *Dependables* prefer well-established places where the amenities are already known in order to avoid risk. On the other hand, *Venturers* are travelers who avoid well-established venues and prefer unusual destinations for their travel. Based on Plog's intuition, we make a working hypothesis assuming that travelers' personality types play an important role in the selection of a venue and this input should therefore be considered, with a self-assessment by the traveler.

Another theoretical observation that can be made in this regard is that apart from personality, *Dependables* and *Venturers* can also be classified depending on the expected utility that they have in mind when they make their decision to travel to a certain place. The *Dependable* type links their satisfaction with their holiday experience closely to the accommodation and conditions of the trip, and expects them to be conventional 'as advertised', while *Venturers* readily accept inadequate or unconventional kinds of accommodation since they feel that this is an integral part of the uniqueness of the vacation. Hotels in traditional destinations are therefore more liable to be frequented and by extension reviewed by *Dependables,* for whom every deviation from the 'perfect' holiday would pull the rating toward the negative pole, even when this does not necessarily concern the hotel itself. *Venturers* would be less concerned about external factors and characteristics pertaining to the hotels of their choice, which would be in more exotic and unpopulated destinations, thereby, generating less digression in the season curves for these hotels.

The motivation behind this study is based on the simple observation that user-generated content in the travel sector has become particularly important and that it can provide useful data combined with the different evaluations of traveler types (Johnson, Sieber, Magnien, & Ariwi, 2011). This can enable the development of systems that can help travelers decide based on an a-priori self-assessment of their profile, as well as the prioritization of the various quality indicators the traveler deems important for his/her choice.

2.1 Service Quality indicators in online travel

Service quality in the travel sector is based on the general characterization of hotels and other travel service providers (e.g. airlines) within a star rating system, generally using a five star ordinal categorization system. The highest classification is the luxury class, which represents a superior performance in several different categories of services, such as location, cleanliness, staff, etc. For example in ref. (Hudson, Hudson, & Miller, 2004) Hudson et al., measure service quality in the tour operator sector using survey techniques based on the notion of perceived quality, and the difference between customer expectations and perceived quality in particular. Likewise, when broken down to individual quality indicators, reviews on online travel websites also reflect this difference, making this evaluation (when it exists) useful to consider.

Other factors may play a role, such as perceived risk, which might contribute to anxiety (Reisinger & Mavondo, 2005). Nusair and Kandampully (Nusair & Kandampully, 2008) performed a content analysis on the 6 most popular travel websites in order to evaluate the attributes that enhance customer satisfaction in the context of their websites. Consumer types greatly contribute to these attitudes, due to the distance between the expected and perceived base among different traveler types. For example in ref. (Atilla, 2006) reported that business travelers have the highest expectations when evaluating hotel quality due to their experience gained from frequent traveling.

2.2 Recommender systems in tourism and travel

Since the early days of the web, recommender systems have become an important research stream for online services, mainly as a result of providing a way to transfer preferences between users with similar profiles (Resnick & Varian, 1997). While several approaches have been used on providing recommendations to a user the most prominent in literature has been the case of collaborative recommender systems where collaborative filtering algorithms are used in order to extract common preferences among users with similar profile characteristics (Colomo-Palacios, García-Crespo, Soto-Acosta, Ruano-Mayoral, & Jiménez-López, 2010; Sarwar, Karypis, Konstan, & Reidl, 2001). Collaborative filtering uses other consumers with similar preferences and social characteristics (Adomavicius & Tuzhilin, 2005; García-Crespo, Colomo-Palacios, Gómez-Berbís, & Ruiz-Mezcua, 2010) in order to learn their taste and provide a recommendation that is closer to the users taste using other users' actions as an input.

While collaborative filtering is used to infer preferences from actions, in the domain of our study learning preferences from travelers actions might not always reveal their true intentions as well as the critical mass that is required in order to "learn" a collaborative filtering system. Therefore, the direction that we take on this paper is based on the concept of demographic recommender systems(Konstan & Riedl, 2012) and in particular the advantage it has over collaborative and content based filtering due to the fact that the matching process can be accomplished without the requirement of a history of user ratings(Burke, 2007) as well as the use of archetype "travelers" where the users can be compared with.

Recommender systems in online travel have found various applications in several stages of a consumer's purchase decision process. For example, a variant of recommendation systems referred to as destination recommendation systems (Fesenmaier, Wöber, & Werthner, 2006) are involved in the general selection process of travel destinations, by providing the consumer with a list of alternatives and advice based on the customer's different decision attitudes (e.g. accompanying persons, cost factors, etc.).

Apart from recommending destinations, destination recommender systems can also provide recommendations for different alternatives in the same destination. Bassala and Klenosky(Basala & Klenosky, 2001), for example, examined the effect of type of accommodation, type of travel companions and language, using conjoint profiles to assess the relationship between implicit and explicit quality indicators affecting consumer choice for the same destination. In ref. (Shapira, Taieb-Maimon, & Moskowitz, 2006) Shapira *et al.*, in a more generalized context, evaluated the relationship between explicit and implicit feedback, and found that relative quality indicators are more indicative than non-relative ones when the relation is based on the traveler type. Similar approaches with travel recommendation have been also implemented with more sophisticated systems, such as knowledge bases (Delgado & Davidson, 2002; García-Crespo, López-Cuadrado, Colomo-Palacios, González-Carrasco, & Ruiz-Mezcua, 2011). Most of these approaches fall within the category of demographic recommendations, where the assessment of the traveler's profile is based on his/her co-alignment with a specific group of travelers usually categorized by age group and/or destination preference (Ricci, 2002).

Another application area where recommender systems in tourism are used is the formation of ad-hoc travel groups, because consumer uncertainty is usually reduced by traveling in a group, and greater economies of scale are achieved for travel operators (Buhalis & Law, 2008).

Having described the basic theoretical intuitions behind this study, we proceed with the practical implementation and subsequent experimental evaluation of the recommendation mechanism in the following sections.

3. Data and Implementation

3.1 Dataset description and problem definition

The dataset was comprised of reviews publicly available on the popular Internet travel website Booking.com[1] which were collected using an automated web crawler. The website provides an intuitive interface for searching for hotels in a given city, listing all hotels in the

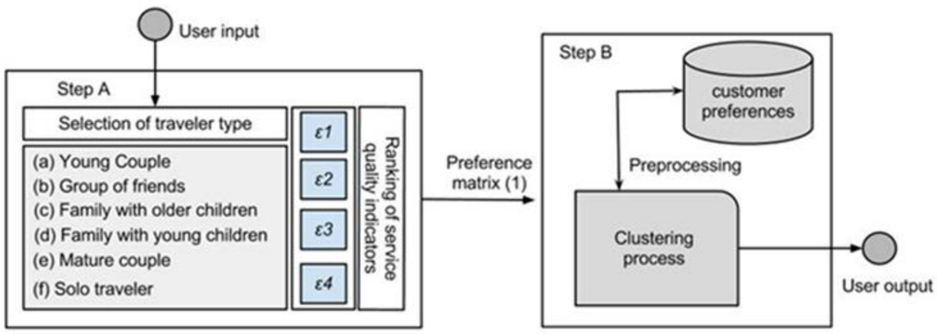

Figure 1: The basic architecture of the collaborative filtering system used in this study. User inputs are provided in Step A and the preference matrix is directed on the clustering process in step B, where the recommendation output is provided to the user.

city by page, without entering specific booking dates. The main reason for selecting Booking.com as an example application of the proposed system is the rating breakdown it provides for individual hotels, as well as the information about the users who rated this accommodation. We gathered data from the website using a web crawler for a period of one month, collecting a total of 185,700 reviews comprised of hotels of various star classifications. Table 1 shows the descriptive statistics for the four quality indicators provided in our dataset and the total score encompassing them. The hotel ratings in Booking.com are generally above average (1-10 rating scale), with a deviation of 1 unit by different types of indicator, with the smallest deviation in the "value for money" rating. We used these data in order to ascertain consumer preferences for a demographic filtering system that performs in two steps, as depicted in Figure 1. The first step (step A) requires the user to provide a self-assessment of his traveler profile based on accompaniment, and in the second step (step B) the user provides an ordinal classification of his/her preference for each service quality indicator involved. The traveler types and service quality indicators are outlined in Table 2, and are based on the classification provided by Booking.com which is the source of the customer preferences data used in step B.

For a user u_i belonging to a category of traveler type $\mu=\{a,b,c,d,e,f\}$ we have the preference matrix, constructed as follows:

$$u_i^{\lambda} = [\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4] \quad (1)$$

where $\varepsilon_1..\varepsilon_4$ corresponds to the different service quality indicators as depicted in Table 1. Taking a list of hotels with evaluations broken down by quality indicator and traveler type, we want to extract the optimal ranking that provides the best categorization available given the user-supplied data from its preference matrix.

Table 1: Number of Reviews and assigned weight indexes for quality indicators and different traveler classes

Number of reviews for different traveler classes	Assigned weight indexes of Hotel's Quality			
	Score Clean	Score Location	Score Services	Score Value For Money
Young Couple (*a*)	*k*1	*l*1	*m*1	*n*1
Group of friends (*b*)	*k*2	*l*2	*m*2	*n*2
Family with older children (*c*)	*k*3	*l*3	*m*3	*n*3
Family with young children (*d*)	*k*4	*l*4	*m*4	*n*4
Mature couple (*e*)	*k*5	*l*5	*m*5	*n*5
Solo traveler (*f*)	*k*6	*l*6	*m*6	*n*6
Rating vector (i=1..6 cases)	*dk*(i)	*dl*(i)	*dm*(i)	*dn*(i)

Table 2: Descriptive statistics for the service quality indicators used in constructing the user profiles

Indicator	Min.	Max.	Mean	St.dev
Value for money	3.2	10	7.45	0.81
Services	2.5	10	7.27	1.0
Location	3	10	8.48	0.9
Cleanliness	2.9	10	7.97	0.95
Total Score	3.6	10	7.75	0.76

In this case, we deal with the extraction of feature weight of the quality indexes of the hotel rating taking into account the number of reviews for a specific traveler type (Table 2).
The problem formulation above is based on the consideration that the weight indexes of each class present a strong degree of homogeneity. Before proceeding with the estimation of the weight indexes, we therefore conduct a preprocessing evaluation. As a first step in the preprocessing stage, we use an unsupervised system to create groups of set hotel's reviews which present correlated features. Once the preprocessing stage is completed, we calculate the feature weights (*k*, *l*, *m* and *n*) by solving the linear equation system dimensions $i*j$. We outline the preprocessing procedure in the section below.

3.2 Preprocessing Stage

Before continuing with the processing of the data for the recommendation provision by learning the individual user profiles, we first check whether the data are homogeneous. The hypothesis that the groups of the points (data review) satisfies the condition of homogeneity is resolved by a well-fitted unsupervised clustering *k*-means method, where the partitions of the points in the i-by-j data matrix A (see the processing stage) are grouped into k clusters.

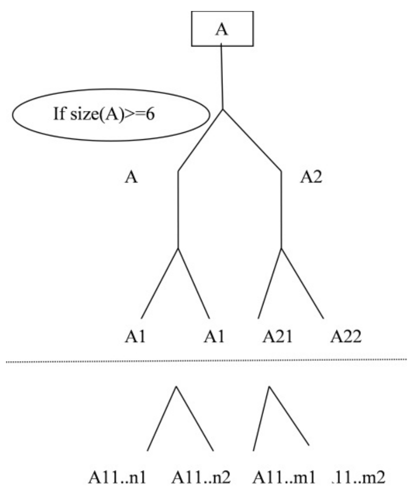

Figure 2: Decimation of the A data by k-means clustering, using hotel service quality indicators obtained from the dataset.

This iterative partitioning minimizes the sum, over all the clusters, of the within-cluster sums of point-to-cluster-centroid distances. The rows of A correspond to points (see equation 2), while the columns correspond to variables. The *k*-means process returns an i-by-1 vector containing the cluster indices of each point. By default, *k*-means uses squared Euclidean distances (Kanungo et al., 2002a, 2002b). As a result, in our case we used as value *k*=2 as the initial number of clusters, and we intuitively classified a given data set using a certain number of clusters (assume *k* clusters) established beforehand. Furthermore, the K-means algorithm is simple and fast. The time complexity of K-means is *O*(k*l*N), where *l* is the number of iterations, *k* is the number of clusters and *N* the total sample (Huang, 1998).

The main idea is to define one *k* centroids for each cluster. These centroids should be placed carefully because different locations provide different results. A loop is generated, which shows that the *k* centroids change their location step by step until no more changes are made. In other words, the centroids do not move any more. In this case, we submitted the dataset from the indicators (provided in table 1) into *k*=2 clusters using the following objective function:

$$J = \sum_{j=1}^{k} \sum_{i=1}^{N} \left\| A_i^{(j)} - c_j \right\|, \quad (2)$$

However, we have to divide all the hotels based on the homogeneity features of the groups. To that end, we adopted a technique involving the Danielson–Lanczos lemma (Press, Flannery, Teukolsky, & Vetterling, 1992) with data review, which is based on a butterfly classification in which the original population is split into 2 orders (which justifies *k*=2 of equation 2), along with the homogeneity features of the grouping of the data review. The butterfly classification is implemented by a butterfly diagram (provided in Figure 1) which depicts the iterative split *k*- means procedure until a convergence is archived, as explained in the previous section.

Having completed the pre-processing stage, we proceed with the processing stage, in which the weight indexes are extracted algebraically using matrix analysis.

3.3 Processing Stage

The problem in our case concerns combining two sets of values. On the one hand, we have the number of the reviews that were voted, and the other hand we have the aggregated results of grades per class (*k*, *l*, *m* and *n*) as shown in Figure 2. We therefore consider the following sets of equations:

$a_1 \times \kappa_1 + b_1 \times k_2 + c_1 \times k_3 + d_1 \times k_4 + e_1 \times k_5 + f_1 \times k_6 = dk$ (3)

$a_1 \times l_1 + b_1 \times l_2 + c_1 \times l_3 + d_1 \times l_4 + e_1 \times l_5 + f_1 \times l_6 = dl$ (4)

$a_1 \times m_1 + b_1 \times m_2 + c_1 \times m_3 + d_1 \times m_4 + e_1 \times m_5 + f_1 \times m_6 = dm$ (5)

$a_1 \times n_1 + b_1 \times n_2 + c_1 \times n_3 + d_1 \times n_4 + e_1 \times n_5 + f_1 \times n_6 = dn$ (6)

Each equation represents the fragmentation of section feature weights (k, l, m and n) correspondence and the coefficients a_1, b_1, d_1, e_1 and f_1 represent the percentage of the participant reviewers for each class. As a result, in order to calculate these we constructed a set of linear equations.

$$A \times k = dk \quad (7)$$
$$A \times l = dl \quad (8)$$
$$A \times m = dm \quad (9)$$
$$A \times n = dn \quad (10)$$

Where:

$$A = \begin{pmatrix} a_{11} & a_{12} & . & . & . & a_{1j} \\ a_{21} & a_{22} & . & . & . & a_{2j} \\ . & . & . & . & . & . \\ a_{i1} & a_{i2} & . & . & . & a_{ij} \end{pmatrix} \quad (11)$$

$$k = \begin{pmatrix} k_1 \\ k_2 \\ . \\ . \\ . \\ k_i \end{pmatrix} \text{ (12) , } \quad l = \begin{pmatrix} l_1 \\ l_2 \\ . \\ . \\ . \\ l_i \end{pmatrix} \text{ (13) , } \quad m = \begin{pmatrix} m_1 \\ m_2 \\ . \\ . \\ . \\ m_i \end{pmatrix} \text{ (14) , } \quad dk = \begin{pmatrix} dk_1 \\ dk_2 \\ . \\ . \\ . \\ dk_i \end{pmatrix} \text{ (15), } \quad dl = \begin{pmatrix} dl_1 \\ dl_2 \\ . \\ . \\ . \\ dl_i \end{pmatrix} \text{ (16) ,}$$

$$dm = \begin{pmatrix} dm_1 \\ dm_2 \\ . \\ . \\ . \\ dm_i \end{pmatrix} \text{ (17) and } \quad dn = \begin{pmatrix} dn_1 \\ dn_2 \\ . \\ . \\ . \\ dn_i \end{pmatrix} \text{(18)}$$

The values of k, l, m, and n are obtained from the solutions of the linear systems of equations (3), (4), (5) and (6) by the least square algorithm and with Gaussian elimination in cases where A is not a square matrix. We then extract the k, l, m and n as follows:

$$k = dk \backslash A \quad (19)$$
$$l = dl \backslash A \quad (20)$$
$$m = dm \backslash A \quad (21)$$
$$n = dn \backslash A \quad (22)$$

We subsequently create a new matrix w, which is responsible for each node in Figure 2 of k, l, m and n matrices as follows:

$$w_{A_1} = \begin{pmatrix} k_1 & l_1 & m_1 & n_1 \\ k_2 & l_2 & m_2 & n_2 \\ k_3 & l_3 & m_3 & n_3 \\ k_4 & l_4 & m_4 & n_4 \\ k_5 & l_5 & m_5 & n_5 \\ k_6 & l_6 & m_6 & n_6 \end{pmatrix} \quad (23)$$

We normalize these in the normalized W_{A1} table shown in equation (24):

$$w_{A_1} = \frac{\begin{pmatrix} k_1 & l_1 & m_1 & n_1 \\ k_2 & l_2 & m_2 & n_2 \\ k_3 & l_3 & m_3 & n_3 \\ k_4 & l_4 & m_4 & n_4 \\ k_5 & l_5 & m_5 & n_5 \\ k_6 & l_6 & m_6 & n_6 \end{pmatrix}}{\begin{pmatrix} \sum_{i=1} k_i + l_i + m_i + n_i \\ \sum_{i=2} k_i + l_i + m_i + n_i \\ \sum_{i=3} k_i + l_i + m_i + n_i \\ \sum_{i=4} k_i + l_i + m_i + n_i \\ \sum_{i=5} k_i + l_i + m_i + n_i \\ \sum_{i=6} k_i + l_i + m_i + n_i \end{pmatrix}} = \begin{pmatrix} \hat{C}_{x1} \\ \hat{C}_{x2} \\ \hat{C}_{x3} \\ \hat{C}_{x4} \\ \hat{C}_{x5} \\ \hat{C}_{x6} \end{pmatrix} \quad (24)$$

After the processing stage, we proceed with the evaluation of the process using an identification test in order to use it to evaluate the closest candidate profiles.

3.4 Identification Test

In this stage, the extracted sets of weighted coefficients $\widehat{C_x}$ from the first level of nodes in Figure 2 were submitted to the cross-correlation procedure along with another set $\widehat{C_y}$ using formula (23):

$$r = \frac{\sum_{i=1}^{k-1} (\hat{C}x_i - \bar{\hat{C}}x_i)(\hat{C}y_i - \bar{\hat{C}}y_i)}{\sqrt{\sum_{i=1}^{k-1} (\hat{C}x_i - \bar{\hat{C}}x_i)^2 \sum_{i=1}^{k-1} (\hat{C}y_i - \bar{\hat{C}}y_i)^2}} \quad (25)$$

The extracted cross-correlation coefficient is a number between -1 and 1, which measures the degree to which two variable sets are linearly related. In our study, we considered that an unknown data set has a perfect positive linear relationship with the auto-correlated set $\widehat{C_x}$ when the cross-correlation coefficient is approximately 1. Those records, which best satisfy this criterion, are then selected as the perfect weighted coefficient. The unknown data set $\widehat{C_y}$ is thus sorted in the sets of nodes which present the best approximation to the value of 1.

1	scoreClean	scoreLocation	scoreServices	scoreValForMoney	total numer of reviews	num reviews Young_couple	num reviews Group of friends	num reviews Family with young childr	numreviews Family with older childre	num reviews Mature couple	numreviews Solo traveller
2	7,3	9,3	6,9	6,7	7399	1432	1922	214	826	1422	1583
3	6,5	7,8	7,2	7,3	2094	994	526	25	70	166	313
4	8,8	7	8,3	8,1	2283	905	489	57	85	302	445
5	9,3	9,5	8,9	8	2361	630	289	72	120	852	398
6	9,3	8,9	8,9	8,4	2087	790	216	99	125	626	231
7	8,8	9,4	8	7,2	2719	736	523	85	107	642	626
8	7,9	9,7	7,4	7,2	2523	530	680	44	114	612	543
9	7,8	9,7	7,5	8,1	2210	147	829	232	609	233	160
10	6,3	5,8	5,8	6,5	2214	657	675	121	140	230	391
11	8,5	9,4	7,8	7,6	2467	686	501	56	112	516	596
12	8,6	8,3	8,3	8,7	2212	680	465	33	126	520	388
13	7,2	8,7	6,6	6,7	3497	693	565	157	196	731	1155
14	8,8	8,4	7,9	8,4	2807	708	624	76	231	559	609
15	7,9	8,2	7,1	6,9	3414	696	709	146	388	734	741
16	8,3	9,3	7,9	7,8	2093	492	210	82	118	627	564
17	7,4	8,3	7,8	7,5	2083	243	334	640	543	163	160
18	8,2	8,7	7,2	7,1	2680	508	578	108	321	517	648
19	7,8	8,8	7,1	6,4	1997	344	322	90	193	580	468
20	7,2	8	6,1	5,9	1985	529	302	75	140	349	590
21	7,8	8,1	7	7,9	1949	255	243	60	92	475	824
22	6,9	9,4	6,3	6,2	1842	542	301	128	208	250	413
23	7,5	8,7	6,3	6,5	2591	392	596	210	531	379	483
24	7,8	8,9	6,9	7,3	1919	359	463	159	237	502	199
25	9	9,5	9	8	2253	581	391	184	429	353	315
26	7,5	8,3	6,3	6,1	1853	331	266	56	177	221	802
27	9	8	8,9	8,7	1834	712	300	38	87	456	241
28	8,4	9,6	7,7	6,9	1830	271	304	281	520	253	201
29	7,8	7,2	7,7	8,4	1830	567	429	218	244	187	185
30	7,3	8,4	7	6,8	1792	122	419	444	495	127	185
31	6,6	8	5,6	6	1791	425	359	45	150	361	451
32	7,3	9,1	6,2	7,1	1791	413	435	138	291	188	326
33	7,5	7,3	6,8	7,6	1744	605	251	59	63	210	556
34	8,8	7,3	8,4	7,9	1732	787	238	26	40	210	431
35	8,3	9,8	7,7	7,8	1712	342	228	83	95	471	493
36	9	8	7,9	8,3	1702	647	215	41	36	182	581
37	6,7	9	6	6,6	1699	143	425	118	258	212	543
38	8	8,6	6,8	6,6	1697	539	249	56	94	301	458

Figure 3: Part of the area of 100 data set records, with the average score value of the reviewers highlighted in blue, while the number of reviewers participating is shown in black.

3.5 Statistical evaluation test

The significance of the relation degree *r* (see equation 25) between two variable weights coefficient sets $\widehat{C_x}$ and $\widehat{C_y}$ using the *t*-test control is calculated during this stage. We used the correlation of the coefficient *r*, which is an estimate of the population parameter of the data review clusters that was sampled via the weights coefficient sets. This parameter is denoted by ρ, lowercase Greek rho. We then considered that the null hypothesis is: $H_0{:}\rho_0=0$.The sampling distribution of r for a population that has zero correlation ($\rho=0$) therefore has a mean value of $\mu=0$. A t-statistic can hence be calculated as:

$$\sigma = \sqrt{\frac{(1-r^2)}{k-2}}$$

$$t = \frac{r-\mu}{\sigma} = \frac{r}{\sqrt{\frac{(1-r2)}{k-2}}} = \frac{r\sqrt{m-2}}{\sqrt{1-r^2}} \cdot \quad (26)$$

The next step was to determine the appropriate value of the *r* coefficient in order to characterize it as defining a significant linear relationship between the correlated sets in our experiment for both methods.

4. Experimental Evaluation

The experimental evaluation process involves three steps: data selection, normalization and clustering in order to find the closest type using a correlational approach. The main objective is to check whether the clustering process converges from the preprocessing stage, and to evaluate the coefficient of identification that the system provides when it comes to classifying new preference matrices submitted by users, using an identification test. We outline these steps in the following sections.

4.1 Step 1: Data selection

In the experimental stage, we selected a data set of 100 records (see Figure 3) in order to implement the aforementioned stages. We followed a general rule of thumb of 10% condition, suggested by Miligan and Cooper(Milligan & Cooper, 1986) in order to ensure enough variance for our evaluation.

4.2 Step 2: Data normalization

The determination of the nodes is done using k-means unsupervised clustering. In this stage we implement the step by step clustering procedure using the A table (see equation 11), which are the last five columns in Figure 3. This set of records is provided in the normalization procedure, and is shown in Figure 4.

0,193539668	0,259764833	0,028922827	0,111636708	0,192188134	0,21394783
0,474689589	0,251193887	0,011938873	0,033428844	0,079274117	0,1494746
0,396408235	0,214191853	0,024967148	0,037231713	0,132282085	0,19491896
0,266836086	0,12240576	0,030495553	0,050825921	0,360864041	0,16857263
0,378533781	0,103497844	0,047436512	0,059894586	0,299952084	0,11068519
0,270687753	0,192350129	0,031261493	0,039352703	0,236116219	0,23023170
0,21006738	0,269520412	0,017439556	0,045184304	0,242568371	0,21521997
0,066515837	0,375113122	0,104977376	0,275565611	0,105429864	0,0723981
0,296747967	0,304878049	0,054652213	0,063233966	0,103884372	0,17660343
0,278070531	0,203080665	0,022699635	0,04539927	0,209160924	0,24158897
0,307414105	0,210216998	0,014918626	0,056962025	0,235081374	0,17540687
0,19816986	0,161567057	0,044895625	0,056048041	0,209036317	0,330283
0,252226576	0,222301389	0,027075169	0,082294264	0,199144995	0,21695760
0,203866432	0,207674282	0,042765085	0,113649678	0,214997071	0,21704745
0,235069279	0,100334448	0,039178213	0,056378404	0,299569995	0,26946966
0,116658665	0,160345655	0,30724916	0,260681709	0,07825252	0,0768122
0,189552239	0,215671642	0,040298507	0,119776119	0,192910448	0,24179104
0,172258388	0,161241863	0,045067601	0,096644967	0,290435653	0,23435152
0,266498741	0,152141058	0,037783375	0,070528967	0,17581864	0,29722921
0,130836326	0,124679323	0,030785018	0,047203694	0,243714726	0,42278091
0,294245385	0,163409338	0,069489685	0,112920738	0,135722041	0,22421281
0,151292937	0,230027017	0,081049788	0,204940178	0,146275569	0,18641451
0,187076602	0,241271496	0,082855654	0,123501824	0,261594581	0,10369984
0,257878384	0,173546383	0,081668886	0,190412783	0,156679982	0,13981358
0,17862925	0,143550998	0,030221263	0,095520777	0,119266055	0,43281165
0,388222465	0,163576881	0,020719738	0,047437296	0,248636859	0,13140676
0,148087432	0,166120219	0,153551913	0,284153005	0,138251366	0,10983606
0,309836066	0,23442623	0,119125683	0,133333333	0,102185792	0,10109289
0,068080357	0,233816964	0,247767857	0,276227679	0,070870536	0,10323660
0,237297599	0,200446678	0,025125628	0,083752094	0,201563372	0,25181462
0,230597432	0,242881072	0,077051926	0,162479062	0,104969291	0,18202121
0,34690367	0,143922018	0,033830275	0,036123853	0,120412844	0,31880733
0,454387991	0,137413395	0,015011547	0,023094688	0,121247113	0,24884526
0,199766355	0,13317757	0,048481308	0,055490654	0,275116822	0,2879672
0,380141011	0,126321974	0,024089307	0,021151586	0,10693302	0,34136310
0,084167157	0,250147145	0,069452619	0,151854032	0,124779282	0,31959976
0,317619328	0,146729523	0,032999411	0,055391868	0,177371833	0,26988803

Figure 4: A snapshot of the normalization of the data (divided by the number of reviews) for each of the traveler classes outlined in Table 2

4.3 Step 3: Clustering procedure

In this stage, we implemented the iterative clustering procedure as indicated in section 3.2. As an example of this procedure, we present the first clustering (see Figure 5), which is achieved by 14 iterations (k-means clustering) in 5,343 samples with a calculated complexity *O* (14*2*5343)=149604

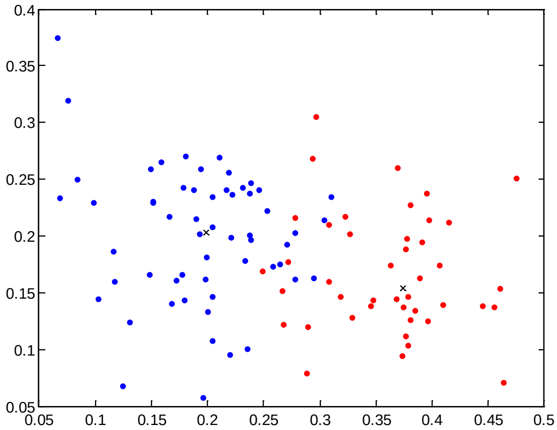

Figure 5: The first clustering. We extract 47 points in the blue group A1, and 5,343 points in the red group.

The procedure is iterated until the *k*-means clustering satisfies the convergence criterion, as mentioned in Section 3.2. The dendrogram resulting from this clustering process is extracted and presented in Figure 6.

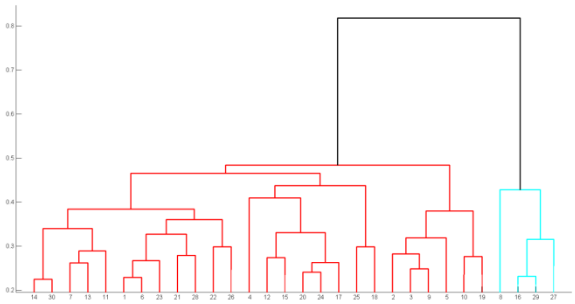

Figure 6: The resulting cluster dendrogram of 5,343 points after running the k-means procedure.

Using this procedure, we created groups with homogeneity features as discussed in Section 3.2. Having completed this stage, we proceeded with the algebraic estimation of the indicators in the following sections.

4.3.1 Solving the linear system of the equations

We then implement the procedure described in the processing stage twice (once each for group a_1 and group a_2), where a_1 is in node 16 of Figure 6, and a_2 is in node 29 in the dendrogram depicted in Figure 6. The results are shown in Tables 3 and 4.

4.3.2 Making the identification test

Having computed the indexed weight for the consumer classes, we then carried out the identification procedure to evaluate which class is the best fit. In the identification test, we investigated cross-correlation with an unknown vector (highlighted in red in Figure 5) in order to classify it in two classes. As we can see in this example identification test, the unknown vector is classified in group *A2*. In the same procedure, steps 2 and 3 are repeated according to Figure 2 until the unknown vector convergence in the last type, which contains at least 6 vectors of data related to consumer preferences. The system thus provides the user with a recommendation output, using only the reviews that best fit the disclosed traveler type and preferences on quality indicators disclosed by the user.

4.3.3 Statistical Evaluation

In this case, we set $m=6$ as the size of the vector w_{A_1} (see equation 24), and as such the resulting degree of freedom is: $v = m - 2 = 4$ and by choosing an alpha level of $\alpha=0.05$ we obtain the critical t value to be $t_{a/2}=2.776$. The significant value of r was then calculated as follows:

$$t_{\alpha/2} = \frac{r\sqrt{m-2}}{\sqrt{1-r^2}} \Rightarrow 2.776^2 = \frac{2r^2}{1-r^2} \Rightarrow r = \pm 0.88$$

(27)

In our case, the coefficient r may be characterized as significant when the null hypothesis is rejected ($1 \le |r| \le 0.88$). As a remark, this statistical process recognizes a correlation between the data of two symmetrical data review clusters as normal when $1 \le |r| \le 0.96$. In other case, the method recognizes these correlated data sets as non-homogeneous (uncorrelated).

Table 3: Weight Indexes of Hotel's Quality (group A1)

Review scores of specific traveler types	Cleanliness Score	Location Score	Services Score	Value for Money Score
(*a*) Young couple	0.2805	0.1727	0.2698	0.2769
(*b*) Group of friends	0.1546	0.3777	0.1410	0.3267
(*c*) Family with older children	0.2259	0.0083	0.3462	0.4196
(*d*) Family with young children	0.2698	0.2612	0.2621	0.2070
(*e*) Mature couple	0.2559	0.3511	0.1787	0.2144
(*f*) Solo traveller	0.2605	0.4260	0.1957	0.1178

Table 4: Weight Indexes of Hotel's Quality (group A2)

Review scores of specific traveler types	Cleanliness Score	Location Score	Services Score	Value for Money Score
(*a*) Young couple	0.2521	0.2716	0.2357	0.2405
(*b*) Group of friends	0.2675	0.2601	0.2436	0.2288
(*c*) Family with older children	0.3665	0.0891	0.2996	0.2448
(*d*) Family with young children	0.1909	0.3863	0.1621	0.2607
(*e*) Mature couple	0.2528	0.2702	0.2449	0.2321
(*f*) Solo traveller	0.2560	0.2655	0.2290	0.2495

Table 5: Result of the identification test

Type	Cleanliness Score	Location Score	Services Score	Value for Money Score	Cross-Corr.
A1	0.2805	0.1727	0.2698	0.2769	-0.0133
	0.321	0.256	0.213	0.210	
	0.1546	0.3777	0.1410	0.3267	-0.2705
	0.321	0.256	0.213	0.210	
	0.2259	0.0083	0.3462	0.4196	-0.4834
	0.321	0.256	0.213	0.210	
	0.2559	0.3511	0.1787	0.2144	0.4401
	0.321	0.256	0.213	0.210	
	0.2605	0.4260	0.1957	0.1178	0.4469
	0.321	0.256	0.213	0.210	
A2	0.2521	0.2716	0.2357	0.2405	0.4785
	0.321	0.256	0.213	0.210	
	0.2675	0.2601	0.2436	0.2288	0.8880

	0.321	0.256	0.213	0.210	
	0.3665	0.0891	0.2996	0.2448	0.3091
	0.321	0.256	0.213	0.210	
	0.1909	0.3863	0.1621	0.2607	-0.0357
	0.321	0.256	0.213	0.210	
	0.2528	0.2702	0.2449	0.2321	0.4945
	0.321	0.256	0.213	0.210	
	0.2560	0.2655	0.2290	0.2495	0.5477
	0.321	0.256	0.213	0.210	

5. Discussion, Conclusions and Limitations

We presented a procedure and architecture for a demographic recommender system that takes service quality indicators and consumer/traveler types into account in order to establish the recommendation process. The main theoretical intuition behind the design of such a system is the use of implicitly defined indicators that appeal to different types of travelers, as can be seen on online travel review systems such as Booking.com. The method we present involves a preprocessing stage, a normalization stage with a clustering procedure, and an identification stage where the individual travelers profile is matched with the data provided.

A particular limitation that is evident in this study is the absence of a real-case user evaluation scenario where the evaluation stage could be extended. This is demanding due to the computationally intensive process that is required for the preprocessing of the user-generated content but also for the elicitation of the experimental sample needed to do a more extensive evaluation. This is a future step direction that can be investigated further.

The approach presented here could be further extended using other sources of implicit indicators such as attachment into groups. For example, with regard to the clustering procedure into individual traveler types, it would be possible to take into account fuzzy related measures of attachment to an individual profile, as in the context of a social network (N. Korfiatis & Sicilia, 2007; Sicilia, Garcia-Barriocanal, & Korfiatis, 2008). Another aspect of a secondary input to the recommendation system could be time-dependent preferences as different groups of travelers are known to prefer different seasons (Cai, Feng, & Breiter, 2004; Goodrich, 1978).

With regard to the recommender systems literature, this study contributes on the demographic recommender systems class, by providing a two-dimensional approach to evaluating the preference matrix using service quality indicators and consumer types extracted by user generated content. Computationally, the latter could be quite challenging and, therefore, a limitation of complexity related with the processing time of the clustering procedure outlined in Section 3 could be a significant limitation, since online travel websites receive millions of reviews every day. Another approach that could also be considered as an additional input to the system is the consideration of qualitative properties of online reviews, such as review text readability (García-Barriocanal, Sicilia, & Korfiatis, 2010; Nikolaos Korfiatis, García-Bariocanal, & Sánchez-Alonso, 2012; Wu, Heijden, & Korfiatis, 2011) and fake reviews generated by request of the owners(O'Connor, 2008).

References

Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, *17*(6), 734–749.

Archak, N., Ghose, A., & Ipeirotis, P. (2011). Deriving the pricing power of product features by mining consumer reviews. *Management Science*, *57*(8), 1485–1509.

Atilla, A. (2006). Measuring service quality in the hotel industry: A study in a business hotel in Turkey. *International Journal of Hospitality Management*, *25*(2), 170–192. doi:10.1016/j.ijhm.2005.08.006

Basala, S. L., & Klenosky, D. B. (2001). Travel-style preferences for visiting a novel destination: A conjoint investigation across the novelty-familiarity continuum. *Journal of Travel Research*, *40*(2), 172–182.

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.

Burke, R. (2007). Hybrid Web Recommender Systems. In P. Brusilovsky, A. Kobsa, & W. Nejdl (Eds.), *The Adaptive Web* (Vol. 4321, pp. 377–408). Springer Berlin /

Heidelberg. Retrieved from http://www.springerlink.com/content/9157265121203453/abstract/

Cai, L. A., Feng, R., & Breiter, D. (2004). Tourist purchase decision involvement and information preferences. *Journal of Vacation Marketing*, *10*(2), 138–148.

Chen, Y., Xie, J., & Hall, M. C. (2008). Online consumer review: Word-of-mouth as a new element of marketing communication mix. *Management Science*, *54*(3), 477–491.

Chevalier, J. A., & Mayzlin, D. (2006). The Effect of Word of Mouth on Sales: Online Book Reviews. *Journal of Marketing Research*, *43*(3), 345–354.

Colomo-Palacios, R., García-Crespo, Á., Soto-Acosta, P., Ruano-Mayoral, M., & Jiménez-López, D. (2010). A case analysis of semantic technologies for R&D intermediation information management. *International Journal of Information Management*, *30*(5), 465–469.

Delgado, J., & Davidson, R. (2002). Knowledge bases and user profiling in travel and hospitality recommender systems. In *Proceedings of the 9th International Conference on Information Technologies in Tourism* (pp. 1–16). Presented at the ENTER’02, Innsbruck, Austria: Springer, Heidelberg.

Dulleck, U., & Kerschbamer, R. (2006). On doctors, mechanics, and computer specialists: The economics of credence goods. *Journal of Economic Literature*, 5–42.

Dulleck, U., Kerschbamer, R., & Sutter, M. (2011). The Economics of Credence Goods: An Experiment on the Role of Liability, Verifiability, Reputation, and Competition. *The American Economic Review*, *101*(2), 526–555.

Fesenmaier, D. R., Wöber, K. W., & Werthner, H. (2006). *Destination recommendation systems: Behavioral foundations and applications*. CABI.

García-Barriocanal, E., Sicilia, M.-A., & Korfiatis, N. (2010). Exploring hotel service quality experience indicators in user-generated content: a case using Tripadvisor data. In

Proceedings of the 5th Mediteranian Conference on Information Systems (MCIS 2010). Presented at the MCIS 2010, Tel-Aviv, Israel: Association of Information Systems (AIS).

García-Crespo, Á., Colomo-Palacios, R., Gómez-Berbís, J. M., & Ruiz-Mezcua, B. (2010). SEMO: a framework for customer social networks analysis based on semantics. *Journal of Information Technology*, *25*(2), 178–188.

García-Crespo, Á., López-Cuadrado, J. L., Colomo-Palacios, R., González-Carrasco, I., & Ruiz-Mezcua, B. (2011). Sem-Fit: A semantic based expert system to provide recommendations in the tourism domain. *Expert systems with applications*.

Goodrich, J. N. (1978). The relationship between preferences for and perceptions of vacation destinations: Application of a choice model. *Journal of Travel Research*, *17*(2), 8–13.

Gretzel, U., & Yoo, K. H. (2008). Use and impact of online travel reviews. *Information and communication technologies in tourism 2008*, 35–46.

Huang, Z. (1998). Extensions to the k-means algorithm for clustering large data sets with categorical values. *Data Mining and Knowledge Discovery*, *2*(3), 283–304.

Hudson, S., Hudson, P., & Miller, G. A. (2004). The measurement of service quality in the tour operating sector: A methodological comparison. *Journal of Travel Research*, *42*(3), 305–312.

Johnson, P. A., Sieber, R. E., Magnien, N., & Ariwi, J. (2011). Automated web harvesting to collect and analyse user-generated content for tourism. *Current Issues in Tourism*, 1–7.

Kanungo, T., Mount, D. M., Netanyahu, N. S., Piatko, C. D., Silverman, R., & Wu, A. Y. (2002a). A local search approximation algorithm for k-means clustering. In *Proceedings of the Eighteenth Annual Symposium on Computational Geometry* (pp. 10–18). ACM.

Kanungo, T., Mount, D. M., Netanyahu, N. S., Piatko, C. D., Silverman, R., & Wu, A. Y. (2002b). An efficient k-means clustering algorithm: Analysis and implementation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, *24*(7), 881–892.

Konstan, J. A., & Riedl, J. (2012). Recommender systems: from algorithms to user experience. *User Modeling and User-Adapted Interaction*, 1–23.

Korfiatis, N., & Sicilia, M. A. (2007). Social Measures and Flexible Navigation on Online Contact Networks. In *Proceedings of the IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)* (pp. 1–6). Imperial College, London UK.

Korfiatis, Nikolaos, García-Bariocanal, E., & Sánchez-Alonso, S. (2012). Evaluating content quality and helpfulness of online product reviews: The interplay of review helpfulness vs. review content. *Electronic Commerce Research and Applications*, *11*(3), 205–217. doi:10.1016/j.elerap.2011.10.003

Kumar, N., & Benbasat, I. (2001). Shopping as experience and Web site as a social actor: Web interface design and para-social presence. In *Proceedings of the Twenty-Second International Conference on Information Systems* (pp. 449–454).

Milligan, G. W., & Cooper, M. C. (1986). A study of the comparability of external criteria for hierarchical cluster analysis. *Multivariate Behavioral Research*, *21*(4), 441–458.

Nelson, P. (1970). Information and Consumer Behavior. *Journal of Political Economy*, *78*(2), 311.

Nusair, K. (Khal), & Kandampully, J. (2008). The antecedents of customer satisfaction with online travel services: a conceptual model. *European Business Review*, *20*(1), 4–19. doi:10.1108/09555340810843663

O'Connor, P. (2008). User-generated content and travel: A case study on Tripadvisor. com. *Information and communication technologies in tourism 2008*, 47–58.

Plog, S. C. (1974). Why destination areas rise and fall in popularity. *Cornell Hotel and Restaurant Administration Quarterly*, *14*(4), 55–58.

Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vetterling, W. T. (1992). *Numerical Recipes in C: The Art of Scientific Computing, Second Edition* (2nd ed.). Cambridge University Press.

Reisinger, Y., & Mavondo, F. (2005). Travel Anxiety and Intentions to Travel Internationally: Implications of Travel Risk Perception. *Journal of Travel Research*, *43*(3), 212 –225. doi:10.1177/0047287504272017

Resnick, P., & Varian, H. R. (1997). Recommender systems. *Communications of the ACM*, *40*(3), 56–58.

Ricci, F. (2002). Travel recommender systems. *IEEE Intelligent Systems*, *17*(6), 55–57.

Sarwar, B., Karypis, G., Konstan, J., & Reidl, J. (2001). Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th International Conference on World Wide Web* (pp. 285–295). ACM.

Shapira, B., Taieb-Maimon, M., & Moskowitz, A. (2006). Study of the usefulness of known and new implicit indicators and their optimal combination for accurate inference of users interests. In *Proceedings of the 2006 ACM Symposium on Applied Computing* (pp. 1118–1119). ACM.

Sicilia, M. A., Garcia-Barriocanal, E., & Korfiatis, N. (2008). Using Imperfect Information in Online Social Systems: Applications and Measures of Betweenness. In *Proceedings of the 12th International Conference on Information Processing and Management of Uncertainty*. Presented at the IPMU 2008, Malaga, Spain.

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

Wolfinbarger, M., & Gilly, M. C. (2001). Shopping Online for Freedom, Control, and Fun. *California Management Review*, *43*(2).

Wu, P. F., Heijden, H. V. D., & Korfiatis, N. (2011). The Influences of Negativity and Review Quality on the Helpfulness of Online Reviews. In *Proceedings of the International Conference on Information Systems, 2011*. Presented at the International Conference on Information Systems, Shanghai, China.

Ye, Q., Law, R., & Gu, B. (2009). The impact of online user reviews on hotel room sales. *International Journal of Hospitality Management*, *28*(1), 180–182.

Ye, Q., Law, R., Gu, B., & Chen, W. (2010). 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), 634–639.