Multi-criteria ratings for Recommender Systems: An Empirical Analysis in the Tourism Domain

Matthias Fuchs[1] and Markus Zanker[2]

[1] Mid Sweden University,
83125 Östersund, Sweden
matthias.fuchs@miun.se
[2] Alpen-Adria-Universität Klagenfurt,
9020 Klagenfurt, Austria
markus.zanker@aau.at

Abstract. Most recommendation systems require some form of user feedback such as ratings in order to make personalized propositions of items. Typically ratings are unidimensional in the sense of consisting of a scalar value that represents the user's appreciation for the rated item. Multi-criteria ratings allow users to express more differentiated opinions by allowing separate ratings for different aspects or dimensions of an item. Recent approaches of multi-criteria recommender systems are able to exploit this multifaceted user feedback and make personalized propositions that are more accurate than recommendations based on unidimensional rating data. However, most proposed multi-criteria recommendation algorithms simply exploit the fact that a richer feature space allows building more accurate predictive models without considering the semantics and available domain expertise. This paper contributes on the latter aspects by analyzing multi-criteria ratings from the major etourism platform, TripAdvisor, and structuring raters' overall satisfaction with the help of a Penalty-Reward Contrast analysis. We identify that several a-priori user segments significantly differ in the way overall satisfaction can be explained by multi-criteria rating dimensions. This finding has implications for practical algorithm development that needs to consider different user segments.

1 Introduction

Recommender systems (RS) are tools for consumer decision support that help to overcome information overload in online environments. Their purpose is to point users to items that best match their presumed preferences and needs. Different basic paradigms of recommender systems exist: *collaborative filtering* builds on the assumption that peers with similar ratings and behavior in the past will also have comparable preferences in the future; *content-based filtering* assumes that users' tastes can be semantically described and therefore proposes items whose content descriptions are similar to what is already known that the user likes; *knowledge-based recommendation* systems try to mimic sales agents that exploit domain expertise in order to best possible match elicited customer needs to items.

In the RS literature different interpretations of *knowledge* are common [1], for instance, explicitly encoded knowledge bases that contain sets of logical sentences [2] and case-based recommendation approaches that require domain knowledge for defining similarity functions [3]. The first two families of algorithms tackle recommendation as a Machine Learning task, where from known training data models are learned and subsequently employed to predict unseen or withheld data in a second step. In contrast knowledge-based recommendation in the sense of [2] does not follow an inductive learning approach but requires explicitly coded expertise, for instance, in the form of business rules and constraints that should constitute plausible heuristics and procedures in the eyes of experienced sales personnel. For more details on recommender systems and the genesis of the research field see [1, 4].

The present work now tries to build bridges between recommendation systems that are mainly driven by codified knowledge such as [5, 2] or [6] and majorly learning and data-driven approaches. Algorithms for multi-criteria recommender systems typically learn predictive models by exploiting the different rating dimensions without exploiting any domain specific knowledge [7, 8, 9]. However, we propose to explore models that are in accordance with theoretical findings in the respective application field. In our case the application domain is tourism where empirical findings about the structure of service quality judgments need to be considered. Therefore, we propose to consider user segmentation and to incorporate models from consumer (i.e. tourist) satisfaction research such as the Kano model [10]. Results from analyzing ratings on the major tourism platform, TripAdvisor.com, indicate users belonging to different market (i.e. travel) segments. We apply Brandt's [11] Penalty-Reward-Contrast analysis in order to explore if and how different rating criteria constitute dissatisfiers (i.e. hygiene factors) or excitement factors in different travel segments.

According to Compete, Inc. (2007) in 2006, 52% of U.S. online shoppers visited at least one community website before having bought their travel and tourism services. Moreover, 26% of U.S. tourists deliver feedback on a community website in connection with their trip. The TripAdvisor portal[1] is considered as the biggest and most famous tourism-related social network site worldwide. In numbers, the strongly interlinked portal counts more than 7,000 URLs and shows above 30 million unique users [12]. Visitors around the globe constantly use this web 2.0 portal to write and read assessments about service quality experiences concerning specific hotels. At TripAdvisor user evaluations are recorded both in the form of standardized items and free-texts. Although the online platform is specialized in hotel and accommodation products, a series of chat-room services related to travel services, restaurant services, as well as trip ideas are additionally provided [13].

Next, related work on multi-criteria recommendation is discussed. Section 3 describes the applied methodology, details empirical results and discusses implications for algorithm development. Finally, perspectives on future work are discussed and a conclusion section is provided.

---

[1] www.TripAdvisor.com

2 Related Work

Several lines of work have successfully exploited multi-criteria ratings to improve
the accuracy of recommendations. An early and very encompassing article that
proposes a *contextualized* view on ratings is Adomavicius et al. [7]. Although
their users still provide unidimensional ratings, the situational context of users
adds additional dimensionality to the ratings. This means that ratings are la-
beled with contextual parameters such as user's age, sex or weekday and that
recommendations for a male user in the twenties on a Monday are presumably
more accurate if the system preferably only exploits ratings that have been
added by users in their twenties, who are males and who have rated the item on
a Monday. Adomavicius & Kwon [8] proposed also a recommendation approach
that can exploit real multi-criteria ratings, i.e. a user provides rating values that
appreciate different aspects of the same item. In order to determine the overall
rating of a specific user for an item their algorithm performs three steps: first,
like in traditional recommender systems working with unidimensional ratings
for each criterion a rating value is computed; second an aggregation function
is estimated that allows computing an overall rating from the multiple criteria
ratings; finally, third the overall rating value is computed and recommendable
items are ranked according to the estimated overall rating value.

Jannach et al. [9] further developed the ideas of Adomavicius & Kwon. They
also employed accommodation ratings from a major tourism platform in their
evaluation scenario. They compare regression models that constitute specific
aggregation functions for each user and each item. Their results are, for instance,
that regression models learned with a classifier using a support vector machine
perform better than linear least squares regression models and that a weighted
combination of user and item specific regression models perform best in their
case.

TripAdvisor data is particulary of interest in tourism research, for instance
doing research on complaint management or aggregating review data to the
destination level. However, the work of Graebner et al. [14] is also focusing on
predicting users' rating values. In contrast to our work, Graebner et al. exploit
the users' textual reviews in order to predict the overall rating value. Using both
textual review and the multi-criteria ratings in order to predict the user's overall
assessment value could be a future extension for this work.

3 Empirical analysis

Today, TripAdvisor represents the world largest and most successful social net-
working and community site in tourism comprising over 25 Mio unique users
[12]. The platform facilitates the reviewing of hotels around the world and brings
together individuals in discussion forums and provides users with independent
travel reviews and comments. Figure 1 depicts TripAdvisor's view on the rating
feedback for an accommodation of an arbitrary user. Users can rate a hotel ac-
cording to 7 different dimensions: value for money, quality of rooms, location of

Fig. 1. Detailed view on user rating

the hotel, cleanliness of the hotel, quality of check-in, overall quality of services and particular business services. In addition, users provide unidimensional overall ratings on hotels (not depicted in Figure 1). These standardized evaluation-items are consistently measured on the base of a 5-point scale (i.e. from excellent to terrible). Furthermore, users are explicitly asked if they would recommend the hotel to a friend. By contrast, the recommendation to visit a hotel is measured by two separate binary ratings (i.e. recommend: yes/1; no/0). Standardized assessments of 62,290 unique users concerning hotels from 14 touristic cities (e.g. Vienna, Munich, London, New York, Singapore, San Francisco, Hong Kong, Sydney, Orlando, etc.) were collected during January 2010 by processing web crawling data. Controlled by the user profile (e.g. age, travel motive, trip type, etc.) weighting schemes related to (e.g. hotel) service quality domains are considered as ideal input to parameterize electronic recommender systems ([6]). The goal of the subsequent data analysis is to identify empirical relationships between the users' willingness to recommend a hotel, their partial assessments of the different rating dimensions as well as user profile and context information about their stay.

As a first methodological step the strength of the empirical dependency between users' willingness to recommend the hotel to a friend and their overall rating has been quantified by using a logistic regression. As to be expected, the empirical results clearly show that the overall quality assessment related to a hotel is an excellent determinant for both positive and negative hotel recommendations. From Table 1 emerges that Nagelkerke's R^2's rank well above the threshold value of 0.2 and about 90% of cases are correctly classified by the estimated logistic function ([15]). The difference between the sum of YES and NO values for the dependent variable willingness to recommend and the overall N results from missing values, i.e. users can provide an overall rating and a textual review, but they are not forced to provide detailed ratings for the 7 dimensions or answer if they would recommend the hotel to a friend. As the strong dependency of willingness to recommend on the overall rating value has been confirmed, we will analyze next if the overall rating value strongly depends on the 7 rating dimensions.

Table 1. Results from Logistic Regression - Hotel recommendation and Overall Assessment

Variable in Equation	Regression Coefficient	Wald Statistic	Sig.	Exp(B)
Overall assessment	1.948	13,160.401	.000	7.016
Constant	-5.665	8,852.921	.000	.003
Nagelkerke's R^2	.574			
Log-Likelihood	34,198.27			
% correctly classified:	89.4%			
N	62,290			
Dependent Variable: Recommendation = YES (49.872)/1				
Overall assessment	-2.591	10,619.689	.000	.075
Constant	7.171	7,802.032	.000	1,301.034
Nagelkerke's R^2	.696			
Log-Likelihood	22,161.47			
% correctly classified:	92.4%			
N	62,290			
Dependent Variable: Recommendation = No (9.933)/0				

3.1 Segmentation of user base

The above empirical confirmation of a strong and significant relationship between hotel recommendation and overall quality assessment leads to the next question about the role of the various antecedents (i.e. hotel quality domains) in affecting overall assessments. More precisely, in arriving at a holistic assessment concerning a particular (e.g. hotel) service experience, consumers typically 'weight' their overall assessment according to the relative importance of particular quality dimensions [16]. Applied to our TripAdvisor data, the relative level of determinance of the above mentioned seven hotel quality domains on overall assessment is identified by using multiple regressions [17, 18]. Moreover, in order to show the adequateness of the proposed approach to generate useful input data for electronic recommender systems, regressions were run with respect to four tourist segments previously defined by TripAdvisor data: segment 1: senior couples (i.e. age above 50, leisure trip, staying with spouse in 4/5 star hotel), segment 2: business tourist solo (i.e. age between 35 and 50, business trip, staying alone in 4/5 star hotel), segment 3: budget family tourist (i.e. age between 35 and 50, leisure trip, staying with partner & children in 0-3 star hotel), and, finally, segment 4: youth tourists (i.e. age below 25, leisure trip, staying with friends in 0-3 star hotel). The emerging weighting schemes related to the quality domains in determining the overall assessment are shown in Table 2.

To start with, all models show a strong explanation power (Adj. R^2), are statistically significant (F-Value), and are free of auto-correlated residuals (Durbin Watson) or multi-correlated variables (Variance Inflation Factor). Thus, the quality of TripAdvisor data looks satisfactory for being used to identify how

Table 2. Multiple Regression Results - Determinance of Overall Assessment by Partial Quality Domains*

A-priori Segments	Segment 1: Senior Tourist Couples			Segment 2: Business Tourist Solo		
	Adj. R^2 = .787 F = 114.21 DW = 1.91 N = 1,284 Share = 8.8%			Adj. R^2 = .807 F = 260.24 DW = 1.85 N = 1,366 Share = 9.3%		
Quality domains	Beta	T-Value	VIF	Beta	T-Value	VIF
Value	0.384	8.274	2.169	0.332	9.211	2.927
Rooms	0.247	4.861	2.610	0.271	7.386	3.031
Locations	0.039	1.157	1.175	0.091	3.788	1.273
Cleanliness	0.128	2.692	2.276	0.122	3.264	3.138
Checkin	0.081	1.755	2.089	0.048	1.475	2.386
Service	0.096	1.742	3.057	0.161	4.242	3.221
Business	0.178	4.427	1.625	0.252	2.885	1.718
A-priori Segments	Segment 3: Budget Family Tourist			Segment 4: Youth tourist & friends		
	Adj. R^2 = .769 F = 183.36 DW = 2.23 N = 2,302 Share = 15.7%			Adj. R^2 = .698 F = 45.99 DW = 2.19 N = 875 Share = 6%		
Quality domains	Beta	T-Value	VIF	Beta	T-Value	VIF
Value	0.444	11.026	2.687	0.266	3.442	2.692
Rooms	0.203	4.684	3.103	0.459	5.411	3.245
Locations	0.055	2.010	1.236	0.185	3.443	1.308
Cleanliness	0.179	4.342	2.820	0.081	0.947	3.272
Checkin	0.107	1.991	1.990	0.021	0.287	2.458
Service	0.044	1.077	2.760	0.131	1.662	2.801
Business	0.058	1.952	1.455	0.093	1.514	1.684

* The (adj.) Coefficient of Determination R^2 is the proportion of variability in data accounted for by the statistical model;

Beta is a measure of how strongly a predictor influences the dependent variable;

An F- or a T-test are statistical tests in which the test statistic has an F or a T-distribution under the null hypothesis.

The null hypothesis is rejected if the F or T-value calculated from the data is greater than the critical value of the F- or T-distribution for some desired false-rejection probability (e.g. 0.05).

The Durbin Watson (DW) Test detects autocorrelation (i.e. residuals from a multiple regression model are independent).

Variance Inflation Factor (VIF) quantifies the degree of multicollinearity (i.e. correlated predictor variables) in regression analyses.

various (i.e. hotel) quality domains determine the overall assessment among different customer segments ([18]). For instance, the results show that the relatively strongest and most general determinance stems from both, the perceived 'value for money' and the 'room quality' (i.e. Beta, T-Value > 2; Tab. 2). However, for youth tourists 'room quality' becomes most important, while for the budget family segment 'value-for money' is the most critical unique quality domain.

Interestingly enough, the remaining quality domains are playing completely different roles in determining overall quality assessments among the hotel customer segments. For instance, not only for business tourists, but also for senior tourist couples a convenient 'business environment' seems to be a crucial quality domain, whereas the 'location' factor only plays an insignificant role. Moreover, for business tourists, in contrast to all other segments, the general 'service quality' of the hotel is the third most important quality domain. Furthermore, budget family tourists put the third biggest emphasis on the 'cleanliness of the hotel', while for the youth tourist segment the 'location factor' becomes a relatively important determinant of their overall quality judgement (Table 2).

3.2 Penalty-Reward-Model

Next to the purely quantitative role of quality domains in determining overall assessments, literature also discusses their relevance from a qualitative point of view [19]. Already since the nineties researchers have begun to tackle empirical problems of service quality perception with a multi-factor-structure of customer satisfaction [20]. This model has been adopted and empirically validated both, in a service marketing and tourism context, respectively [21, 22, 23, 24, 25]. The three-factor structure of customer satisfaction was first defined by Kano [10]. Based on his model, quality attributes may be grouped into three categories, each of which exerts a different impact on customer satisfaction:

– *Basic factors* are minimum requirements that cause dissatisfaction if not fulfilled but do not lead to customer satisfaction if fulfilled or exceeded; negative performances with these quality domains has a greater impact on overall satisfaction than a positive one. Hence, basic factors are expected by the customer (i.e. regarded as prerequisites).
– *Excitement factors* are factors that increase customer satisfaction if delivered but do not cause dissatisfaction if they are not delivered; positive performance on these quality dimensions has a greater impact on overall satisfaction than a negative one.
– *Performance factors* lead to satisfaction if performance is high and lead to dissatisfaction if performance is low. In this case, the attribute performance-overall satisfaction relationship is linear and symmetric [24].

Based on Brandt's [11] Penalty-Reward-Contrast analysis a method to empirically decipher the factor-structure of customer satisfaction is presented next. The method employs a dichotomised regression analysis using dummy variables ([18]). More precisely, one set of dummy variables exemplifies in quantitative form excitement factors, while a second set expresses basic factors. In order to

carry out the analysis using our TripAdvisor data the 5 point scales of the in-
dependent variables (i.e. from 5= excellent-1= terrible) were recoded in a way
that scores of 5 were used to form a first dummy variable (i.e. representing
the quantification of the excitement factor with a value of 1). Due to the em-
pirical distribution of the (i.e. independent) variables scores of 3, 2 and 1 (i.e.
relatively low satisfaction) were chosen to create a second dummy representing
the quantification of the basic factor with a value of 1. Although this approach
shows some degree of arbitrariness, the empirical distribution of the raw data is
taken into consideration and is, thus, recommended by [19, 22, 23, 24]. Finally,
empty cells of both dummies were recoded with a value of zero. With the help
of this recoding multiple regression analyses were carried out to quantify basic
requirements and excitement factors using the *overall rating* assessment as the
dependent variable and the two dummy variables for each of the seven quality
domains as independent variables. 'Penalties' can now be expressed as the incre-
mental decline associated with low levels of satisfaction, while 'rewards' become
expressed as the incremental increase associated with high satisfaction to be
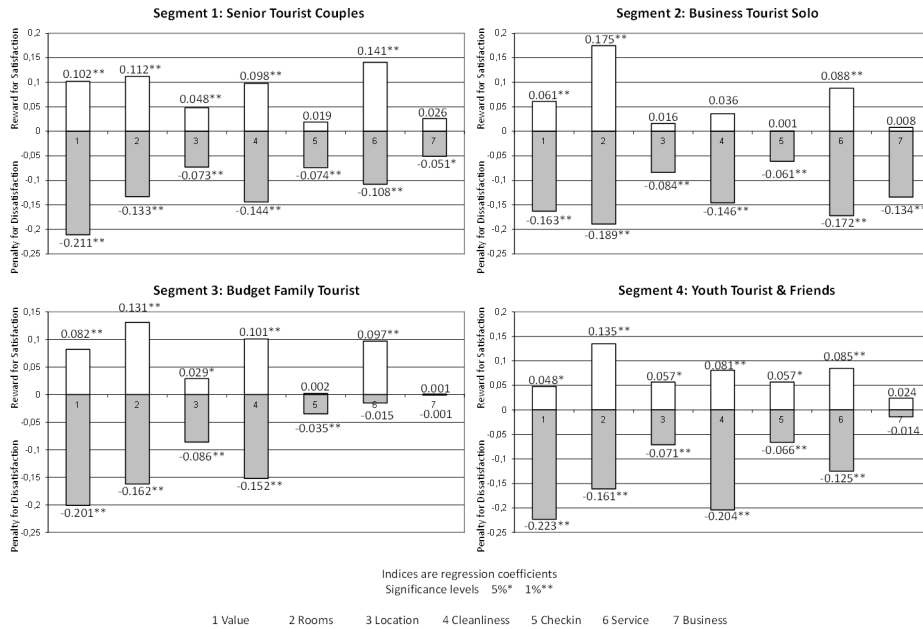observed within a certain hotel quality domain. Thus, if penalty levels surpass

Fig. 2. Penalty Reward Contrast Analysis

reward levels the respective quality domain is a basic factor. Otherwise, if the
reward index surpasses the penalty value the quality dimension should be inter-
preted as an excitement factor. Finally, if reward and penalty values are rather
similar, the quality domain will contribute to tourist satisfaction only when its
level of performance is high. It will lead, at the same time to dissatisfaction when

the performance is low (i.e. performance factor). Using the described Penalty-Reward approach and applying it to the seven hotel quality domains gathered from TripAdvisor data the following results emerged (Fig. 2).

As with the previous tests, again all regressions show a strong explanation power (Adj. R^2 between 0.681 and 0.723), are statistically significant (F-Value between 74.28 and 429.24), and are free of auto-correlated residuals (Durbin Watson between 1.87 and 2.02) and multi-correlated variables (Variance Inflation Factor between 1.224 and 2.087). Thus, the adequateness of TripAdvisor data for indentifying the factor-structure of customer satisfaction for (i.e. hotel) quality domains among various segments could be well confirmed ([18]).

Interestingly enough, for all customer segments results revealed a lack of pure delighting factors (i.e. the positive reward index surpassing the negative penalty value; Figure 2). Obviously, the (i.e. hotel) quality dimensions measured by TripAdvisor are generally perceived as performance or basic factors, respectively. More precisely, while for all customer segments the 'room quality' shows a relatively strong potential to increase overall satisfaction (i.e. if its performance level is high), for senior tourist couples also the 'general service quality' shows potentials to 'delight'. However, at the same time the quality domains 'value for money' 'room quality' and 'cleanliness' show large penalty potentials (i.e. to decrease overall assessment if performance is low). A completely different picture emerged for the business tourist segment, since both 'business convenience' and 'general service quality' emerged as quality dimensions with relatively large penalty potentials (Figure 2). Moreover, for the budget family tourists segment, fully consistent with previous results, next to 'room quality' and 'cleanliness' also the 'location factor' emerged as a quality dimension with strong penalty potentials. Finally, results for the youth tourist segment may be interpreted in analogous fashion.

To summarize, the proposed approach revealed significantly differing determinance profiles (Table 2) and penalty-reward profiles (Figure 2) between the examined customer segments. These insights are particularly valuable, since weighting schemes attached to (e.g. hotel) quality domains build the basis to form overall quality assessments [16, 21]. Thus, segment-specific recommendation strategies might simultaneously consider determinance and penalty-reward profiles, consequently recommending those products (e.g. hotels) that show highest performance values in those quality domains that emerged as significant in respective weighting schemes.

3.3 Discussion and implications

Obviously, the willingness to recommend a hotel to a friend strongly depends on the overall rating value as has been shown in Table 1. The 7 specific rating dimensions from "value for money" to "business services" capture most of the signal to determine the overall rating value (Adj. R^2 clearly above 0.7 for most regression models in Table 2). However, this relationship between multi-criteria ratings and overall rating is clearly moderated by the tourist segment (user profile data and travel context), i.e. the relative influence of the specific rating

dimensions changes for the different segments. Furthermore, a penalty-reward-contrast analysis unveils a qualitative interpretation of multi-criteria ratings as basic and excitement factors and indicates differences for the different segments. To summarize, the proposed methods revealed both, plausible and statistically significant results that have implications for the development of future recommendation algorithms:

– Conversational and knowledge-based recommendation systems can explicitly consider these differing weights users attribute to different criteria when appreciating items and can elicit users' travel segment affiliation in order to adjust utility weights when ranking search results.
– Collaborative multi-criteria recommendation approaches could apply segment specific weights when determining overall ratings and recommended items. Furthermore, the existence of basic and excitement factors indicate non-linear relationships between factor ratings and overall ratings that could be better modeled when ratings are re-coded according to the Kano model.
– Knowing about the qualitative differences in the appreciation of different criteria can be used to generate segment-specific item descriptions and explanations [26, 27] in order to not only more accurately predict items of interest but also to create more persuasive [28] interaction experiences.

Consequently, future research is needed to investigate whether recommendation strategies using segment-specific weighting schemes and a re-coding of ratings according to the Kano model that differentiates between basic and excitement factors outperforms traditional regression models. The approach of [9] actually learns user-specific and item-specific regression models, which means that weights are not adapted on the segment level but instead on the more fine-granular user and item level. However, the approach of [9] cannot be applied to personalize offerings to novel or *cold-start* users and ignores that the same user identity might travel in different contexts, e.g. business trip vs. family.

As a sidenote authors also would like to mention that TripAdvisor has modified its rating criteria since data extraction for this paper took place. The two dimensions *Checkin* and *Business Services* have been deleted and the dimension *Sleep Quality* has been introduced. This partly corresponds to our findings (Figure 2) as Business Services only insignificantly influences overall rating assessments for Segments 3 and 4. However, Checkin has been shown to be a basic factor for all four segments.

4 Conclusions

The paper presented an empirical analysis of multi-criteria ratings harvested from the TripAdvisor portal. The results indicate a significant and strong moderating effect of travel segment in multiple regression models where dimensional ratings predict the users overall rating. Furthermore, a Kano model that qualitatively differentiates rating dimensions into basic and excitement factors indicates non-linear relationships between multi-criteria ratings and overall rating.

The paper outlines how the theoretically plausible and statistically significant findings can serve as a basis for further refinements of future recommendation algorithms.

Acknowledgment

References

1. Jannach, D., Zanker, M., Felfernig, A., Friedrich, G.: Recommender Systems An Introduction. Cambridge University Press (2010)
2. Zanker, M., Jessenitschnig, M., Schmid, W.: Preference reasoning with soft constraints in constraint-based recommender systems. Constraints **15**(4) (2010) 574–595
3. Ricci, F., Venturini, A., Cavada, D., Mirzadeh, N., Blaas, D., Nones, M.: Product recommendation with interactive query management and twofold similarity. In: 5th International Conference on Case-Based Reasoning, Trondheim, Norway, Springer (2003) 479–493
4. Ricci, F., Rokach, L., Shapira, B., Kantor, P., eds.: Handbook on Recommender Systems. Springer (2011)
5. Zanker, M., Bricman, M., Gordea, S., Jannach, D., Jessenitschnig, M.: Persuasive online-selling in quality & taste domains. In: Proc. 7th International Conference on Electronic Commerce and Web Technologies (EC-Web), Krakow, Poland, Springer (2006) 51–60
6. Jannach, D., Zanker, M., Fuchs, M.: Constraint-based recommendation in tourism: A multi-perspective case study. Information Technology & Tourism **11**(2) (2009) 139–155
7. Adomavicius, G., Sankaranarayanan, R., Sen, S., Tuzhilin, A.: Incorporating contextual information in recommender systems using a multidimensional approach. ACM Transactions on Information Systems **23(1)** (2005) 103–145
8. Adomavicius, G., Kwon, Y.: New recommendation techniques for multicriteria rating systems. IEEE Intelligent Systems **22** (2007) 48–55
9. Jannach, D., Gedikli, F., Karakaya, Z., Juwig, O.: Recommending hotels based on multi-dimensional customer ratings. In: International Conference on Information and Communication Technologies in Tourism (ENTER), Springer (2012) 320–331
10. Kano, N.: Attractive Quality and Must-be Quality. Hinshitsu: The Journal of the Japanese Society for Quality Control **14**(2) (1984) 39–48
11. Brandt, R.D.: How service marketers can identify value enhancing service elements. Journal of Services Marketing **2**(3) (1988) 35–41
12. O'Connor, P.: User-generated content and travel - a case study on tripadvisor.com. In: International Conference on Information and Communication Technologies in Tourism (ENTER), Springer (2008) 47–58

13. Dippelreiter, B., Gruen, C., Poettler, M., Seidel, I., Berger, H., Dittenbach, M., Pesenhofer, A.: Online tourism communities on the path to web 2.0 - an evaluation of virtual communities in travel and tourism. Information Technology & Tourism **10**(4) (2007) 329–353

14. Graebner, D., Zanker, M., Fliedl, G., Fuchs, M.: Classification of customer reviews based on sentiment analysis. In: 19th Conference on Information and Communication Technologies in Tourism (ENTER), Springer (2012) 460–470

15. Hosmer, D., Lemeshow, S.: Applied Logistic Regression. 2nd edn. Wiley, New York (2000)

16. Klaus, P.: Quality Epiphenomenon: The Conceptual Understanding of Quality in Face-to-Face Service Encounters. In: The Service Encounter: Managing Employee Customer Interaction in Service Business. Lexington (1985) 17–33

17. Weiermair, K., Fuchs, M.: Measuring tourist judgments on service quality. Annals of Tourism Research **26**(4) (1999) 1004–1021

18. Hair, J.E., Anderson, R.E., Bubin, B.J., Tatham, R.L., Black, W.C.: Multivariate Data Analysis. 6th edn. Prentice-Hall, New York (2006)

19. Busacca, B., Padula, G.: Understanding the relationship between attribute performance and overall satisfaction: Theory, measurement and implications. Marketing Intelligence & Planning **23**(6) (2005) 543–561

20. Johnston, R.: The determinants of service quality: Satisfiers and dis-satisfiers. International Journal of Service Industry Management **6**(1) (1995) 53–71

21. Fuchs, M., Weiermair, K.: New perspectives on satisfaction research in tourism destinations. Tourism Review **58**(3) (2003) 6–14

22. Matzler, K., Sauerwein, E.: The factor structure of customer satisfaction: An empirical test of the importance grid and the penalty-reward-contrast analysis. International Journal of Service Industry Management **13**(4) (2002) 314–332

23. Matzler, K., Bailom, F., Hinterhuber, H., Renzl, B., Pichler, J.: The asymmetric relationship between attribute-level performance and overall customer satisfaction: A reconsideration of the importance-performance analysis. Industrial Marketing Management **33** (2004) 271–277

24. Fuchs, M., Weiermair, K.: Destination benchmarking: An indicator-system's potential for exploring guest satisfaction. Journal of Travel Research **42** (2004) 212–225

25. Mikulic, J., Prebeac, D.: Prioritizing improvement of service attributes using impact range-performance analysis and impact-asymmetry analysis. Managing Service Quality **18**(6) (2008) 559–576

26. Zanker, M., Ninaus, D.: Knowledgable explanations for recommender systems. In: Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI/IAT), IEEE (2010) 657–660

27. Friedrich, G., Zanker, M.: A taxonomy for generating explanations in recommender systems. AI Magazine **32**(3) (2011) 90–98

28. Yoo, K.H., Gretzel, U., Zanker, M.: Persuasive Recommender Systems - Conceptual Background and Implications. Springer (2013)