CrossMark

ORIGINAL RESEARCH

Predicting happiness: user interactions and sentiment analysis in an online travel forum

Julia Neidhardt[1] · **Nataliia Rümmele[1]** ·
Hannes Werthner[1]

Abstract Web sources of tourism services provide valuable resources of knowledge not only for the travellers but also for the companies. Tourism operators are increasingly aware that user related data should be regarded as an important asset. Furthermore, as data is permanently generated and always available, the landscape of empirical research is changing. In this paper, user activities and interactions in the tourism domain are analysed. In particular, the emotions of the users regarding their forthcoming trips are studied with the objective to characterize interdependencies between them. Social network analysis is applied to examine interactions between the users. To capture their emotions, text mining techniques and sentiment analysis are applied to construct a measure, which is based on free-text comments in a travel forum. The experimental outcome provides some evidence that the network has an effect on the sentiment of the users.

Keywords Social network analysis · Social influence · Network effects · Text mining · Sentiment analysis · Online travel forum

✉ Julia Neidhardt
  julia.neidhardt@ec.tuwien.ac.at

[1] E-Commerce Group, TU Wien, Favoritenstrasse 9-11/188, 1040 Vienna, Austria

1 Introduction

In recent years, the impact of the World Wide Web on almost all areas of modern society has tremendously increased. The tourism landscape has also been profoundly affected by the Web, giving rise to new directions of research in eTourism (Werthner et al. 2015). Together with this development, a number of online communities and their importance have grown. Today, they serve as platforms for people to communicate and to interact—both in people's private lives and in business environments. As a consequence, the amount of available data and user generated content has exploded. Thus, this high quantity of data is a valuable resource for research because it enables to study the behaviour of people as well as their interactions. Furthermore, the huge amount of data has become an important asset of tourism companies. The advantages of properly handling data are manifold: from improving customer relationship management, both in terms of attracting new travellers and maintaining the existing ones, to identifying points for improvement and existing issues in the business. However, new challenges arise: how to manage data and to ensure its quality, how to preserve privacy of the customers, and how to mine valuable knowledge from it. With the development of new computational, mathematical and statistical methods that are able to process and to analyse large amounts of data, there are now a high number of techniques to analyse textual and relational data.

In this paper, user activities and interactions in the tourism domain are analysed. The objective of the study is to determine whether the users are influencing each other. Here, in particular, the emotions of the users are taken into consideration. Thus, the goal is to find out whether these emotions are interdependent. This leads to the following research question: Am I happy because my peers are happy?

To study this question, a travel related online forum is used where users are discussing their forthcoming trips. The main goal of the company owning this forum is to bring users together, thus it is crucial to understand their interactions and influence processes. Also, from the company's perspective, exploring the user generated comments in the forum presents a great opportunity to better understand the needs of their customers, which helps to enrich the user model. A more comprehensive and more accurate view of the users in turn can be utilized for better strategic planning, product design as well as trip recommendations.

For our empirical research we apply social network analysis to characterize the interactions between the users. To capture their emotions, a measure, which is constructed based on free-text comments in the forum, is assigned to the users. Here, text mining techniques and sentiment analysis are applied.

The main results of the study are:

1. A measure to capture the emotions of the user generated text is constructed.
2. A social influence model is built upon the network of users to capture interdependency between user emotions.
3. The experimental outcome provides some evidence that the users are influenced by the sentiments of their peers.

The rest of the paper is structured as follows. The state-of-the-art is discussed in the next section. In Sect. 3 the travel online forum, which provides data for the experimental setup, is described. In Sect. 4 the construction of the network of users of this forum is explained. The calculation of the measure to capture the sentiments in the text is presented in Sect. 5. Statistical models to evaluate the inter-dependency between the calculated measures are described in Sect. 6 together with the discussion of the experimental results. Finally, conclusions are drawn and directions for future work are outlined.

2 State-of-the-art

Social network analysis has a long tradition in the social sciences (Wasserman and Faust 1994). In recent years this method has also become increasingly popular in other disciplines and in particular in computer science. The reason is the tremendous amount of relational data that are available today, coming from the Web and other sources making the design and development for large-scale computational, mathematical and statistical techniques inevitable.

One focus of social network analysis is the study of social influence (Sun and Tang 2011). Social influence occurs when individuals adapt their behaviour according to the behaviour of others in the network. In terms of social network analysis this means that given the edges (i.e., connections) between the nodes (i.e., actors) in the network, the nodal attributes (i.e., behaviour of the actors, or their opinions or sentiments) are influencing one another (i.e., the behaviour is contagious). In our case we want to investigate sentiment changes of a node, which represents a person, in accordance to the emotional status of its network.

It is quite challenging to verify whether social influence mechanisms in fact occur in a network. If the outcome behaviour is binary (e.g., a user is smoker or non-smoker) and longitudinal data exists, SIENA models can be applied (Steglich et al. 2010). If cross-sectional data exists, Autologistic Actor Attribute Models can be used (Daraganova and Robins 2013). The latter can be seen as a generalization of logistic regressions for networks. However, if the outcome behaviour is continuous as the emotion-based measure that we construct in the presented work, Linear Network Autocorrelation Models are appropriate. Those models are related to spatial regression methods. They can be considered as extensions of ordinary least squares (OLS) for networks since they can incorporate local effects (covariates) and interaction effects (network structure) (Leenders 1997, 2002). In literature, they are also called Network Effects Models (Doreian 1989). However, all these models are very complex and/or do not scale.

There is a branch of research that addresses the influence maximization problem in social networks: the goal is to maximize the adoption of a product or the spread of an opinion by identifying appropriate seed users. Typically diffusion models and other computational models are used (e.g., Kempe et al. 2003). However, we focus on statistical inference, which is usually not possible in such models. There is also research that aims at identifying influential users in online discussion forums. Here, typically users with high network centrality measures such as PageRank are

considered as influential. As in our work, forum threads are often used to derive user interaction networks as a basis for the analysis (e.g., Zhang et al. 2007).

The role of emotions of users interacting in online forums and micro-blogging Websites is the focus of several studies. These works illustrate why and how studying user interactions and emotional shifts in online communities can be beneficial for businesses and improvement of user experience. In Mitrović et al. (2010) Blog data is used to demonstrate that user communities emerge around certain topics. The evolution of these communities, i.e., whether they grow or shrink, is related to the emotional content of relevant posts. Posts from Blogs and BBC forums are studied in Chmiel et al. (2011b). This work examines how discussion evolves based on emotional contents, and it shows that the emotions of community members are likely to influence one another. In BBC online forums, where political discussions are taking place, negative emotions are dominating (Chmiel et al. 2011a). Connected users on the Chinese micro-blogging site Weibo show a strong sentiment correlation, especially if they interact a lot. However, negative emotions seem to have a higher impact than positive emotions (Fan et al. 2014). Instead, in the context of MySpace comments positive emotions appear to have a higher impact (Thelwall et al. 2010). It was also observed that there are clear gender differences. Female users express positive emotions more often than male users. In Kramer et al. (2014), the so-called "Facebook Study", experimental evidence for massive-scale contagion of emotional content on Facebook is given. In the study, the messages that are displayed to the users are filtered in a way that some users receive less positive contents and some less negative. It turns out that the users start to behave accordingly in their own messages, i.e., they produce fewer positive and negative contents accordingly.

Since we want to study the contagiousness of emotions in online communities, we need to assess the sentiment of user-generated content. For this purpose supervised machine learning methods are commonly used. However Kramer et al. (2014) applies a lexical-based approach, which is also done in our work. To study correlations and interdependencies between user sentiments various techniques are used, such as temporal approaches including time series and diffusion models (Mitrović et al. 2010; Fan et al. 2014), agent based models (Chmiel et al. 2011a), anova tests (Thelwall et al. 2010), conditional probabilities (Chmiel et al. 2011b), and regression methods (Kramer et al. 2014). We are not aware of any other work, where statistical social network models are applied to relate the sentiments of different users.

We choose lexical-based sentiment analysis to quantify the emotionality of a text or a user since this approach is often applied in the context of tourism. The term sentiment analysis refers to approaches that aim to extract subjectivity from text either to decide whether a text is objective or subjective, or whether a subjective text is positive or negative. The lexicon-based approach utilizes sentiment dictionaries to quantify the subjective of a text by aggregating the sentiments assigned to the words in that text (Taboada et al. 2011). In Gräbner et al. (2012) a lexicon-based approach is applied to relate tourism related reviews to their numerical rating. Using such an approach, the authors are able to classify reviews as "good" or "bad" in a quite accurate way. In Schmunk et al. (2013) statements about product properties of hotel

reviews are extracted. The statements are tested to determine if they are subjective, and if so, whether they are positive or negative. The authors show that for subjectivity recognition the lexical based approach performs better than various supervised machine learning techniques. In Garcia et al. (2012) an approach is introduced that makes use of lexical databases to calculate sentiment scores of tourism related reviews. In Rossetti et al. (2016) Topic-Sentiment Criteria (TSC) Models are presented to extract these aspects, i.e., topics as well as sentiments, from textual reviews. The TSC Models make use of pre-defined word lists containing the sentiment polarity of these words (Lin et al. 2012). In García-Pablos et al. (2016) travel related reviews in several languages are analysed also with respect to their sentiment polarity. Also here a lexical-based approach is applied. Unlike these studies, the goal of this work is to extract sentiments from online travel forum and to identify the inter-dependency between them. Moreover, the suggested approach considers the emoticons as well as negation present in the text.

3 Data sample

The analysis is done within a project with a start-up company. The name of the company cannot be disclosed due to contractual commitments.[1] This company is an online marketplace where group tours to over 200 countries of the world can be compared, booked and discussed. Details about a tour including the points of interests that are visited, the length of the tour, etc. are provided by the respective tour operator. After the tour, a traveller can leave a tour review on the platform. These reviews contain free-text and a five-stars rating for several categories (see Neidhardt et al. 2015).

An important feature of the platform is the discussion within so-called *meets*. In these meets users are given the opportunity to engage online with co-travellers before the tour starts. Typically tour related questions are discussed here. The messages are usually short and are often written in moments when users are excited, i.e., after booking a tour or before the departure. Meets are organized as threads, i.e., sequences of messages that are posted as replies to one another. Every user can start a meet and several meets related to one tour can exist. Meets provide the opportunity to study interactions and possible influence between users, thus they are the focus of the work presented here. From the company's perspective, discussions in the so-called meets present an opportunity to better understand users, their ailments and aspirations, thus leading to insights how to improve user experience and to attract new customers.

The data for the study was received from the company as a dump of their MySQL database. In Fig. 1 we show the entity relationship diagram of the data sample which we used for our analysis. Apart from the user generated free-text within the meets and reviews (see fields 'text' in tables 'meets_comments' and 'reviews' in Fig. 1), the database contains meta-information about meets (table 'meets_comments'), tours (tables 'tours' and 'tour_locations') and users (table 'users') of the

---

[1] In order to ensure reproducibility, the disclosure for interested researchers is possible.
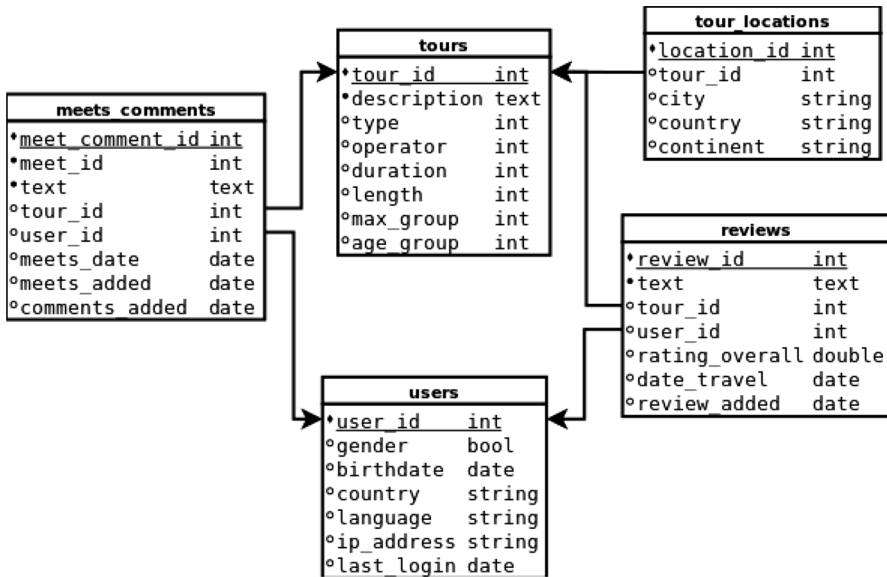
Fig. 1 The entity relationship diagram of the data sample for the analysis

platform. For each meet (see table 'meets_comments' in Fig. 1) it is known when it was started (field 'meets_added'), the comments it includes, when (field 'comments_added') and by whom (field 'user_id') these comments were posted. Here the IP-addresses of the users (field 'ip_address' in table 'users') are also stored. Furthermore, it is known to which tour the meet (field 'tour_id' in table 'meets_-comments') is assigned to as well as the date when the respective tour started (field 'meets_date' in table 'meets_comments'). However, the latter has to be indicated by the users themselves with no restrictions and is, thus, a bit noisy. The available information about tours (see table 'tours' in Fig. 1) encompasses a number of attributes including tour length (field 'length'), destination (see table 'tour_locations' in Fig. 1), tour operator (field 'operator'), maximum possible group size (field 'max_group'), and preferable age of the participants (field 'age_group'). User attributes (see table 'users' in Fig. 1) include gender, location, birth date and language. Except for gender, these details are missing for the majority users. It is also known how active users are overall on the platform.

The study that is presented here, aims to extend the analysis in Neidhardt et al. (2016) to the entire year 2013, i.e., all meets that were posted on the platform from January 1, 2013 to December 31, 2013, are considered. The resulting data sample comprises 32,704 comments posted in 4821 meets by 9881 distinct users. Thus, on average, each meet has 6.8 comments and each user posts 3.3 comments. Furthermore, the 4821 meets are related to 635 tours, i.e., per tour there are on average 7.6 meets taking place. Note that one tour typically has several departure dates, i.e., the same tour is typically offered repeatedly. Out of the 635 tours, 207 (i.e., 32.6%) are taking place in Asia, 161 (i.e., 25.4%) in Europe, 108 (i.e., 17%) in Africa, 79 (i.e., 12.4%) in North America, 58 (i.e., 9.1%) in South America and 22

Table 1 User activity in 2013—comments per month (percentages)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
7.3%	7.9%	8.2%	9.1%	9.6%	9.8%	8.9%	9.6%	7.6%	7.5%	7.9%	6.5%

(i.e., 3.5%) in Australia and Oceania (again, this refers to different types of tours each of which typically has various departure dates). The lengths of the tours vary among one day (or less) and on year. However, on average users participate in tours of 2.8 weeks length and the median is 2.1 weeks.

In Table 1 the distribution of the comments over different months is displayed. From April to August (i.e., during the spring and summer months) the user activity is clearly higher than in the rest of the year. In December the lowest number of comments is posted.

Based on the IP-addresses, the map in Fig. 2 illustrates the geographic location of a user when posting a comment. When comparing the time of these postings to the departure dates of the respective tours, it turns out that almost all comments are posted before the tour starts, only 453 comments out of the 32,704 (i.e., 1.4%) are posted after the beginning of the tour. We make use of this fact to deduce the users' countries of origin, which are not explicitly known, with the help of the IP-addresses, i.e., we use a script to map the IP-addresses to the respective country-codes.

For the vast majority of users (i.e., 9090 or 92.1%) all their comments have the same country-code. For the rest (i.e., 32 or 2.5%), different country-codes are assigned to their comments. Here, the country-code of her/his first comment is assigned to a user leading to the following distribution: 3902 users or 39.5% were from Australia, 2356 or 23.8% from the United Kingdom, 893 or 9.0% from Canada, 657 or 6.5% from the US, 549 or 5.6% from New Zealand, 180 or 1.8% from South Africa and 133 or 1.3% from Germany. There are 93 further countries occurring in the sample but less than 1% of the users were located in each of them so they are not considered further, in particular as we deduce the countries from the users' IP-addresses. Although in general this approach can be considered as sufficiently accurate (What Is My IP Address 2016), we only focus on the bigger countries in our sample in order to draw more reliable conclusions. The resulting distribution of the biggest countries shows that back in 2013 mainly Australians and people from other English speaking countries were using the platform. This clearly makes sense since the company was founded in Australia and only later moved to Europe.

Travellers from Australia mainly book a tour in Europe (2359 travellers or 60.5%), followed by Asia (936 travellers or 24.0%). People from the UK mainly participate in tours taking place in Asia (748 or 31.7%), Australia and Oceania (620 or 26.3%) and Europe (570 or 24.2%). Travellers from Canada clearly favour tours in Asia (512 or 57.3%) followed by Europe (189 or 21.2%). People from the US also prefer tours that take place in these two continents but the distribution is more even i.e., 211 travellers or 32.1% visit Asia, 201 or 30.6% visit Europe. The vast majority

Fig. 2 Locations of users when posting comments (typically before a tour)

of New Zealanders books tours in Europe (348 travellers of 63.4%), followed by Asia (122 travellers or 22.2%). More than 80% of South Africans in our sample travel to Europe (148 travellers or 82.2%) and none of them participates in a tour in their own area. Finally, the travellers from Germany favour tours in Asia (55 travellers or 41.4%), followed by North America (25 travellers or 18.8%) and Africa (24 travellers or 18%). In Table 2 the detailed distributions are displayed.

With respect to gender one can observe that the female/male ratio is almost 3/1: among 9881 users 7195 are female and 2682 are male.

In Table 3 we compare some statistics of the data sample used in this study and the one used in Neidhardt et al. (2016), and we see that overall they are quite similar. However, in April meets comprise on average 3.9 comments whereas this number increases when the entire year is considered. Thus, the users are writing slightly more comments (on average 3.1 rather than 2.4) in a lower number of meets, i.e., more users interact. The median number of comments by a user is two, with a minimum number of one and a maximum of 146.

Table 2 Distribution of continents visited by travellers from most represented countries

	Number of travellers	Africa (%)	Asia (%)	Australia/ Oceania (%)	Europe (%)	North America (%)	South America (%)
Australia	3902	6.2	24.0	2.7	60.5	4.6	2.1
UK	2356	10.5	31.7	26.9	24.2	3.9	3.3
Canada	893	5.7	57.3	4.0	21.2	8.4	3.4
USA	657	13.4	32.1	1.5	30.6	14.0	8.4
New Zealand	549	6.2	22.2	1.3	63.4	5.3	1.6
South Africa	180	8.3	7.8	–	82.2	1.1	0.6
Germany	133	18.0	41.4	3.8	14.3	18.8	3.8

	April 2013	Entire year 2013
Number of comments	3066	32,704
Number of meets	789	4821
Number of distinct users	1270	9881
Comments per meet (avg.)	3.9	6.8
Comments per user (avg.)	2.4	3.1
Users from Australia	38.7%	39.5%
Users from UK	24.5%	23.8%
Users from Canada	11.3%	9.0%
Users from USA	5.5%	6.5%
Users from New Zealand	5.7%	5.6%
Users from South Africa	2.0%	1.8%
Users from Germany	<1%	1.3%
Users from Ireland	1.0%	<1%
Females/Males	74.1%/25.9%	72.8%/27.2%

Table 3 Summary statistics comparing the previous data sample (i.e., April 2013) to the current one (i.e., entire year 2013)

4 User network

To study interactions and influence between the users on the platform, an undirected network is constructed the following way: the vertices of the network represent all users that were writing a comment in a meet in the selected period. Two users are connected by an edge if they were engaged in the same meet. Furthermore, a weight is assigned to the edge, which represents the number of different meets two connected users were part of. This process is illustrated in Fig. 3.

In the resulting network, the 9981 travellers are connected by 33,181 edges. Thus, on average, each user interacts with 6.7 others. This is also the average degree in the network. The highest degree is 139, i.e., there is one user who interacted with 139 others. On the other hand, there are 1853 isolates (18.8%) in the network. These users tried to initiate a conversation but nobody replied. The Kolmogorov–Smirnov test (with $D < 0.05$) shows that the degree distribution only follows a power-law for the degrees 15 and higher (Clauset et al. 2009). However, the first part of the
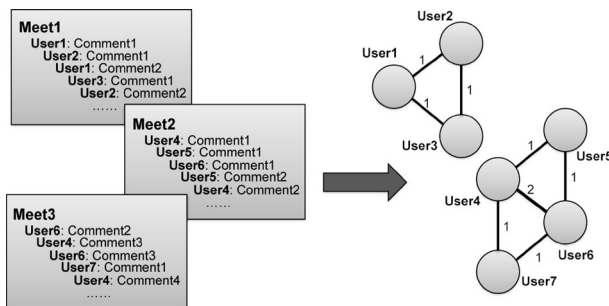

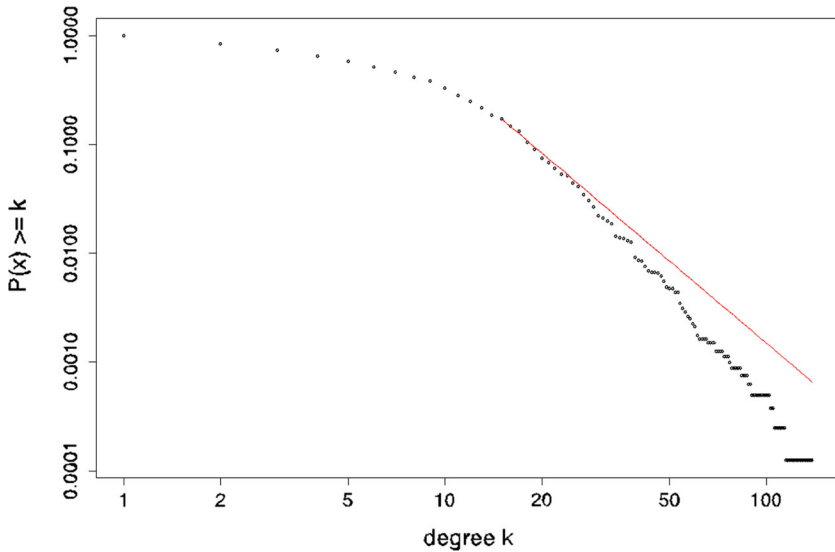**Fig. 3** Construction of a user network based on the meets

Fig. 4 Cumulative degree distribution of the user network (log–log scale)

distribution does not resemble a power-law. This is also illustrated in Fig. 4, where the cumulative degree distribution is shown on a log–log scale. For degrees 15 and higher the distribution can be approximated by a power-law distribution with exponent 3.5 (represented by the red line).

Almost all edges have a weight equal to one (32,821 edges or 98.92%); 340 edges or 1.02% have a weight equal to two, 14 edges or 0.04% have a weight equal to three and only 6 edges or 0.02% have a weight larger than three. This implies that 340 pairs of travellers met in two different meets; and 20 pairs of users even met in three or more different meets.

The network has a high number of small connected components. There are 1069 connected components that consist of at least two nodes, and the largest component has 3357 nodes (i.e., 34% of all nodes). Thus, different regions of the network are hardly connected, but the nodes within a region are densely connected. This can be seen in Fig. 5, where the global structure of the network is displayed. This is not surprising and only reflects the semantics of the constructed network, namely, that each user is typically going only on one tour in a certain period and is, thus, participating only in those meets which are related to that specific tour.

Due to the design of the network, the average clustering coefficient is very high (0.93). The clustering coefficient captures the probability that two randomly selected neighbours of a node (i.e., nodes that share an edge with the node) are also connected by an edge, and thus characterizes the local structure of a network (Newman 2010).

There is no significant difference between the average degree of male (i.e., 6.63) and female users in the network (i.e., 6.74). As statistical tests show, there are also no significant differences between the average degrees of travellers from the most represented countries in the data sample: for travellers from Australia it is 6.53,
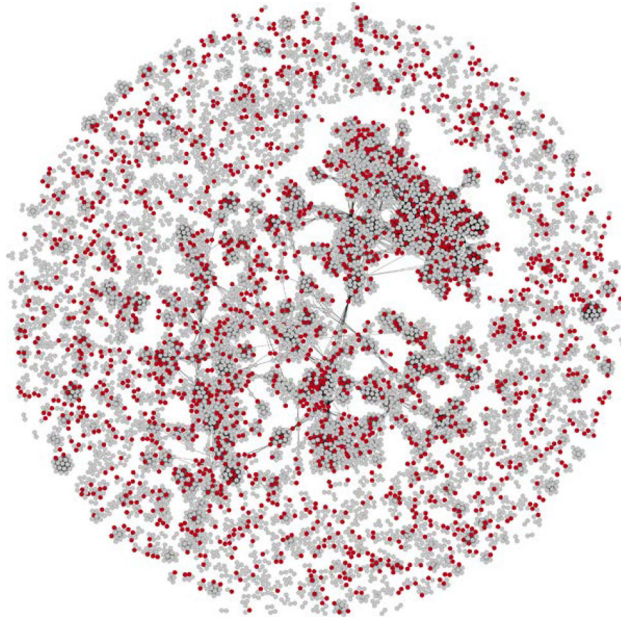
Fig. 5 Global structure of the user network (without isolates)—the *red/darker* nodes represent users with non-positive sentiment (see Sect. 5) (color figure online)

Table 4 Network statistics comparing the user interaction network of the previous data sample (i.e., April 2013) to the current one (i.e., entire year 2013)

	April 2013	Entire year 2013
Number of nodes	1270	9881
Number of edges	2055	33,181
Average degree	3.2	6.7
Maximum degree	3.9	139
Isolates	27.2%	18.8%
Connected components of size ≥ 2	228	1069
Nodes in largest component	4.0%	23.8%
Average clustering coefficient	0.95	0.93

from the UK it is 6.85, from Canada it is 6.52, from the US it is 7.19, from New Zealand it is 7.05, from South Africa it is 7.02 and from Germany it is 6.26.

In Table 4 the network statistics of the presented user interaction network is compared with the one presented in Neidhardt et al. (2016). The average degree increases from 3.9 (April only) to 6.7 (entire year 2013), which corresponds to the observation that the number of comments per meets gets bigger (see Table 3). Furthermore, the number of isolates decreases and the largest component gets

bigger. Thus, over time users, who try to start a conversation, are more likely to receive an answer. However, the number of isolates is still rather high. It is often observed empirically that the size of the largest component increases over time (Easley and Kleinberg 2010).

5 Sentiment scores

Focus of this work is the analysis of the emotions of the users and the interdependencies between those emotions. Thus, a measure, called sentiment score, is constructed with the aim to capture the state of mood of each user. This sentiment score is obtained with the help of a text mining procedure and is based on all free-text comments that a user posted in 2013.

The procedure is as follows. Firstly, tokenization and part-of-speech (POS) tagging of the comments are performed (Bird et al. 2009); afterwards, SentiWordNet (Esuli and Sebastiani 2006; Baccianella et al. 2010) is applied. However, note that in SentiWordNet a word with a specific meaning and POS tag is represented as a synset. Since a word can have different meanings depending on the context, a word can have several synsets, and all of them can have different positive and negative scores. For example, an adjective "poor" has three synsets. All of them have positive score equal to 0, but the first one has a negative score 0, the second one has 0.125, and the last one has 0.5. To resolve this issue, the average of the scores of all synsets is used (Taboada et al. 2011).

Furthermore, the presence of negation in the text is addressed as follows. Once a negation is encountered in the sentence, positive and negative scores for the rest of the tokens in the sentence are swapped (Miller et al. 2011). In this approach, emoticons are also taken into account. A sentiment score of 1.0 is assigned to positive emoticons and -1.0 to negative emoticons. Their values are not swapped after a negation.

For each sentence the sentiment score is calculated as a difference between positive and negative scores per each word and then summed up. Such approach allows to accurately capturing the overall sentiment in the sentence. For example, a sentence with an overall negative sentiment is "Sorry guys I've had to postpone my trip to Africa due to some unforeseen circumstances." whereas "Woo can't wait :)" has an overall positive sentiment score. "How's everyone's packing lists going?", on the other hand is a rather neutral sentence.

Though SentiWordNet does not cover all words used in the comments (either due to misspellings or due to the absence of the corresponding word in the dictionary), we could identify sentiment scores for all meets. Firstly, whenever possible, we have substituted colloquial expressions with synonyms by using a dictionary of spoken English from the natural language toolkit and WordNet. Secondly, this shortcoming is also compensated by the fact that we consider emoticons, which are a good measure for users' mood when writing the comment.

Now, for each user her/his sentiment score is determined as an average of the scores of all sentences in all her/his comments posted in 2013. The sentiment score
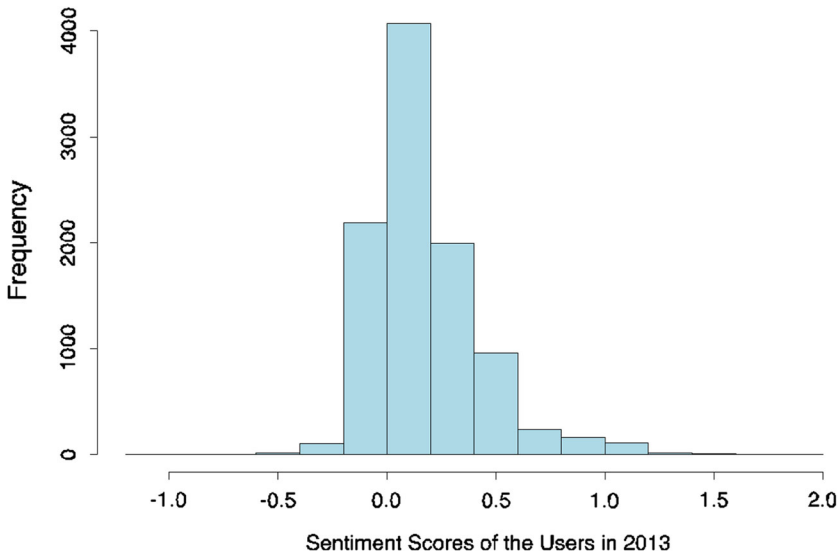
Fig. 6 Distribution of the users' sentiment scores

of user 6 in Fig. 3, e.g., is the average of the sentiment score of the sentences in her/his comment 1, comment 2 and comment 3.

In Fig. 6 the overall distribution of the sentiment scores is displayed. The average sentiment score is 0.17; the minimum sentiment score is -1.17, the maximum 1.93 and the median 0.11. Thus, most of the sentiment scores are positive. This can also be observed in the network displayed in Fig. 5. The positivity of the posted messages might be explained by the fact that future travellers are usually excited about their forthcoming tour.

When considering female and male users separately, it turns out that there is a significant difference between their average sentiment scores (0.19 vs. 0.12, $p < 0.001$).

Regarding the origin of users, the average sentiment scores of users from Australia are significantly higher than those of users not from Australia (0.19 vs. 0.16, $p < 0.001$), the same applies to travellers from the UK (0.19 vs. 0.17, $p < 0.01$) and Germany (0.23 vs. 0.17, $p < 0.05$). On the other hand, the average sentiment scores of users from Canada are significantly lower than the average sentiment scores of users not from that country (0.12 vs. 0.18, $p < 0.001$), and the same applies to travellers from the US (0.10 vs. 0.18, $p < 0.001$). For the other high-represented countries there are no significant differences. In Fig. 7 we summarize these findings.

In Table 5 the summary statistics of the sentiment scores of the users for the entire year 2013 are compared to April 2013. Although the range is bigger in the larger sample, with a minimum of -1.17 and a maximum of 1.93 compared to -0.6 and 1.63 in April, the overall tendency is the same: the average is 0.17 for both and also the medium value only slightly changes, i.e., 0.11 in the entire year 2013 and 0.13 in April. In both cases, moreover, there are significant differences between the
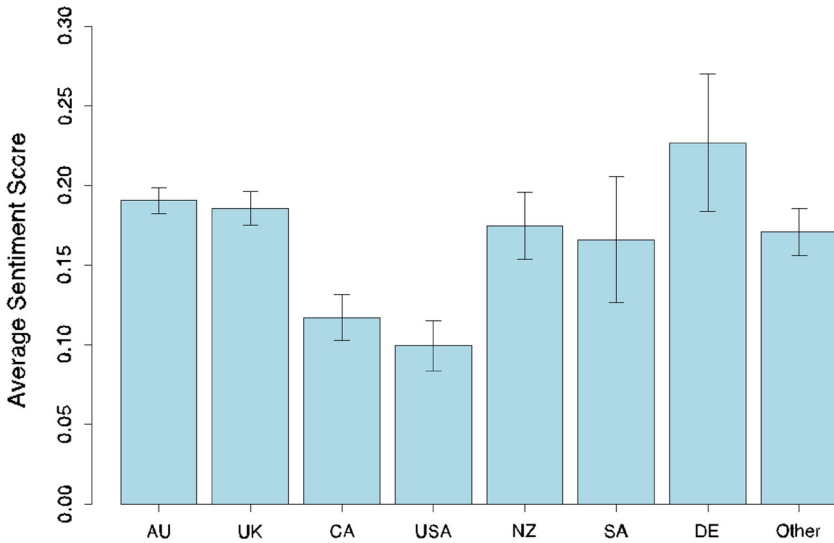
Fig. 7 Average sentiment scores and standard error per country

Table 5 Comparison of the sentiment scores of the users in the previous sample (i.e., April 2013) and in the current sample (i.e., entire year 2013)

	April 2013	Entire year 2013
Average sentiment score	0.17	0.17
Minimum sentiment score	−0.6	−1.17
Maximum sentiment score	1.63	1.93
Median sentiment score	0.13	0.11
Females/males (avg. sentiment score)	0.19/0.11	0.19/0.12
UK/not UK (avg. sentiment score)	0.20/0.16	0.19/0.17
Canada/not Canada (avg. sentiment score)	0.12/0.18	0.12/0.18
USA/not USA (avg. sentiment score)	0.10/0.17	0.10/0.18

average sentiment score of female and male travellers. Regarding the origin of the users, in both samples there are differences between the sentiment scores of users from different countries, the overall tendency is the same. In Table 5 only those countries are listed, where significant differences in both samples occur.

6 Network effect models and results

To test whether the users influence each other regarding their emotions, *Linear Network Autocorrelation Models* are developed. These models are defined by the following equation:

$$y = \rho Wy + X\beta + \varepsilon. \tag{1}$$

Here, the vector y represents the outcome variable, i.e., the sentiment scores of the users in the network. However, y also appears on the right hand side of the equation as predictor variable. This captures the idea that the sentiment score of a user is influenced by the sentiment scores of all users that user is connected to. Thus, these scores are outcome and predictor variable at the same time. The weighted matrix W represents the structure of the network. This implies that only users can influence each other that are connected. The scalar ρ is called *autocorrelation* or *network effects parameter* and represents the strength of the impact of the network on the outcome variable. Thus, the first term in Eq. 1 captures the *contagion effect*. Furthermore, matrix X contains other predictor variables (covariates) and the vector β the corresponding parameters. Thus, the second term in Eq. 1 captures the *intrinsic opinion* of the users. The error term is represented by ε. If there are no network effects, i.e., the first term equals 0, the model is equivalent to Ordinary Least Square Regression (OLS) (Doreian 1989; Leenders 1997, 2002).

For our analysis, we aim to fit a Linear Network Autocorrelation Model based on the data sample, which comprises the whole year 2013. However, to better assess its results we consider the model presented in Neidhardt et al. (2016) as a baseline, i.e., the model for user interactions taking place in April 2013. Comparing the two models helps to gain insights into the stability of the results.

As discussed in Sect. 5, there is a difference in sentiment scores for females and males. Thus, gender is included as a predictor variable. Furthermore, differences with respect of the countries of origin of the users are considered. As we aim to do a comparison with the model presented in Neidhardt et al. (2016), we focus on the countries presented in Table 5. Two dummy variables are constructed: the first indicates whether a user is from the US or Canada and the second whether a user is from the UK. Those dummy variables are included into the model as predictor variables since users from these countries have on average a significant smaller (and respectively larger) sentiment score compared to the other users. The length of a tour (in weeks) and the number of comments written by a user are included as control variables.

In Table 6 the results of the models are displayed. Model 1 is based on the analysis of the user activities in April 2013 and was presented in Neidhardt et al. (2016). It shows that females and users from the UK are more likely to have a higher sentiment score. On the other hand, users from the US and Canada have typically a lower sentiment score. The model also shows that users who plan a longer trip are more likely to have a higher sentiment score. Furthermore, the model detects a positive and significant network effect. It implies that the sentiment scores of the users are not independent but interrelated.

Although we extend the sample considerably, the results are almost the same (see Model 2 in Table 6), which confirms our initial findings. The impact of gender slightly increases. On the other hand the countries of origin considered in the models become slightly less important. The impact of the length of a tour stays the same and the number of comments is significant in Model 2. The network effect slightly decreases but stays strongly significant. Furthermore, it has to be underlined

Table 6 Linear Network Autocorrelation Models

	Model 1: April 2013	Model 2: entire year 2013
Gender female	0.14 (0.01)***	0.15 (0.00)***
User from USA or Canada	−0.04 (0.02)*	−0.03 (0.01)***
User from the UK	0.06 (0.01)***	0.05 (0.01)***
Length of tour in weeks	0.01 (0.00)**	0.01 (0.00)***
Number of comments by user	0.001 (0.00)	0.001 (0.00)***
Network effects	0.02 (0.01)**	0.01 (0.00)***
R^2	0.11	0.09
AIC	−382.3	−255.1
N	1270	9881

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

that the variance explained by Model 2 is almost as high as the variance explained by Model 1 (R^2 is equal to 0.11 in Model 1 and equal to 0.09 in Model 2), which clearly implies that we identified overall patterns rather than random fluctuations. Thus, the representation of the data by these models is reasonable. Furthermore, to test the stability of the results, we developed models, where on one hand, additional predictor variables were taken into account and, on the other hand, without taking the network structure into account, which is equivalent to OLS.

We also conducted a logarithmic transformation of the dependent variable, i.e., the sentiment scores of the users, as these scores are positively skewed (Tabachnick and Fidell 2007). This step did not considerably change the coefficients and their significance but helped to improve the R^2 value. However, we do not report these results in detail, as the focus is the comparison with the previous model.

Overall, the models imply that the more connections a user has, the higher the contribution of the network on her/his sentiment scores. Also, if two users meet in more than one discussion, the impact of this connection gets more important. Thus, your sentiment scores can in fact be predicted by looking at the network connections of a user.

7 Conclusions

The main goal of this work is to determine whether the emotions of a user are influenced by the emotions of her/his peers. Based on the communication threads of the users, a network is constructed. To capture the interdependencies between the sentiments of the users, statistical models for networks are used. The results imply that the emotions of the users are interdependent; a user seems to be influenced by the emotions of all her/his network connections. In particular, the analysis presented here confirms the findings in Neidhardt et al. (2016): although the data sample has been considerably extended, the overall patterns are identical.

In order to understand customer preferences, companies often focus on individual characteristics of their customers. However, typically users are not isolated actors but are rather interrelated. Thus, it clearly makes sense to go beyond the individual level and also take the network level into consideration. This is also shown by our results. Our work implies that in order to understand the preference of a user, also her/his peers as well as influence processes among them should be considered. Combining the individual and the network level helps to establish a more complex user model, which in turn can be used to get a better understanding of the communities a company wants to address. Insights that are gained, moreover, can help to design better products or to recommend specific products to specific groups of user. Furthermore, the user interaction network can be used to identify influential users, e.g., with the help of network centrality measures such as PageRank. This knowledge enables a company to develop strategies for targeting, e.g., non-central user to get them better involved.

Our results clearly imply that the sentiment scores of the users are interrelated. To show this, we utilize models from the literature as explained in the previous section. These models focus on social influence mechanisms. However, part of the detected effect might be due to social selection processes (Sun and Tang 2011); for instance, users who are in a good mood might rather participate in conversations where positive sentiments are already prevalent. What we observe is typically a consequence of all these mechanisms. Thus, further analyses are necessary to distinguish these effects more clearly.

Regarding the choice of the model, Linear Network Autocorrelation Models are appropriate as the outcome variable, i.e., the sentiment score of a user, is continuous. However, these models do not scale well; fitting a model for all the users in the current sample, i.e., 9881 individuals, takes approximately 20 h. Thus, it is very time-consuming to test different models that comprise different combinations of predictor variables. In future work other approaches will be explored. Here, in particular conditional random field models might be an option (Neidhardt 2016).

One assumption of this study is that all users in a thread are interacting with each other, i.e., their interactions are represented by an undirected network. This assumption is reasonable because all users are typically engaged in these discussions shortly before the beginning of a tour. However, in this analysis it is not taken into account how many messages are posted within one thread. In a next step this will be taken into consideration when constructing the weighted network as more interactions might reinforce the influence.

The sentiment scores are extracted and assigned using an automated procedure. Although this approach has its limitations, it is state-of-the-art and well-accepted. However, compared to other studies, positive emotions are prevalent in the presented setting. In BBC online forums where political discussions are taking place, negative emotions are dominating (Chmiel et al. 2011a). This clearly makes sense as people are typically in a good mood and excited when thinking about upcoming vacations. Here, no controversial discussions usually take place. The positive mood seems even reinforced by peers and co-travellers. Thus, the results imply that in the context of tourism positive emotions can be seen as an asset that influences others. However, the same is true for negative emotions. Future work will

further deal with such questions, e.g., if bad mood in a forum can be changed by positive influence. Another issue is how sentiments in discussions before the tour influence the formation of the destination image and affect the overall satisfaction from the travel experience. This would enhance the study of destination branding and image (Költringer and Dickinger 2015). Unfortunately we cannot relate in our sample the comments of the users before a tour to reviews posted on the platform after the tour. This is due to inconsistencies in the data (see Neidhardt et al. 2015).

Our results confirm findings from the literature that there are differences between female and male users with respect to the expression of sentiments (Thelwall et al. 2010). It has also been shown that there are cultural differences regarding the polarity of travel reviews (García-Pablos et al. 2016). However, their reviews in different languages are analysed while we exclusively focus on English text. The differences that we detect with respect to the average sentiment scores of the users from different countries are consistent for the data sample comprising April 2013 and the sample comprising the entire year 2013. However, we plan to conduct a more detailed analyses of the language used by the travellers to obtain a clearer picture why people from Canada and the USA have significantly lower sentiment scores than the rest.

References

Baccianella S, Esuli A, Sebastiani F (2010) SentiWordNet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. LREC 10:2200–2204

Bird S, Klein E, Loper E (2009) Natural language processing with Python. O'Reilly

Chmiel A, Sobkowicz P, Sienkiewicz J, Paltoglou G, Buckley K, Thelwall M, Hołyst JA (2011a) Negative emotions boost user activity at BBC forum. Phys A 390(16):2936–2944

Chmiel A, Sienkiewicz J, Thelwall M, Paltoglou G, Buckley K, Kappas A, Hołyst JA (2011b) Collective emotions online and their influence on community life. PLoS One 6(7):e22207

Clauset A, Shalizi CR, Newman ME (2009) Power-law distributions in empirical data. SIAM Rev 51(4):661–703

Daraganova G, Robins G (2013) Autologistic actor attribute models. In: Lusher D, Koskinen J, Robins G (eds) Exponential random graph models for social networks: theory, methods and applications. Cambridge University Press, Cambridge, pp 102–114

Doreian P (1989) Models of network effects on social actors. Research methods in social network analysis, pp 295–317

Easley D, Kleinberg J (2010) Networks, crowds, and markets: reasoning about a highly connected world. Cambridge University Press, Cambridge

Esuli A, Sebastiani F (2006) Sentiwordnet: a publicly available lexical resource for opinion mining. In: Proceedings of LREC, vol 6, pp 417–422

Fan R, Zhao J, Chen Y, Xu K (2014) Anger is more influential than joy: sentiment correlation in weibo. PLoS One 9(10):e110184

Garcia A, Gaines S, Linaza MT (2012) A Lexicon based sentiment analysis retrieval system for tourism domain. e-Rev Tourism Res (eRTR) 10:35–38

García-Pablos A, Duca AL, Cuadros M, Linaza MT, Marchetti A (2016) Correlating languages and sentiment analysis on the basis of text-based reviews. Information and communication technologies in tourism 2016. Springer International Publishing, New York, pp 565–577

Gräbner D, Zanker M, Fliedl G, Fuchs M (2012) Classification of customer reviews based on sentiment analysis. Information and communication technologies in tourism. Springer, Vienna, pp 460–470

Kempe D, Kleinberg J, Tardos É (2003) Maximizing the spread of influence through a social network. In: Proceedings of the ninth ACM SIGKDD international conference on knowledge discovery and data mining. ACM, pp 137–146

Költringer C, Dickinger A (2015) Analyzing destination branding and image from online sources: a web content mining approach. J Bus Res 68(9)

Kramer AD, Guillory JE, Hancock JT (2014) Experimental evidence of massive-scale emotional contagion through social networks. Proc Natl Acad Sci 111(24):8788–8790

Leenders RTAJ (1997) Longitudinal behavior of network structure and actor attributes: modeling interdependence of contagion and selection. Evol Soc Netw 1

Leenders RTAJ (2002) Modeling social influence through network autocorrelation: constructing the weight matrix. Soc Netw 24(1):21–47

Lin C, He Y, Everson R, Ruger S (2012) Weakly supervised joint sentiment-topic detection from text. IEEE Trans Knowl Data Eng 24(6):1134–1145

Miller M, Sathi C, Wiesenthal D, Leskovec J, Potts C (2011) Sentiment flow through hyperlink networks. In: ICWSM

Mitrović M, Paltoglou G, Tadić B (2010) Networks and emotion-driven user communities at popular blogs. Eur Phys J B 77(4):597–609

Neidhardt J (2016) Modeling and understanding social influence in groups and networks. Dissertation, TU Wien

Neidhardt J, Pobiedina N, Werthner H (2015) What can we learn from review data? e-Rev Tour Res (eRTR) 6

Neidhardt J, Rümmele N, Werthner H (2016) Can we predict your sentiments by listening to your peers? In: Information and communication technologies in tourism 2016. Springer International Publishing, pp 593–603

Newman M (2010) Networks: an introduction. Oxford University Press, Oxford

Rossetti M, Stella F, Zanker M (2016) Analyzing user reviews in tourism with topic models. Inf Technol Tour 16(1):5–21

Schmunk S, Höpken W, Fuchs M, Lexhagen M (2013) Sentiment analysis: extracting decision-relevant knowledge from UGC. In: Information and communication technologies in tourism 2014, pp 253–265

Steglich C, Snijders TA, Pearson M (2010) Dynamic networks and behavior: separating selection from influence. Soc Methodol 40(1):329–393

Sun J, Tang J (2011) A survey of models and algorithms for social influence analysis. In: Social network data analytics. Springer US, pp 177–214

Tabachnick BG, Fidell LS (2007) Using multivariate statistics, 5th. Allyn & Bacon, Needham Height

Taboada M, Brooke J, Tofiloski M, Voll K, Stede M (2011) Lexicon-based methods for sentiment analysis. Comput Ling 37(2):267–307

Thelwall M, Wilkinson D, Uppal S (2010) Data mining emotion in social network communication: gender differences in MySpace. J Am Soc Inf Sci Technol 61(1):190–199

Wasserman S, Faust K (1994) Social network analysis: methods and applications, vol 8. Cambridge University Press, Cambridge

Werthner H, Alzua-Sorzabal A, Cantoni L, Dickinger A, Gretzel U, Jannach D, Neidhardt J, Pröll B, Ricci F, Scaglione M, Stangl B, Stock O, Zanker M (2015) Future research issues in IT and tourism. Inf Technol Tour 15(1):1–15

What Is My IP Address (2016) How accurate is IP GeoLocation? http://whatismyipaddress.com/geolocation-accuracy Accessed Oct 2016

Zhang J, Ackerman MS, Adamic L (2007) Expertise networks in online communities: structure and algorithms. In: Proceedings of the 16th WWW conference. ACM