PROFILING AND RATING PREDICTION FROM MULTI-CRITERIA CROWD-SOURCED HOTEL RATINGS

Fátima Leal
Universidade de Vigo
School of Telecommunications Engineering
INESC TEC, Porto
E-mail: fatimaleal2@gmail.com

Horacio González–Vélez
National College of Ireland
Cloud Competency Centre
E-mail: horacio@ncirl.ie

Benedita Malheiro
Instituto Politécnico do Porto
School of Engineering
INESC TEC, Porto
E-mail: mbm@isep.ipp.pt

Juan Carlos Burguillo
Universidade de Vigo
School of Telecommunications Engineering
E-mail: J.C.Burguillo@uvigo.es

KEYWORDS

Collaborative Filtering, Personalisation, Prediction Models, Multi-criteria Ratings, Tourism, Crowd-Sourcing, Recommender Systems, Data Analytics.

ABSTRACT

Based on historical user information, collaborative filters predict for a given user the classification of unknown items, typically using a single criterion. However, a crowd typically rates tourism resources using multi-criteria, *i.e.*, each user provides multiple ratings per item. In order to apply standard collaborative filtering, it is necessary to have a unique classification per user and item. This unique classification can be based on a single rating – single criterion (SC) profiling – or on the multiple ratings available – multi-criteria (MC) profiling. Exploring both SC and MC profiling, this work proposes: ($\imath$) the selection of the most representative crowd-sourced rating; and ($\imath\imath$) the combination of the different user ratings per item, using the average of the non-null ratings or the personalised weighted average based on the user rating profile. Having employed matrix factorisation to predict unknown ratings, we argue that the personalised combination of multi-criteria item ratings improves the tourist profile and, consequently, the quality of the collaborative predictions. Thus, this paper contributes to a novel approach for guest profiling based on multi-criteria hotel ratings and to the prediction of hotel guest ratings based on the Alternating Least Squares algorithm. Our experiments with crowd-sourced Expedia and TripAdvisor data show that the proposed method improves the accuracy of the hotel rating predictions.

INTRODUCTION

Information and Communication Technology has revolutionised the tourist behaviour. In particular, the mobile technology provides tourists with permanent access to endless web services which influence their decisions. Well-known tourism business-to-customer on-line platforms (*e.g.*, TripAdvisor, Expedia, airbnb, *etc.*) aim to support travellers by providing additional information regarding tourism resources. Furthermore, the on-line tourism-related services enable tourists to share (*e.g.*, photos or videos), comment (*e.g.*, reviews or posts) and rate (*e.g.*, ratings or likes) their travel experiences. Consequently, these tourism services become, while gatherers of voluntarily shared feedback information, crowdsourcing platforms (Egger et al. 2016). The value of the crowd-sourced tourism information is crucial for businesses and clients alike. In the case of this work, it enables the modelling of tourists and tourism resources using multi-criteria ratings to produce suitable recommendations.

Personalised recommendations are often based on the prediction of user classifications, whereby, accurate prediction is essential to generate useful recommendations. Typically, the crowd-sourced classification of hotels involves multi-criteria ratings, *e.g.*, hotels are classified in the Expedia platform in terms of *cleanliness*, *hotel condition*, *service & staff*, *room comfort* and *overall* opinion. We argue that the personalised combination of multi-criteria item ratings improves the tourist profile and, consequently, the accuracy of the collaborative predictions.

Collaborative filtering is a classification-based technique, *i.e.*, depends on the classification each user gave to the items he/she was exposed to (Breese et al. 1998). Typically, this classification corresponds to a unique rating. Whenever the crowd-sourced data holds multiple ratings per user and item, first, it is necessary to decide which user classification to use in order to apply collaborative filtering. This work explores both single criterion (SC) – chooses the most representative of the crowd-sourced user ratings (Leal et al. 2017) – and multi-criteria (MC) profiling approaches –combines the different crowd-sourced user ratings per item, using the Non-Null Rating Average (NNRA) or the Personalised Weighted Rating Average (PWRA), *i.e.*, based on the individual user rating profile.

This research contributes to guest and hotel profiling –

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

based on multi-criteria ratings – and to the prediction of hotel guest ratings – based on the the Alternating Least Squares with Weighted-λ-Regularisation (ALS-WR) matrix factorisation algorithm. Our experiments with crowd-sourced Expedia and TripAdvisor data proved that the proposed profiling method improves the ALS-WR prediction accuracy of unknown hotel ratings. In particular, when faced with null multi-criteria user ratings, the most accurate predictions were achieved with the Personalised Weighted Rating Average combination.

This paper is organised as follows. The related work section reviews personalisation via crowd-sourced ratings. The proposed method section describes the approach and algorithms used. The experiments and tests section reports the data set, tests performed and the results obtained. Finally, the conclusions section summarises and discusses the outcomes of this work.

RELATED WORK

Technology plays an important role in the hotel and tourism industry. Both tourists and businesses benefit from technology advances regarding communication, reservation and guest feedback services. Tourists use tourism Web services to organise trips, *i.e.*, to search, book and share their opinions in the form of ratings, textual reviews, photos, *etc.*, creating a digital footprint. This permanent interaction between tourists and tourism Web services and mobile applications generates large volumes of precious data.

The tourist profiles, which are based on the individual digital footprints, are used by recommendation systems to personalise recommendations. Thus, refined tourist profiles will increase the quality of the recommendations and, ultimately, the tourist experience.

Collaborative filtering is a popular recommendation technique in the tourism domain. It often relies on rating information voluntarily provided by tourists, *i.e.*, crowd-sourced ratings, to recommend unknown resources to other tourists. Well-known tourism crowdsourcing platforms, *e.g.*, TripAdvisor or Expedia, allow users to classify tourism resources using multi-criteria, *e.g.*, *overall*, *service*, *cleanliness*, *etc*. Thus, profiling and prediction using tourism crowd-sourced multi-criteria ratings is an important research topic for the hospitality industry.

Adomavicius and Kwon (2015), Bilge and Kaleli (2014), Lee and Teng (2007), Jhalani et al. (2016), Liu et al. (2011), Manouselis and Costopoulou (2007), and Shambour et al. (2016) have explored the integration of multi-criteria ratings in the user profile, mainly using multimedia data sets to validate their proposals. However, scant research considers crowd-sourced multi-criteria ratings for profiling and rating prediction applied to the tourism domain.

Jannach et al. (2012) apply the Adomavicius and Kwon (2007) methods to incorporate multi-criteria ratings in the tourist profile based on Support Vector Regression (SVR). It combines a user and item models, using a weighted approach, to provide better recommendations. The evaluation was performed with a data set provided by HRS.com.

Fuchs and Zanker (2012) perform multi-criteria rating analysis based on a TripAdvisor data set. First, they use Multiple Linear Regression (MLR) to identify correlations, patterns, and trends among the TripAdvisor data set parameters. Then, the authors apply the Penalty-Reward-Contrast analysis proposed by Randall Brandt (1988) to establish tourist satisfaction levels based on multi-criteria ratings. This work proposes a methodology for MC rating analysis.

Nilashi et al. (2015) propose a SC profiling approach together with a hybrid hotel recommendation model for multi-criteria recommendation. They employed: (*i*) Principal Component Analysis (PCA) for the selection of the most representative rating (dimensionality reduction); (*ii*) Expectation Maximisation (EM) and Adaptive Neuro-Fuzzy Inference System (ANFIS) as prediction techniques; and (*iii*) TripAdvisor data for evaluation.

Farokhi et al. (2016) explore SC profiling together with collaborative filtering. First, the authors selected the *overall* as the most representative rating after determining the correlation between the multiple ratings, then applied data clustering (Fuzzy c-means and k-means) to find the nearest neighbours and, finally, predicted the unknown hotel ratings using the Pearson Correlation coefficient. The evaluation was performed with TripAdvisor data.

Ebadi and Krzyzak (2016) developed an intelligent hybrid multi-criteria hotel recommender system. The system uses both textual reviews and ratings from TripAdvisor. Regarding the ratings, it adopts SC profiling to learn the guest preferences and Singular Value Decomposition (SVD) matrix factorisation to predict unknown ratings.

Contributions

This paper explores profiling and prediction using tourism crowd-sourced multi-criteria ratings. The main goal is to refine guest and hotel profiling by reusing the multiple hotel ratings each guest shares. According to Nilashi et al. (2015) and Adomavicius and Kwon (2015), collaborative filtering with multi-criteria item ratings has been unexplored when compared with its single criterion item rating counterpart. The current work proposes and compares different ways of utilising multi-criteria user ratings to improve the accuracy of predictions. Furthermore, when compared with other research found in the literature (Table 1), our work uses: (*i*) single and multiple rating profiling; (*ii*) employs ALS-WR as predictive technique; and (*iii*) Expedia (E) and TripAdvisor (TA) crowd-sourced data for evaluation.

TABLE 1: Comparison of Multi-Criteria Research Approaches

Approach	Evaluation	Profiling	Prediction
Jannach et al. (2012)	HRS	MC	SVR
Fuchs and Zanker (2012)	TA	MC	–
Nilashi et al. (2015)	TA	SC	ANFIS
Farokhi et al. (2016)	TA	SC	k-means
Ebadi and Krzyzak (2016)	TA	SC	SVD
Leal et al. (2017)	**TA & E**	**SC & MC**	**ALS-WR**

PROPOSED METHOD

Typically, a collaborative recommendation filter relies on an unique rating to produce recommendations, *i.e*, in standard rating-based recommendation systems, the user is modelled using a single rating. However, in tourism crowdsourcing platforms, the tourist-related data encompasses multi-criteria ratings.

This paper addresses the problem of personalisation via crowd-sourced multi-criteria tourism ratings. Our proposed method has four modules: (*i*) Data Collection for gathering data from Expedia platform; (*ii*) Rating Analysis for exploring distinct profiling approaches based on multi-criteria ratings; (*iii*) Rating Prediction for predicting unknown ratings; and (*iv*) Evaluation Metrics for assessing the profiling and recommendation results.

Data Collection

Expedia (`http://www.expedia.com`) is a powerful platform which contains large volumes of crowd-sourced hotel opinions. Moreover, Expedia owns a host of on-line brands, including TripAdvisor, Hotels.com or trivago. According to Law and Chen (2000) (Law and Chen 2000), Expedia brands cover researching, booking, experiencing and sharing travels. The platform allows choosing flights or hotels, reading personal reviews of hotels, classifying hotels using textual reviews and ratings as well as planning new travels.

Taking into account these characteristics, we collected different crowd-sourced ratings via the Expedia API (`http://developer.expedia.com/directory`). In the Expedia platform tourists classify hotels using multi-criteria ratings: *overall*, *cleanliness*, *hotel condition*, *service & staff* and *room comfort*. Based on these multiple criteria classifications, we create, using different approaches, unique personalised ratings per tourist and hotel.

Rating Analysis

The rating analysis module explores different profiling approaches based on crowd-sourced multi-criteria ratings. First, we apply a Multiple Linear Regression (MLR) to identify the Most Representative Rating (MRR). Then, we combine the crowd-sourced multi-criteria user ratings into a single rating using NNRA and PWRA.

Multiple Linear Regression is typically applied to multivariate scenarios in order to predict one or more continuous variables based on other data set attributes, *i.e.*, by identifying existing dependencies among variables (Sykes 2000). First, we determine the correlation between the multi-criteria ratings to identify the dependent variable and, then, perform MLR to estimate the relation between the identified dependent variable and the remaining set of explanatory variables. Equation 1 displays the model of the MLR with k regression variables where ϵ_i is the disturbance, β_0 is the intercept and β_i ($i = 1$ to k) are the partial regression coefficients, representing the rate of change of Y as a function of the changes of $X = \{x_1, x_2, ..., x_k\}$ (Tranmer and Elliot 2008).

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + ... + \beta_k x_{ki} + \epsilon_i \quad (1)$$

We use Ordinary Least Squares (OLS) to estimate the unknown parameters (β_{ki}) of this linear regression model. OLS minimises the distance between the observed responses and the responses predicted by the linear approximation (Stone and Brooks 1990). Equation 2 represents the OLS method where x_i and y_i are the observations and $\hat{x}$ and $\hat{y}$ the predictions.

$$\widehat{\beta} = \frac{\sum(x_i - \hat{x})(y_i - \hat{y})}{\sum(x_i - \hat{x})^2} \quad (2)$$

Rating Combination explores two multi-criteria item rating combination methods: (*i*) the Non-Null Rating Average (NNRA); and (*ii*) the Personalised Weighted Rating Average (PWRA). The non-null rating average $r_{u,i}$ is defined by Equation 3 where $r_{u,i,j}$ is the non-null rating of type j given by user u to item i and n is the number of non-null ratings given by user u to item i.

$$r_{u,i} = \frac{\sum_{j=1}^{n} r_{u,i,j}}{n} \quad (3)$$

Equation 4 displays the personalised weighted rating average $r_{u,i}$ where $r_{u,i,j}$ is the non-null rating of type j given by user u to item i, $n_{u,j}$ the number of non-null ratings of type j given by user u and n is the total number of non-null multi-criteria ratings given by user u.

$$r_{u,i} = \frac{\sum_{j=1}^{n} n_j r_{u,i,j}}{\sum_{j=1}^{n} n_{i,j}} \quad (4)$$

Rating Prediction

The rating prediction module aims to predict unknown hotel ratings, *i.e.*, hotels not yet classified by the tourists, by implementing a user-based collaborative recommendation filter. We use the Alternating-Least-Squares with Weighted-λ-Regularisation (ALS-WR) algorithm since, according to (Zhou et al. 2008, Hu et al. 2008), it provides better results than other matrix factorisation approaches despite its higher execution time. ALS-WR employs matrix factorisation to represent tourists and hotels as vectors of latent factors. The rating matrix (R_{u*i}) holds for all users and items the corresponding user u item i rating. For recommendations purposes, the algorithm factorises the matrix R_{u*i} into two latent matrices: (*i*) the user-factor matrix P; and (*ii*) the item-factor matrix Q. Equation 5 represents this factorisation where each row p_u of P or q_i of Q represents the relation between the corresponding latent factor and the user u or item i, respectively, and λ regularises the learned factors (Friedman et al. 2016).

$$\min_{P,Q} \sum_{r_{u,i} \in R} \left[(r_{u,i} - p_u q_i^{\intercal})^2 + \lambda \left(||p_u||^2 + ||q_i||^2 \right) \right] \quad (5)$$

Finally, R is approximated by the product of P and Q, *i.e.*, each known rating $r_{u,i}$ is approximated by $\hat{r}_{u,i} = p_u \cdot q_i^{\intercal}$.

Algorithm 1 summarises the ALS-WR iterative implementation. In each iteration, P and Q are sequentially fixed to solve the optimisation problem. Once the latent vectors converge, the algorithm ends. We defined the regularisation weight λ, the dimensionality of latent feature space (k) and the number of iterations (n) based on the above mentioned research works. The final matrix holds all user item rating predictions used for recommendation.

Algorithm 1 ALS-WR

Inputs	User u, Item i and $r_{u,i}$
Outputs	User u, Item i and $\hat{r}_{u,i}$
Step 1	Matrix Factorisation with $\lambda = 0.1$, $k = 20$ and $n = 20$
Step 2	Create the P and Q latent matrices
Step 3	Fix Q and estimate P
Step 4	Apply ALS-WR
Step 5	Fix P and estimate Q
Step 6	Apply ALS-WR
Step 7	Calculate prediction matrix

Evaluation Metrics

The evaluation of recommendation systems involves predictive accuracy and classification metrics. On the one hand, the predictive accuracy metrics measure the error between the predicted rating and the real user rating. It is the case of the Mean Absolute Error (MAE), which measures the average absolute deviation among the predicted rating and the real rating, or the Root Mean Square Error (RMSE), which highlights the largest errors (Herlocker et al. 1999). Equation 6 and Equation 7 represent both error functions where $\hat{r}_{u,i}$ represents the rating predicted for user u and item i, $r_{u,i}$ the rating given by user u to item i, m the total number of users and n the total number of items.

$$MAE = \frac{1}{u} \times \sum_{u=1}^{m} \left(\frac{1}{n} \times \sum_{i=1}^{n} |\hat{r}_{u,i} - r_{u,i}| \right) \tag{6}$$

$$RMSE = \frac{1}{u} \times \sum_{u=1}^{m} \left(\sqrt{\frac{1}{n} \times \sum_{i=1}^{n} (\hat{r}_{u,i} - r_{u,i})^2} \right) \tag{7}$$

On the other hand, the classification accuracy metrics, which include Precision and Recall and range from 1 (best) to 0 (worst) (Basu et al. 1998). Recall determines the percentage of relevant items selected from the total number of relevant items available (Equation 9). Precision defines the percentage of relevant items selected from the total number of items (Equation 8). Equation 9 and Equation 8 detail both metrics where TP is the number of relevant items recommended by the system or true positives, FN is the number of relevant items not recommended by the system or false negatives and FP corresponds to the number of irrelevant items recommended by the system or false positives.

$$Precision = \frac{TP}{TP + FP} \quad (8) \quad Recall = \frac{TP}{TP + FN} \quad (9)$$

Finally, the quality of the top N recommendations can be determined using Recall@N metric. In particular, Nilashi et al. (2015) define Recall@N according to Equation 10, where TP is the number of true positive or relevant items and Top_N is the list of the top N recommended items.

$$Recall@N = \frac{TP \cap Top_N}{TP} \tag{10}$$

EXPERIMENTS AND RESULTS

We conducted several off-line experiments with the HotelExpedia data set (`http://ave.dee.isep.ipp.pt/~1080560/ExpediaDataSet.7z`) and the TripAdvisor data set (Wang et al. 2010) to evaluate the proposed method. The data processing was implemented in Python using the *scikit-learn* library (`http://scikit-learn.org`). Our system is running on a cloud OpenStack instance, holding 16 GB RAM, 8 CPU and 160 GB hard-disk. The experiments involved MRR, NNRA and PWRA profiling, rating prediction and rating prediction evaluation.

Data Sets

The experiments were performed with the HotelExpedia and TripAdvisor data sets. The data set was randomly partitioned into training (75 %) and test (25 %) in order to perform the off-line profiling and rating prediction.

Expedia Table 2 describes the contents of the data set. It is composed of 6276 hotels, 1090 identified users and 214 342 reviewers from 11 different locations. Each user classified at least 20 hotels and each hotel has a minimum of 10 ratings. Our experiments, which rely on the hotel, user and hotel user review data, use, specifically, the user nickname, the hotel identification and, as multi-criteria ratings, the *overall*, *cleanliness*, *service & staff*, *hotel condition* and *room comfort*. This data set does not contain null ratings, *i.e.*, all users rated the hotels according to the multiple criteria.

TABLE 2: HotelExpedia Data Set

Entities	**Features**
Hotels	`hotelId`, `description`, `latitude-longitude`, `starRating`, `guestReviewCount`, `price`, `amenity`, `overall`, `recommendedPercent`, `cleanliness`, `serviceAndstaff`, `roomComfort`, `hotelCondition`
Users & Reviews	`nickname`, `userLocation`, `hotelId`, `overall`, `cleanliness`, `hotelCondition`, `serviceAndStaff`, `roomComfort`, `reviewText` `timestamp`

TripAdvisor Table 3 describes the contents of the data set, which is composed of 9114 hotels, 7452 users and 235 793 hotel reviews. Our experiments reuse the user and hotel identification and, as multi-criteria ratings, the *overall*, *value*, *rooms*, *location*, *cleanliness*, *service*, and *sleep quality*. This data set contains 14 % of null ratings.

TABLE 3: TripAdvisor Data Set

Entities	**Features**
Hotels	`name`, `hotelURL`, `price`, `hotelID`, `imgURL`
Users & Reviews	`authorLocation`, `title`, `author`, `reviewID`, `reviewText`, `date`, `overall`, `value`, `rooms`, `location`, `cleanliness`, `service`, `sleepQuality`

Rating Analysis and Prediction

First, we analysed the available multi-criteria guest ratings per hotel and, then, applied Algorithm 1 to predict the unknown hotel ratings. The rating analysis comprised two different approaches: (*i*) the identification of the most representative hotel rating; and (*ii*) the combination of the multi-criteria guest ratings per hotel into a unique guest rating per hotel.

Most Representative Rating This rating analysis determines the correlation between the multiple hotel ratings to recognise the most correlated rating and, then, estimates and quantifies the relationship between this rating (dependent variable) and the remaining ratings (independent variables) using Multiple Linear Regression. The *overall* rating resulted as the most correlated rating (dependent variable) and, thus, can be estimated in terms of the remaining ratings (independent variables) for both HotelExpedia and TripAdvisor data sets. Table 4 displays the OLS MLR results where β_i are the regression coefficients and R^2 quantifies the response variable variation that is explained by the model.

TABLE 4: MLR Results for the Overall Rating

Data Set	Independent Features	β_i	R^2
Hotel Expedia	Service & Staff	0.32	0.80
	Hotel Condition	0.30	
	Room Comfort	0.29	
	Cleanliness	0.11	
Trip Advisor	Value	0.23	0.78
	Service	0.22	
	Rooms	0.18	
	Cleanliness	0.14	
	Location	0.12	
	Sleep Quality	0.10	

In the case of HotelExpedia, the results show that the independent variables (*cleanliness*, *hotel condition*, *room comfort* and *service & staff*) are capable of explaining approximately 80 % of the dependent variable. The regression was performed with 214 343 multi-criteria ratings. In the case of TripAdvisor, Leal et al. (2016) report that the independent variables (cleanliness, location, rooms, service, sleep quality and value) are capable of explaining approximately 78 % of the dependent variable (*overall*). Based on these results, we chose the *overall* rating as the Most Representative Rating (MRR) of both HotelExpedia and TripAdvisor and, then, performed the *overall* rating prediction. Figure 1 plots the Normalised RMSE (NRMSE) of the predictions for the training and test data partitions of both data sets. In both cases the NRMSE decreases monotonically and converges over time to approximately 0.138 (training) and 0.196 (test) using Expedia data and 0.05 (training) and 0.215 (test) using TripAdvisor data.

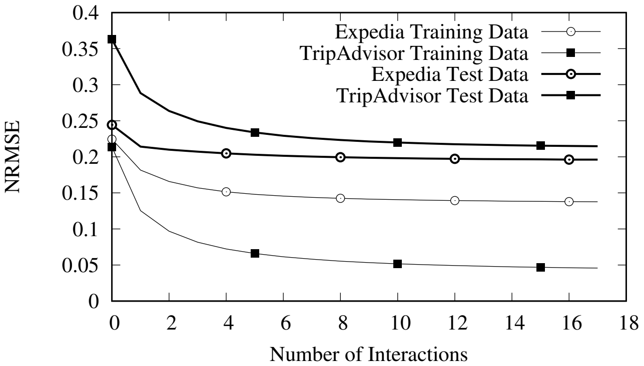

Fig. 1. NRMSE of the Predictions with MRR Profiling

Rating Combination The second rating analysis combines the multi-criteria guest ratings per hotel into a single guest rating per hotel. As a first approach, we calculated the Non-Null Rating Average (NNRA) with Equation 3 and performed the rating prediction using the NNRA rating. Figure 2 plots the NRMSE of the predictions for the training and test data partitions. In both cases the NRMSE decreases monotonically and converges over time to 0.123 (training) and 0.167 (test) using Expedia data and 0.045 (training) and 0.191 (test) using TripAdvisor data.

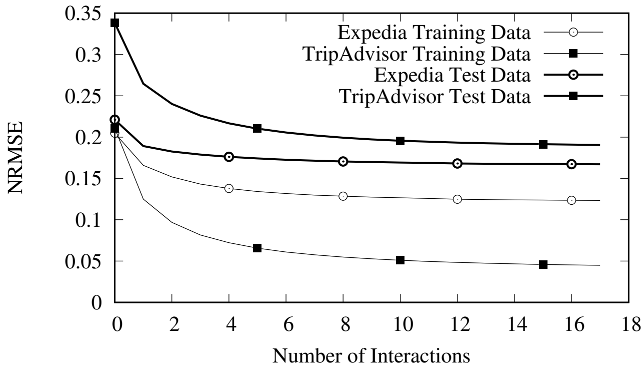

Fig. 2. NRMSE of the Predictions with NNRA Profiling

As an alternative combination approach, we applied the Personalised Weighted Rating Average (PWRA), according to Equation 4, and generated the predictions. Figure 3 plots the NRMSE of the training and test data rating predictions based on the PWRA rating. The NRMSE decreases monotonically and converges over time to 0.123 (training) and 0.167 (test) for Expedia and 0.045 (training) and 0.186 (test) for TripAdvisor.

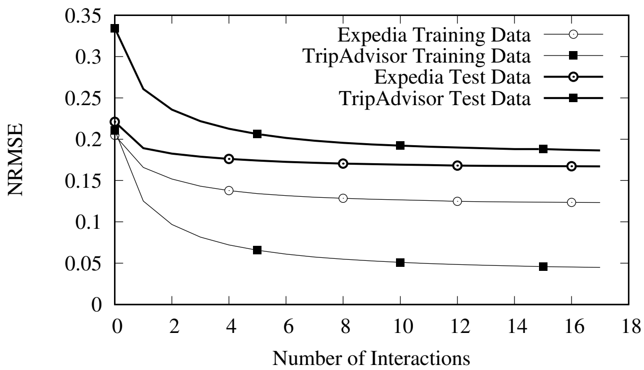

Fig. 3. NRMSE of the Predictions with PWRA Profiling

Recommendations

Finally, we recommend hotels to potential guests with the support of ALS-WR predictions and PWRA profiling. Our novel profiling approach reuses the complete collection of multi-criteria hotel classifications available. To obtain recommendations, the user introduces a desired location and the application provides the top 10 hotel predictions for the user and location. The effectiveness of this recommendation engine was measured using the Recall@10, *i.e.*, considering the top 10 hotel predictions per user.

Discussion

Table 5 compares the global predictive (NRMSE and NMAE), and classification (Recall and Recall@10) accuracy for the test data with the most representative rating (MRR), the Non-Null Rating Average (NNRA) and the Personalised Weighted Rating Average (PWRA) profiling approaches. The results correspond to the average of ten tests. Lower error values and higher classification values indicate higher prediction accuracy. Since the global Precision is one (1), we only present Recall-based classification results. The MMR profiling, which corresponds to the usage of the standard *overall* rating, is the base profiling approach.

The NNRA and PWRA results with the HotelExpedia data set, which has no null ratings, are naturally equal, whereas, with the TripAdvisor data set, which includes 14 % of null ratings, they are not only distinct, but favourable to PWRA. In particular, with HotelExpedia, the NNRA and PWRA profiling, when compared with the MMR approach, improve the NRMSE 14.8 %, the NMAE 26.6 %, the Recall 5.5 % and the Recall@10 17.9 %. In the TripAdvisor case, the results with the PWRA profiling, when compared with those of the MMR approach, improve the NRMSE 13.5 %, the NMAE 14.7 %, the Recall 6.6 % and the Recall@10 16.1 %.

TABLE 5: Prediction Metrics Results

	Profiling	NRMSE	NMAE	Recall	Recall@10
Hotel Expedia	MRR	0.196	0.173	0.254	0.801
	NNRA	0.167	0.127	0.268	0.944
	PWRA	0.167	0.127	0.268	0.944
Trip Advisor	MRR	0.215	0.143	0.351	0.753
	NNRA	0.191	0.125	0.363	0.822
	PWRA	0.186	0.122	0.374	0.874

In terms of the accuracy of the rating predictions, these results show that: (*i*) NNRA and PWRA are preferable to MRR profiling; and (*ii*) PWRA, when faced with null multi-criteria user ratings, outperforms both MMR and NNRA profiling.

CONCLUSIONS

Tourism crowdsourcing platforms, *e.g.*, Expedia and TripAdvisor, collect large volumes of feedback data regarding tourism resources, including multi-criteria ratings, textual reviews, photos, *etc*. The crowd-sourced tourist profile corresponds this individual digital footprint.

The present work explores crowd-sourced multi-criteria rating profiling together with collaborative filtering to provide hotel recommendations. In order to apply standard collaborative filtering, it is necessary to provide the filter with a single classification per user and item. To address this problem, *i.e.*, use multi-criteria ratings for profiling, we designed and experimented with two main approaches: (*i*) the identification of the most representative rating (MRR) with MLR; and (*ii*) the combination of the multi-criteria ratings into a single rating per user and item using NNRA and PWRA. The predictions were performed using the ALS-WR matrix factorisation technique.

The experiments, which were conducted with Expedia and TripAdvisor crowd-sourced multi-criteria hotel ratings, showed that the highest ALS-WR prediction accuracy occurs with the personalised weighted rating average profiling. Based on these results, we adopted the PWRA profiling for the prediction of hotel guest ratings.

In terms of contributions, this research work provides a novel profiling approach based on crowd-sourced multi-criteria ratings which improves the ALS-WR hotel rating prediction accuracy.

As future work, we intend to: (*i*) cluster hotels taking into account their crowd-sourced value for money; and (*ii*) explore multi-criteria recommendation using both textual reviews and multi-criteria ratings.

ACKNOWLEDGEMENTS

This work was partially financed by: (*i*) the European Regional Development Fund (ERDF) through the Operational Programme for Competitiveness and Internationalisation - COMPETE Programme - within project ≪FCOMP-01-0202-FEDER-023151≫ and project ≪POCI-01-0145-FEDER-006961≫, and by National Funds through Fundação para a Ciência e a Tecnologia (FCT) - Portuguese Foundation for Science and Technology - as part of project UID/EEA/50014/2013; and (*ii*) ICT COST Action IC1406 High-Performance Modelling and Simulation for Big Data Applications (cHiPSet).

REFERENCES

Adomavicius, G. and Kwon, Y.: 2007, New recommendation techniques for multicriteria rating systems, *IEEE Intelligent Systems* **22**(3).

Adomavicius, G. and Kwon, Y.: 2015, Multi-criteria recommender systems, *Recommender systems handbook*, Springer, pp. 847–880.

Basu, C., Hirsh, H., Cohen, W. et al.: 1998, Recommendation as classification: Using social and content-based information in recommendation, *AAAI/IAAI*, pp. 714–720.

Bilge, A. and Kaleli, C.: 2014, A multi-criteria item-based collaborative filtering framework, *Computer Science and Software Engineering (JCSSE), 2014 11th International Joint Conference on*, IEEE, pp. 18–22.

Breese, J. S., Heckerman, D. and Kadie, C.: 1998, Empirical analysis of predictive algorithms for collaborative filtering, *Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence*, Morgan Kaufmann Publishers Inc., pp. 43–52.

Ebadi, A. and Krzyzak, A.: 2016, A hybrid multi-criteria hotel recommender system using explicit and implicit feedbacks, *World Academy of Science, Engineering and Technology, International Journal of Computer, Electrical, Automation, Control and Information Engineering* **10**(8), 1377–1385.

Egger, R., Gula, I. and Walcher, D.: 2016, *Open Tourism: Open Innovation, Crowdsourcing and Co-Creation Challenging the Tourism Industry*, Springer.

Farokhi, N., Vahid, M., Nilashi, M. and Ibrahim, O.: 2016, A multi-criteria recommender system for tourism using fuzzy approach, *Journal of Soft Computing and Decision Support Systems* **3**(4), 19–29.

Friedman, A., Berkovsky, S. and Kaafar, M. A.: 2016, A differential privacy framework for matrix factorization recommender systems, *User Modeling and User-Adapted Interaction* **26**(5), 425–458.

Fuchs, M. and Zanker, M.: 2012, Multi-criteria ratings for recommender systems: an empirical analysis in the tourism domain, *International Conference on Electronic Commerce and Web Technologies*, Springer, pp. 100–111.

Herlocker, J. L., Konstan, J. A., Borchers, A. and Riedl, J.: 1999, An algorithmic framework for performing collaborative filtering, *Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval*, ACM, pp. 230–237.

Hu, Y., Koren, Y. and Volinsky, C.: 2008, Collaborative filtering for implicit feedback datasets, *2008 Eighth IEEE International Conference on Data Mining*, Ieee, pp. 263–272.

Jannach, D., Gedikli, F., Karakaya, Z. and Juwig, O.: 2012, *Recommending Hotels based on Multi-Dimensional Customer Ratings*, Springer Vienna, Vienna, pp. 320–331.

Jhalani, T., Kant, V. and Dwivedi, P.: 2016, *A Linear Regression Approach to Multi-criteria Recommender System*, Springer International Publishing, Cham, pp. 235–243.

Law, R. and Chen, F.: 2000, Internet in travel and tourism-part ii: Expedia, *Journal of Travel & Tourism Marketing* **9**(4), 83–87.

Leal, F., Dias, J. M., Malheiro, B. and Burguillo, J. C.: 2016, Analysis and visualisation of crowd-sourced tourism data, *Proceedings of the Ninth International C* Conference on Computer Science & Software Engineering*, C3S2E '16, ACM, Porto, pp. 98–101.

Leal, F., Malheiro, B. and Burguillo, J. C.: 2017, Prediction and analysis of hotel ratings from crowd-sourced data, *in* A. Rocha, A. M. Correia, H. Adeli, S. Costanzo and L. P. Reis (eds), *New Advances in Information Systems and Technologies*, Springer International Publishing, Cham, pp. 1–10.

Lee, H.-H. and Teng, W.-G.: 2007, Incorporating multi-criteria ratings in recommendation systems, *Information Reuse and Integration, 2007. IRI 2007. IEEE International Conference on*, IEEE, pp. 273–278.

Liu, L., Mehandjiev, N. and Xu, D.-L.: 2011, Multi-criteria service recommendation based on user criteria preferences, *Proceedings of the fifth ACM conference on Recommender systems*, ACM, pp. 77–84.

Manouselis, N. and Costopoulou, C.: 2007, Analysis and classification of multi-criteria recommender systems, *World Wide Web* **10**(4), 415–441.

Nilashi, M., bin Ibrahim, O., Ithnin, N. and Sarmin, N. H.: 2015, A multi-criteria collaborative filtering recommender system for the tourism domain using expectation maximization (em) and pca–anfis, *Electronic Commerce Research and Applications* **14**(6), 542–562.

Randall Brandt, D.: 1988, How service marketers can identify value-enhancing service elements, *Journal of Services Marketing* **2**(3), 35–41.

Shambour, Q., Hourani, M. and Fraihat, S.: 2016, An item-based multi-criteria collaborative filtering algorithm for personalized recommender systems, *International Journal of Advanced Computer Science and Applications* **7**(8), 275–279.

Stone, M. and Brooks, R. J.: 1990, Continuum regression: cross-validated sequentially constructed prediction embracing ordinary least squares, partial least squares and principal components regression, *Journal of the Royal Statistical Society* pp. 237–269.

Sykes, A. O.: 2000, An introduction to regression analysis, *in* E. A. Posner (ed.), *Chicago Lectures in Law and Economics*, Foundation Press, New York.

Tranmer, M. and Elliot, M.: 2008, Multiple linear regression, *Technical report*, The Cathie Marsh Centre for Census and Survey Research (CCSR).

Wang, H., Lu, Y. and Zhai, C.: 2010, Latent aspect rating analysis on review text data: a rating regression approach, *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, ACM, pp. 783–792.

Zhou, Y., Wilkinson, D., Schreiber, R. and Pan, R.: 2008, Large-scale parallel collaborative filtering for the netflix prize, *International Conference on Algorithmic Applications in Management*, Springer, pp. 337–348.

FÁTIMA LEAL holds a B.Sc. and an M.Sc. in Electrical and Computers Engineering (Major in Telecommunications) from the Instituto Superior de Engenharia do Porto, Portugal. She is currently enrolled in the "Information and Communication Technologies" Ph.D. programme at the University of Vigo, Spain as well as a researcher at INESC TEC, Porto. Her research, which is applied to crowd-sourced tourism data, is focussed on Trust and Reputation, Big Data and Context-aware Recommendation.

HORACIO GONZÁLEZ-VÉLEZ is an associate professor and head of the new Cloud Competency Centre at the National College of Ireland in Dublin. He spent over a decade working in engineering and product marketing for innovation-driven companies such as Silicon Graphics and Sun Microsystems. Award-winning lecturer and researcher, Horacio has also carried out applied research in parallel and distributed computing, funded by a number of public and industrial organisations including the European Commission, UK NESTA, NVidia, Dell, and Microsoft. He holds a Ph.D. in Informatics from the University of Edinburgh.

BENEDITA MALHEIRO holds a Ph.D. and an M.Sc. in electrical and computer engineering from the University of Porto (Faculty of Engineering). She is an adjunct professor at the Polytechnic Institute of Porto (School of Engineering) and a senior researcher at INESC TEC. Her research interests include multi-agent systems, conflict resolution, belief revision, personalisation, recommendation and context-aware systems.

JUAN CARLOS BURGUILLO holds an M.Sc. in telecommunications and a Ph.D. in telematics from the University of Vigo, Spain. He is currently an associate professor at the Department of Telematic Engineering at the same university. He has managed several R&D projects, and has published more than 100 papers in international journals and conferences. His research interests include multi-agent systems, evolutionary algorithms, game theory and telematic services.