Information Technology & Tourism (2018) 20:153–180
https://doi.org/10.1007/s40558-018-0128-5

ORIGINAL RESEARCH

Finding patterns in urban tourist behaviour: a social network analysis approach based on TripAdvisor reviews

Egbert Van der Zee[1] · Dario Bertocchi[2]

Received: 19 October 2017 / Revised: 8 October 2018 / Accepted: 21 November 2018 / Published online: 11 December 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract
Developments in ICT and the massive growth in social media usage have increased the availability of data on travel behaviour. This brings an array of new possibilities to improve destination management through Data-driven decisions. This data, however, needs to be analysed and interpreted in order to be beneficial for destination management. Different kinds of methodologies and data have already been applied to analyse spatial behaviour of tourists between and within destinations. The novelty of our paper in this sense that we apply a relational approach by conducting a network analysis methodology on a readily available big data source: user generated content (UGC) from TripAdvisor. The collected data from the city of Antwerp, Belgium shows how locals, Belgians, Europeans and non-Europeans have distinct review patterns, but also shows recurring behavioural patterns. By comparing the relational constellation of the review network to the spatial distribution of central and peripheral attractions, hotels and restaurants, we discuss the added value of social network analysis on UGC for translating (big) data into applicable information and knowledge. The results show a dominant position of a limited number of clustered attractions in the historic city centre, and shows how geographical proximity and relational proximity are interrelated for international reviewers but less for domestic reviewers. This finding is translated into a set of recommendations for policy makers and destination managers trying to accomplish a better distribution of tourists over the entire destination.

Keywords Social network analysis · Tourism destination management · User generated content · Trip advisor · Relational approach

✉ Dario Bertocchi
dario.bertocchi@unive.it

Egbert Van der Zee
E.L.vanderZee@uu.nl

1 Faculty of Geosciences, Utrecht University, Heidelberglaan 2, Utrecht, The Netherlands

2 Department of Economics, Ca' Foscari University of Venice, Cannaregio 873, 30121 Venice, Italy

1 Introduction

Sharing experiences, images and recommendations through social media platforms has become a central element of contemporary tourism. Social media platforms are an important source for travel information because they are prominently available on online search engines (Xiang and Gretzel 2010), but also because they are highly trusted and readily available (Munar and Jacobsen 2014; Zeng and Gerritsen 2014; Gursoy et al. 2017; Law et al. 2010). While tourists seem to find their way to social media platforms and use them not only to create and share content, but especially to inquire, to get inspired and to assist their decision making (Xiang and Gretzel 2010; Hudson and Thal 2013), these platforms are seldom used by DMO's and other stakeholders to acquire information on how tourists perceive destinations, what they do at the destinations and how they subsequently create and communicate tourism destinations images [but see Hays et al. (2013), Kwok and Yu (2013) or Marine-Roig and Clavé (2015) for notable exceptions].

A growing body of literature discusses the mechanisms behind reviewing behaviour and the reputation sites, hotels, attractions and destinations have online, and how this affects tourist behaviour and purchasing decisions [see e.g. Marchiori and Cantoni (2015), Mkono and Tribe (2017) or Miguéns et al. (2008)]. The question how destinations can use this user generated content (UGC) to inform policy and management is however less often addressed (Zeng and Gerritsen 2014; Marine-Roig and Clavé 2015). The fast growing availability of different forms of UGC brings a manifold of innovative possibilities to the field of tourism management. Previous studies have shown how content from photo sharing websites can indicate tourist hot spots in cities (Kádár 2014), the amount of reviews a destination receives can indicate overnight stays (Tilly et al. 2015), reviewing behaviour can be used to divide tourists into segments based on their interests (Hernández et al. 2018) or how the geographic concentration of reviews can indicate tourismification in urban heritage destinations (van der Zee et al. 2018). In these studies, the added value of UGC analyses over traditional data, characterized by simplistic statistics, like number of overnights stays and money spent in the destination on which decision making in tourism often relies, is shown.

Questions are raised concerning the representativeness of UGC as a data source of tourist behaviour, for example not every tourist creates UGC and the ones that do create content do this in a very selective manner (Hernández et al. 2018). Even though these concerns are valid, destinations are starting to explore the possibilities of big data analysis in order to transform information into knowledge about who uses the destination in what way (Buhalis and Amaranggana 2015; Gretzel et al. 2015a, b; Marine-Roig and Clavé 2015). UGC plays an important role in this quest, as it offers readily available, detailed and voluntarily provided information on a manifold of topics important for destinations. A challenge, however, lies in the question how this unstructured and diverse content can be translated from plain information into knowledge about the destination which can be applied by practitioners. Questions remain about how to collect, manage and translate that information into concrete policy recommendations and applicable content for stakeholders and actors of the destination.

Lu and Stepchenkova (2015) distinguish five approaches how UGC can be analysed and used as data for research and policy recommendations. The first approach uses different types of content and sentiment analysis on texts or images, this can be used for example to study how a destination, attraction or service provider is represented by tourists (e.g. Liu et al. 2013). The second type of studies delves into the reasons why and how tourists create and use UGC by applying quantitative studies among tourists or panels [see e.g. Gursoy et al. (2017) or Munar and Jacobsen (2014)]. A third type of studies applies quantitative analysis on numerical data, like ratings, to analyse tourist experiences and satisfaction (see e.g. Lu and Stepchenkova 2015; Banerjee and Chua 2016). The fourth type of studies applies quantitative analysis and modelling to analyse the effect of UGC on (future) tourism behaviour (see e.g. Ye et al. 2011) while the fifth type of studies applies social network or geographic flow analysis on contextual information like geolocations and review patterns of UGC reviews (see e.g. Leung et al. 2012).

In this last line of research, social network analysis is put forward as a methodology to translate large and complex datasets into meaningful and usable information. By depicting a destination as a network of nodes (attractions, hotels, restaurants etc.) which are related through the behaviour of their users, social network analysis can reveal, sometimes hidden, relational patterns present within destinations (Batty 2013; Hernández et al. 2018). Approaching a destination as a web of relations has been done before, for example by studying hyperlink references between stakeholders (Baggio et al. 2010) or patterns of collaborations within the tourism sector (Wyss et al. 2015). Using behavioural patterns of tourists as a measure for identifying relational patterns within a destination is a fairly new and under researched approach [but see Shih (2006) for a notable exception]. Hernández et al. (2018) showed in a recent study how UGC can be used as a valuable data source for showing relational patterns present within destination and were able to use review data to distinguish segments of tourist attractions based on tourist behaviour.

The goal of this paper is to further develop the network approach and social network analysis for exploring how large amounts of UGC can be analysed and translated into applicable recommendations for destination managers and practitioners and inform data-driven decision making. To achieve this goal, we test how social network analysis can help to organise a large dataset of TripAdvisor reviews from an urban heritage destination and translate it into a visualisation of the current destination system through the eyes of its reviewers. By approaching a destination as a network of relations between attractions, hotels and restaurants created by their users, this approach allows to show which features within the destination are connected through tourist behaviour and whether there are certain thematic or geographic clusters present in the destination (Baggio et al. 2010; Bendle 2015; Shih 2006; Hernández et al. 2018). To explore how relational data analysis of user-generated content can facilitate decision-making in a tourism destination, this paper describes the process of data acquisition and analysis of TripAdvisor reviews in the Belgian city of Antwerp, and translates these findings into concrete policy recommendations.

2 Extracting value from big data

2.1 User-generated contents as a tourism big data source

The last decade has witnessed the emergence of a second generation of web–based services, in which users actively provide and share content. Examples of these services are social networking sites, photo sharing platforms, wikis and folksonomies (O'Reilly 2005). In many of these services, tourism or touristic activities have a central position, as the arrival of these platforms allows tourists to share recounts of their personal experiences online (Xiang and Gretzel 2010; Marine-Roig and Clavé 2015). Not only do these platforms provide input for travel plans, destination and hotel reviews, images or suggestions for tourist experiences (Miguéns et al. 2008), it also caused a democratization of travel writing and information provision as it features freely expressed opinions of tourists who visited the destination or took part on the activity in question (Marine-Roig and Clavé 2016). Lu and Stepchenkova (2015, p. 120) state that the emergence of social media and UGC provides "individuals with unprecedented power to instantaneously add 'digital traces' when performing tasks such as reviewing airline, hotel, and restaurant services, lodging a customer complaint, documenting a travel experience, or uploading photos and videos to the global big data bank".

As more and more tourists maks use of these platforms to share experiences and other content, this form of digital peer-to-peer communication, or word-of-mouse (Govers et al. 2007), is becoming one of the most important agents in the formation of tourism destination images (Gursoy et al. 2017; Zeng and Gerritsen 2014). Sharing pictures or experiences on Facebook (Munar and Jacobsen 2014) was found to be the most prevalent platform where tourists share experiences while online review sites like TripAdvisor are believed to be the most influential sites to inform travel choices and tourist behaviour (Gursoy et al. 2017; Kennell and Rushton 2015). Certain web 2.0 platforms, such as TripAdvisor, are becoming increasingly popular and are evolving into primary online travel information sources, sometimes more comprehensive and more specific than destination management organization websites (Xiang and Gretzel 2010).

Platforms featuring tourism related UGC are increasingly being considered as rich data sources for national tourism organizations (NTOs), destination management organizations (DMOs) and other stakeholders (Fuchs et al. 2014; Edwards et al. 2017; Marine-Roig and Clavé 2016). UGC can be used by business in order to assist them in improving the tourism experience they offer by indicating how to personalise and tailor services and products to different types of visitors (Buhalis and Amaranggana 2015; Marchiori and Cantoni 2015), but also by policy makers and destination managers to generate information about tourist behaviour, the functioning of the wider destination system and the perceived image and quality of the different services offered in the destination (Leung et al. 2012; Lu and Stepchenkova 2015). Analysis the large and continuously expanding amount of data from UGC platforms can provide the input necessary for a more data-driven form of policy making as well as doing business. The key lies in translating data into information and implementing this information in such a way that it generates knowledge not available without the application of (big) data analysis.

While a number of studies shows how UGC can be used to generate input for data-driven destination management, this field merits further research. Next to spatial analysis of review patterns (see e.g. van der Zee et al. 2018), a large share of UGC related to tourism provides information about the tourist experience (e.g. a score or review) about an entity that has a geographic location (a restaurant, museum or a destination). This information is often provided by registered users, which adds a profile ID to the shared information. By combining the different reviewed objects based on profile ID's, relational patterns can be uncovered, e.g. profile A wrote a content on museum X, restaurant Y and hotel Z within the same destination. When combining a significant number of these profiles, relational patterns can be found.

Recently, a number of studies applied network analysis and pattern analysis on this type of user generated content to uncover behavioural patterns in cities or regions. Boy and Uitermark (2016) use Instagram to study segregation in Amsterdam and Copenhagen, while Alvin Chua et al. (2016) use tweets to map intraregional travelling by tourists and locals in Italy. Hernández et al. (2018) applied social network analysis on tripadvisor reviews to segment tourists based on their review behaviour and interests. The relational patterns found in these types of data analysis are believed to be able to address policies, enrich experiences, start data-driven developing processes and stimulating value-adding partnerships based on actual behaviour patterns (Fuchs et al. 2014; Shih 2006; Boy and Uitermark 2016). The ambition of the present paper to build upon these studies and investigate the possibility of analysing and extracting value from UGC data to discover and visualize relational patterns that form the networks and tourism systems within a destination created by the actual online behaviour of tourists.

2.2 Relational analysis of UGC as a proxy for spatial behaviour

While empirical research on spatial behaviour of tourists in an urban destination is scarce (Ashworth and Page 2011; Shoval 2018), it is argued to provide important information for destination managers and other stakeholders. Lew and McKercher (2006) present a number of benefits of knowing tourist preferences and actual behaviour, being able to improve and coordinate transportation planning, tackle overcrowding and connect popular attractions to stimulate flows of tourists between them. By identifying underutilized attractions of clusters of tourism product and service providers and connecting them by creating new routes the economic benefits of tourism can be spread over a larger part of the destination, or new attractions or services can be developed among existing popular routes (Shih 2006).

Different kinds of data have been used to study spatial patterns and behaviour analysis, such as GPS tracking (eg. Shoval and Isaacson 2007; McKercher et al. 2012), field surveys (eg. Russo et al. 2010), signals by mobile phones like Wifi or Bluetooth (Versichele et al. 2012), antenna signals of the telecommunication providers (Hawelka et al. 2014) and geo-located social media data such as Twitter and Instagram (Alvin Chua et al. 2016; Boy and Uitermark 2016). While the majority of studies apply flow analysis, spatial modelling or other GIS techniques to map, visualise and predict spatial behaviour, applying a relational perspective and methodology towards the analysis of spatial behaviour by tourists within and between destinations is gaining momentum (Hwang et al. 2006; Bendle 2015; Leung et al. 2012; Lu and Stepchenkova 2015; Shih 2006; Liu et al 2010; Peng et al. 2016; Hernández et al. 2018). Social network analysis of behavioural data, albeit actual spatial tourist footrpints measured by tracking tourists or digital footprints measured by extracting data from different online platforms, give an "insight into the structure and processes of the complex systems which are inherent in tourism contexts" (Bendle 2015, p. 4). In social network analysis of tourism behaviour, most emphasis is given to the relationship between the studied entities. Attractions or services do not exist in isolation, but are connected to other attractions and services. The constellation of relations make up the relational space of a destination in which tourist flows or itineraries are distributed unevenly over space, with some highlights gaining a lot of attention and a large number of attractions, sites or destinations being sparsely connected (Bendle 2015; Leung et al. 2012; Shih 2006; Liu et al. 2010; Peng et al. 2016).

Visualising (review) behaviour as a social network allows to read and interpret the relational space of a destination and increases the interpretability of large and complex datasets (Leung et al. 2012). While spatial patterns of tourist review behaviour can give an insight into the geographic spread of tourism and its costs and benefits and it shows areas and routes where crowding potentially can become an issue, a relational visualisation based on social network analysis can uncover non-spatial clusters of related attractions, sites, restaurants and hotels within a destination. This makes it possible to distinguish itineraries, clusters of activities and over or under-visited segments of the destination, both based on geography and on thematic grounds (e.g. clusters of thematically-related attractions can be distinguished as not geographically clustered).

Furthermore, social network analysis allows for the conducting of statistical tests on the topology of the entire network (e.g. it's density or degree-distribution) or on the position of individual nodes within the network, e.g. showing the importance of nodes through the number of ties they have (degree). These analysis can indicate the stability of the network, can tell whether it is easy or difficult for newcomers to enter the network, can give an indication into how easy information can flow through the network and can indicate which existing connections can be further utilized or where new connections need to be made to improve the network structure (Baggio 2011; Baggio et al. 2010).

Lastly, these relational patterns based on behaviour of tourists can help to inform future tourists through location-based information services. As patterns reveal a majority of visitors of site A are also interested in site B, specific information provisions based on these patterns can suggest tourists to explore the destination based on previous experiences of peers, travel from A to B, or keep it in mind for a future trip to the destination.

Shih (2006) applies social network analysis on self-drive tourists in Taiwan and by looking at different network-related statistics like centrality, betweenness and closeness; and gives an indication of the different types of destinations within networks and suggests the development of specific facilities based on this. Bendle (2015) studied itinerary networks of tour operators in South East Asia to show the dominance of a small number of hubs. Peng et al. (2016) apply SNA to study the effect of a provincial border on tourism flows between destinations in China, while Hwang et al. (2006) were one of the first to study intra-city travelling in the USA applying a network methodology.

While most studies focus on intra-destination movements, and apply social network analysis data collected by conducting tourist surveys [see e.g. Bendle (2015), Peng et al. (2016), Shih (2006) or Hwang et al. (2006)], the studies by Leung et al. (2012) and Hernández et al. (2018) form an interesting exception. In their study, UGC gathered from online travel diaries by overseas visitors of Beijing (Leung et al. 2012) and TripAdvisor reviews in Florida (Hernández et al. 2018) are analysed for spatial patterns of behaviour. The study by Leung et al. (2012) shows overseas tourists are mainly interested in the traditional attractions with a strong reputation, but through their behaviour create a complex web of separate itineraries with a number of clusters of interesting attractions. Hernández et al. (2018) also showed segmentation of tourists is possible through applying social network analysis on online review behaviour, resulting in different clusters of attractions which are reviewed by the same tourists but are not necessarily located close to each other. The geographical and relational space of a destination can thus be different from each other, and while the former is easy to map and interpret, the latter asks for a more complex methodology and interpretation.

Since the study by Leung et al. (2012), the availability and extent of UGC grew exponentially, paving the road for social network analysis of different types of UGC. This paper therefore explores the opportunities the application of SNA of UGC brings for destination management by studying relational patterns of TripAdvisor reviews within an urban destination.

3 Methodology

Review websites such as TripAdvisor allow users to voluntarily leave some traces of their visit/experience from the destination in the form of a review. Contributors need to create a profile to write a review, while all reviews are available for the wider community without registration. In the example of TripAdvisor, every review combines a description of an experience (qualitative information), a score representing the value ascribed to the experience (quantitative information), a location of a reviewed attraction, restaurant or hotel (geographic information), a time of visit and/or time of review (temporal information) and a profile (personal information). TripAdvisor lists and shares approximately 385 million reviews and opinions and has on average 350 million unique visitors every month (Tripadvisor 2016). When collected, this provides a large dataset which can be analysed for patterns that reflect tourist behaviour.

In this paper, we study intra-destination review patterns by collecting TripAdvisor reviews on attractions, restaurants and hotels within one destination. The destination chosen for this study is the Belgian city of Antwerp. The city of Antwerp is a relatively popular tourism destination, receiving 1.1 million tourists who spend a total of 1.9 million nights in the city in 2017 (Toerisme Vlaanderen 2018). On average, tourists spend 1.75 nights in the destination, which is the lowest length-of-stay among the Flemish Art Cities (Bruges, Ghent, Antwerp, Leuven and Mechelen). The DMO, Visit Antwerpen, aims at increasing the length of stay in the destination and to increase the visitation of areas, attractions and services away from the city centre (van der Zee et al. 2018). At the time of writing, TripAdvisor listed over 90,000 reviews on Antwerp, featuring over 1200 restaurants, over 250 things to do (referred to as attractions in this study) and over 200 accommodation facilities (referred to as hotels in this study).

3.1 Data acquisition

This study applies social network analysis on UGC, in this case reviews from TripAdvisor in a single urban destination. We look specifically at relations between reviewed places. These relations are formed when a reviewer reviews multiple places, and thus connects places through his or her reviewing behaviour. To be able to conduct a relational analysis on this type of UGC, it is vital to collect a sample consisting of reviewers and take into account all reviews written by these reviewers on places in the chosen destinations. To collect a large sample of the 90,000 reviews available on Antwerp, a scraping technique was applied. We used a web-based scraping software (Kimono) and inserted a list of URLs of reviewer profiles as input. The scraping software was then able to create two separate databases, one database with information about the reviewer provided on the profile page (reviewer ID, country of residence, age-group, gender, total number of reviews and the date the reviewer joined TripAdvisor) and one database with all reviews by all reviewers (with reviewer ID, name of reviewed site, location of reviewed site, rating and date of review).

In order to create the list of URLs needed as input for the web scraper, we collected all reviewers who wrote about the Antwerp train station. The Antwerp train station was at the time of writing both the most reviewed place in the destination and was ranked as most popular attraction in the city on the TripAdvisor website. Besides, it is also one of the most used gateway to reach the city and it is an important place for tourists since it hosts a tourist info-point, maps of the city and luggage lockers.

From this point of departure, the collecting procedure followed a number of steps: gathering the URLs of all 4448 users that wrote a review about the central station; collecting profile information from these reviewers; collecting all the reviews the users wrote on TripAdvisor; filtering the review dataset for reviews on places in Antwerp to reduce the size of the database. From the total of 4448 users, 4354 profiles were kept after removing duplicate profiles or scraping errors. These 4354 users wrote a total of 352,790 reviews (an average of 81 reviews per user, of which 20,948 considered sites in Antwerp (an average of 4.81 reviews per user), covering 21.7% of all reviews in Antwerp.

The outputs of the scraping and selecting procedures are two different data sets, one covering the profile data, composed by profile ID of the user, total number of reviews wrote, city of residence, country of residence and country code [1 = Antwerp (Local), 2 = Belgium (National), 3 = Europe, 4 = Rest of the World, 9 = Unknown]. This led to the following division of users and reviews:

- European 1311 users, 5614 reviews (average 4.28 reviews per user).
- Non-European 1029 users 3912 reviews (average 3.80 reviews per user).
- Local 569 users 4967 reviews (average 8,73 reviews per user).
- Belgian 398 users 2214 reviews (average 5,56 reviews per user).
- Unknown 1047 users 4241 reviews (average 4,05 reviews per user).

The second dataset hat consists of the user ID (profile nickname), reviewed place, city of the reviewed place, score of the review (from 1 to 5) and date of the review. This second data set was queried with the help of the profile dataset to create five different matrixes to conduct social network analysis for different groups of reviewers, in which each of the previously distinguished groups based on place of residence were gathered in a separate matrix. The Tableau software package has been used to query the database and create the matrixes used for the data analysis is.

3.2 Data analysis

Social network analysis was applied to uncover relational patterns within the destination created by review behaviour. The matrixes created from the UGC combine user-profiles with reviewed site in the destination. Since the rows and columns in this matrix are of a different nature (a two-mode network of users and reviewed places), the matrix needs to be converted into a one-mode network (showing the relation between places). This goes by the assumption that when a reviewer reviews two different places on TripAdvisor, the reviewer visited both of them during his/her stay in Antwerp. Therefore, the one-mode network takes the form of a matrix listing places in both the rows and columns, and in the cells the number of times the places are combined by users is listed. This creates a weighted one-mode network. The weighted network matrices are then converted into binary, non-weighed network matrices (in which '1' refers to a connection between nodes and '0' refers to no connection), which are used to calculate network measures and map network graphs. The former allows to better understand the structure and density of the network while the while the latter give an insight in the aggregate tourist behaviour creating the networks. While the calculations and visual representation are made using a binary, non-weighted network, the weighted relations are used to allow filtering of sparsely occurring ties and nodes to reduce the complexity and improve the legibility of the visual representation. We use the non-weighted network for the social network analysis since the goal is to show which places are similar do to the other places they are connected to, and use the weighted network to show the most occurring ties in the destination. The constructed matrix makes it possible to analyse the structure and topology of the network (Baggio et al. 2010). In this paper the following quantitative network measures are analysed (following e.g. Baggio et al. 2010; Bendle 2015):

- The density of the total network (the percentage of ties in the network compared to the maximum number of ties if all nodes were connected) to show how heavily Antwerp as destination is used (tourist pressure).
- The topology of the network (the degree distribution of nodes in the network, which can be either a normal distribution where nodes are connected relatively evenly to each other or can be a skewed power-law distribution with a small number of dominant nodes with a very high degree and a large number of nodes with a very small degree); which gives insight into the hierarchy of reviewed places in Antwerp.
- The structure of the network (relative relations of nodes with each other and the total network pattern).
- The centrality of the different nodes within the network (the connections a node has with other nodes in the network, also known as 'degree'); indicates the difference in connectivity into the wider destination system and thus the popularity of the various locations.
- The strength of ties in the network (the number of times a tie occurs between two nodes); to highlight potential and important tourist flows in Antwerp.

Next to a more quantitative approach, a visualisation and more qualitative analysis of patterns and attributes of nodes and ties can help to construct the narrative of the tourism destination as created by its users. For visualisation purposes, a threshold value for the minimum value a tie between two nodes needs to have in order to be included in the network graph is chosen. Since in all four networks a different number of reviewers is present, reviewing a different total number of nodes and creating a different number of ties, a relative threshold value was chosen. By looking at the effect of taking different threshold values (Fig. 1), for every network the first value was chosen where the network stabilized. I.e., where the steep decline of losing ties when the minimum number of ties was expanded by 1 stopped and the relative change compared to the previous step was comparable. For the European and non-European reviewer networks this point was a minimum of five ties, for the local reviewer network the minimum was six ties and for the Belgian network the minimum was four ties. The network graphs were enriched by colouring the nodes consistent with the type of place they represent (e.g. Restaurants, hotels, museums etc.), allowing for a better visual interpretation. The social network analysis and visualisation was conduct using UCINET (Borgatti et al. 2002).

While the social network analysis allows to better understand the relation system of the destination and the interrelatedness of different places within the destination, it does leave some questions as to why the distinguished patterns occur. To further explore this question, the last step in this paper was to visualise the networks applying a geographical methodology. To explore whether there were geographic patterns to be distinguished in the review networks, all reviewed places were mapped displaying their 'degree'. The degree of a place, being the

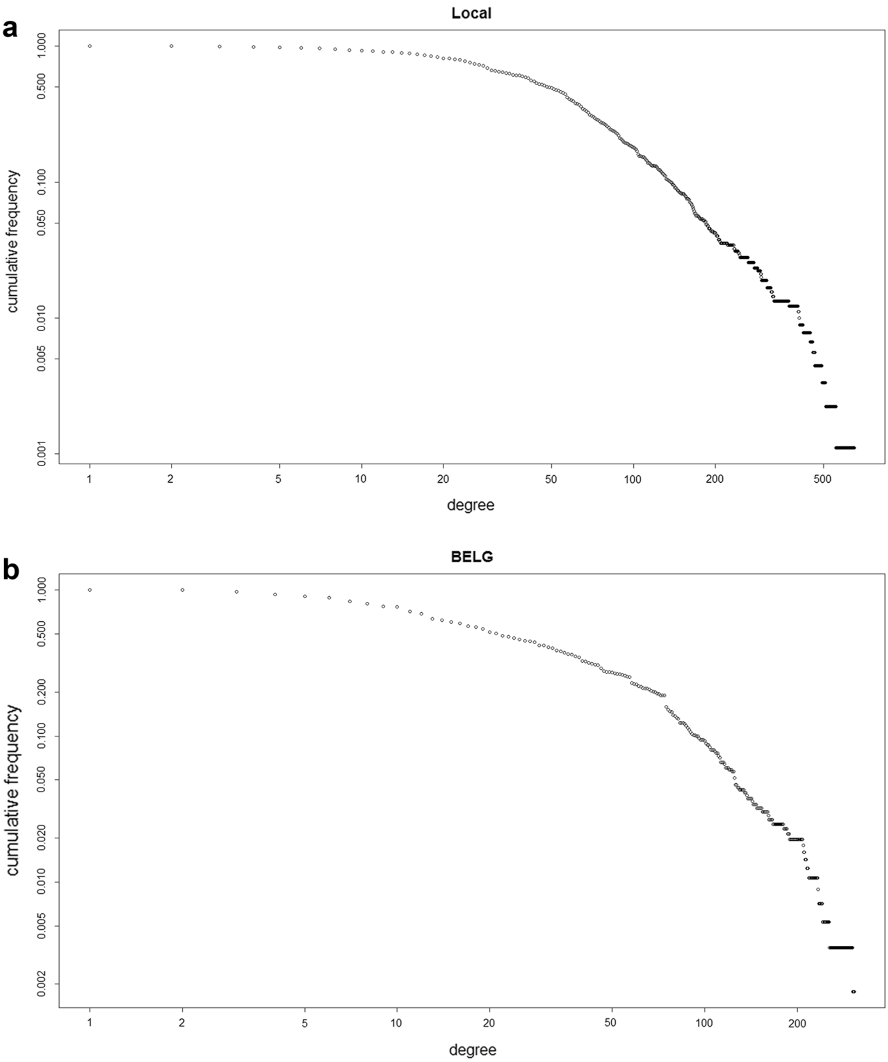

Fig. 1 **a–d** Degree distribution of reviewed places (logarithmic scale, degree on the x-axis and relative frequency of nodes on the y-axis, where the combination 1.000(y) – 1(x) means 100% of the nodes have a degree of 1)

number of other places the place in question is connected to through review behaviour of its visitors, shows its relative position in the network. Mapping the places according to their degree allows to explore whether places which are well connected in the destination network are also located in close proximity. To do so, all places were geo-localized and visualised by displaying the XY coordinates on a map of the destination by their degree (size) and type of place (colour) using the Tableau software package.

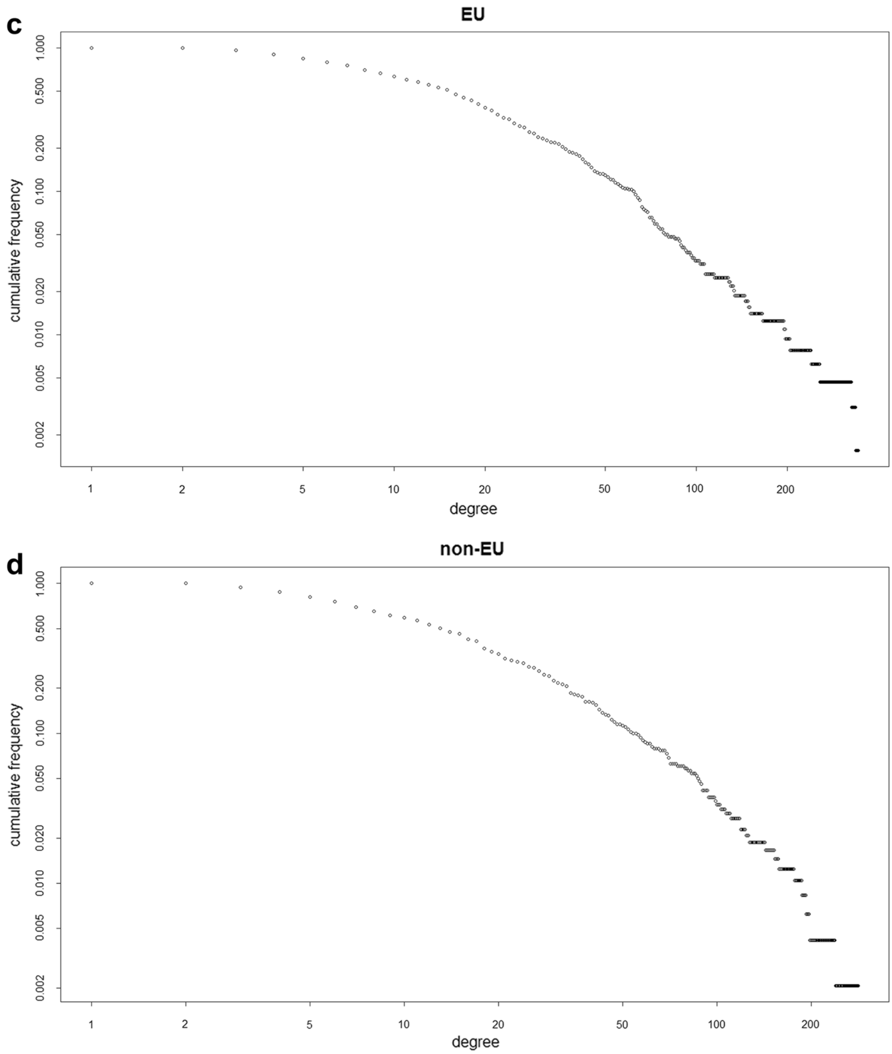

Fig. 1 (continued)

4 Results

The different networks have a similar topology, with some distinctive features. The degree distribution over the nodes in the network is a first insight into the topology of the network. The figures show a relative frequency (y-axis) distribution of centrality (x-axis). The topology of the reviewer networks was found to show features of a Power-law distribution (a skewed distribution with a couple of highly connected nodes and a majority of sparsely connected nodes), which suggests a process of preferential attachment as expected given the data: if locations are often reviewed and therefore more central in the review network, they are likely to attract consecutive visits and reviews leading to the skewed 'fat tail' distributions illustrated in Fig. 1a–d. Three different clusters of nodes can be distinguished, being a cluster of a limited number of nodes having a very high degree (core-cluster), a cluster of a larger group of nodes with a moderate to relatively high degree (ring of connected nodes) and a cluster with the majority of the nodes which are only sparsely connected to other nodes (periphery).

The topology of the network and distribution of degree indicate that new places that would enter the network are most likely to be connected to a number of dominant nodes, raising the degree of these nodes. In other words, reviewers who review one place in Antwerp, are likely to review several other places, with a bias towards the local hotspots that already attract a lot of reviews. These hotspots are mainly located in the historic city centre (e.g. the Cathedral, Grote Markt and Rubens House) with an exception for the MAS and Red Star Line museum at the northern fringe of the city centre (see Table 1). Even though the degrees differ between the groups of reviewers, the most connected places are even stronger connected in the review network by local reviewers compared to the other groups, the lists show lot of similarities. The destination's main museums and heritage attractions are among the best connected in all review networks, and almost all listed places have an explicit reference to the cities mediaeval, baroque and more recent heritage.

Looking more closely at the topology of the network, comparable network densities can be found. The lowest density can be found in the network of European reviewers (0.077), while the highest density can be found in the local network (0.106) (Table 3), suggesting that the different reviewer groups use the city as a tourist destination in a similar way. The high density (in practice a degree between 10^{-1} and 10^{-2} is believed to be high) indicates the review behaviour connects tourist attractions, restaurants and hotels in the city of Antwerp. As the distribution of degrees illustrates, the centrality shows hierarchy as it is not divided evenly over the nodes: only few places have a high centrality, most have a very low degree centrality (see Table 1 for the places with the highest degrees).

All networks (Fig. 2) show a similar network structure. The point of departure, the central station, is connected to all other nodes. The centre of the network graphs also shows an interconnected cluster of nodes consisting mainly of museums, churches, heritage sites and important squares and streets. The configuration of these clusters, as well as the density and the relation with the central hub is different in the different networks. The periphery of the network, which is mainly made up by restaurants for the local and Belgian reviewers and restaurants and hotels for the international reviewers, shows a similar structure in all four networks. The number of nodes is different, but the pattern of a large periphery which is only marginally connected to the central cluster is recurrent in all network visualisations. This means that most restaurants and hotels are reviewed in relative isolation, because reviewers either only review a single or limited number of these facilities, or because the patterns of combined reviews of facilities is not recurrent and combinations between facilities or between facilities and attractions are relatively random.

When combining the visualised network structures, a number of differences between the networks can be uncovered. The main clusters in all four networks are

Table 1 Places with the highest degree in the different review networks

Local reviewers	Degree	Belgian reviewers	Degree	European reviewers	Degree	Non-European reviewers	Degree
MAS Museum	0.73	MAS Museum	0.55	Cathedral	0.54	Cathedral	0.60
Zoo	0.62	Zoo	0.54	MAS Museum	0.53	Grote Markt	0.51
Grote Markt	0.57	Meir	0.46	Grote Markt	0.51	Rubens House	0.42
Cathedral	0.55	Grote Markt	0.43	Rubens House	0.40	MAS Museum	0.41
Sint Annatunnel	0.52	Sint Annatunnel	0.42	Meir	0.38	Town Hall	0.395
Da Giovanni Restaurant	0.51	Cathedral	0.42	Town Hall	0.32	Meir	0.37
Meir	0.50	Middelheimpark	0.39	Zoo	0.31	Plantin Moretus Museum	0.33
Red Star Line Museum	0.47	Cogels Osylei	0.38	Plantin Moretus Museum	0.31	Port	0.32
Rubens House	0.46	Rubens House	0.38	Sint Annatunnel	0.26	Zoo	0.30
Het Pomphuis Restaurant	0.45	Grand Café Horta	0.38	Steen Castle	0.24	Sint Annatunnel	0.27
Cogels Osylei	0.45	Groenplaats	0.37	Red Star Line Museum	0.23	Steen Castle	0.26
Shopping Stadsfeestzaal	0.42	Town Hall	0.34	Da Giovanni Restaurant	0.23	Da Giovanni Restaurant	0.25
Groenplaats	0.37	Red Star Line Museum	0.33	Diamond District	0.21	Carolus Borromeus Church	0.25
Kloosterstraat	0.36	Vlaeykensgang	0.32	Groenplaats	0.21	Hilton Antwerp Old Town	0.23

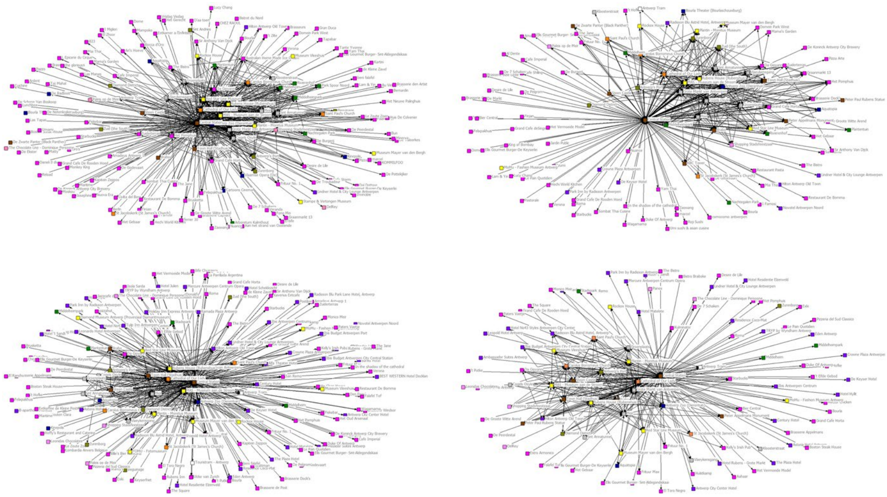

Fig. 2 Patterns of user behaviour

different, both in structure, in position related to the main hub and by the nodes which are part of the cluster. The main differences are summarised in Table 2. While all networks show a core with a central cluster, a moderately connected ring and sparsely connected periphery, some important distinctions can be found, in the structure, narrative and geography of the networks. Looking at the most occurring connections (Table 3), some differences are visible between the different groups. While for the local and Belgian reviewers a number of museums, heritage attractions and the zoo form the core of the most occurring combinations, the Grote Markt (market square) and the adjacent Cathedral are clearly the main hubs for the non-European reviewers.

Both the clustered relational pattern of review behaviour as the list of most occurring connections and most central places in the destination suggest a strong core of different types of heritage attractions forms the touristic DNA of the destination. A selection of museums, churches, squares, streets and monumental buildings are reviewed most often and connected to the majority of other reviewed places in the destination. Figure 3a–d show that apart from a relational cluster, the places can also be seen as a geographical cluster. Mapping the distribution of reviewed places according to their degree shows whether there is a relationship between the geographical location of places and their level of centrality in the review network. The places that form the core of the review network can also be found located in the heart of the historic centre of the destination, with the exception of the MAS and Red Star Line museum which are located to the North of historic city centre, and the central station and zoo which are located to the east of the city centre. Reviewed restaurants and hotels are distributed more evenly over the destination, even though a cluster of restaurants can be found in the direct vicinity of the destinations' core attractions. Comparing the distribution of places with a high degree between the different review groups provides

Table 2 Most occurring ties in the different review networks

Local reviewers		% of total possible occurrences	Belgian reviewers		% of total possible occurrences
Cathedral	Zoo	8.96	MAS Museum	Zoo	6.53
Grote Markt	Cathedral	8.96	Grote Markt	Zoo	5.03
Grote Markt	Zoo	8.79	MAS Museum	Cathedral	5.03
MAS Museum	Cathedral	8.61	Rubens House	MAS Museum	5.03
MAS Museum	Grote Markt	7.73	MAS Museum	Grote Markt	4.77
Rubens House	Cathedral	7.21	Rubens House	Zoo	4.27
Red Star Line Museum	MAS Museum	6.85	Rubens House	Cathedral	4.27
Rubens House	Zoo	6.33	Cathedral	Zoo	3.77
Rubens House	MAS Museum	6.15	Grote Markt	Cathedral	3.77
Sint Anna tunnel	Zoo	5.62	Meir	Zoo	3.77
Cogels Osylei	Zoo	5.27	Groenplaats	Zoo	3.52
MAS Museum	Cogels Osylei	5.27	MAS Museum	Cogels Osylei	3.52
Rubens House	Grote Markt	5.27	MAS Museum	Red Star Line Museum	3.27
Red Star Line Museum	Cathedral	5.10	Rubens House	Grote Markt	3.27
Red Star Line Museum	Zoo	4.92	Sint Annatunnel	MAS Museum	3.27
European reviewers		**% of total possible occurrences**	**Non-European reviewers**		**% of total possible occurrences**
Grote Markt	Cathedral	6.86	Grote Markt	Cathedral	8.61
Rubens House	Cathedral	4.96	Rubens House	Cathedral	6.87
MAS Museum	Grote Markt	4.58	Town Hall	Cathedral	4.40
Rubens House	Grote Markt	4.50	MAS Museum	Cathedral	4.03
MAS Museum	Cathedral	4.42	Rubens House	Grote Markt	4.03
Rubens House	MAS Museum	3.66	Steen Castle	Cathedral	3.66

Table 2 (continued)

European reviewers		% of total possible occurrences	Non-European reviewers		% of total possible occurrences
Town Hall	Cathedral	3.28	Meir	Cathedral	3.30
Meir	Cathedral	3.20	Town Hall	Grote Markt	3.11
Meir	Grote Markt	3.13	Meir	Grote Markt	2.66
MAS Museum	Zoo	2.67	Cathedral	AntwerpsPort	2.56
Cathedral	Zoo	2.52	MAS Museum	Grote Markt	2.56
Steen Castle	Grote Markt	2.52	Steen Castle	Grote Markt	2.47
Town Hall	Grote Markt	2.52	Cathedral	Carolus Borromeus Church	2.38
Rubens House	Meir	2.44	Plantin Moretus Museum	Cathedral	2.38
Grote Markt	Zoo	2.36	Town Hall	Meir	2.11

Table 3 Clusters of the four networks

	Relational structure of the network	Narrative of the network	Geography of the network
Non-European reviewers Nodes: 499 Ties: 11,734 Centrality: 0.092	Loosely connected central cluster, moderately close to the central hub surrounded by a ring of nodes connected to different sections of the cluster. The periphery consists of a low number of nodes with single connections to the main hub. Relatively few nodes but a higher centrality compared to other (international) tourists	Strong cluster of different types of sites and attractions representing historic/medieval Antwerp. The tourist experience is extended by some by different museum visitations, churches, the diamond district or the art nouveau district. The periphery is made up by various hotels and restaurants with a dominant focus on a ‘beer and fries’ type of tourism product	Limited geographic distribution of well-connected places. Clusters can be found surrounding the central station, historic city centre and to a lesser degree the MAS museum. Secondary tourist products with a high degree are located in the proximity of the main core of attractions

Table 3 (continued)

	Relational structure of the network	Narrative of the network	Geography of the network
European reviewers Nodes: 658 Ties: 17,122 Centrality: 0.077	Small dense main cluster close to the main hub, surrounded by an interconnected ring of nodes at some distance from the main cluster and a second tier of mainly hotels and various tourist attractions connected to different sections of the cluster. The periphery consists of a relatively large number of nodes, mainly hotels and restaurants, with connections to the main hub at comparable distances	The network shows multiple narratives. The main clusters consist of strong ties between the main historic centre sites/attractions (Cathedral, Town Hall, Grote Markt and Rubens House). The MAS museum is very central in the network but not fully part of the historic cluster. It forms a bridge to other museum in the city. The visit to the central cluster is extended by visits of other museums, streets, squares and districts with a historical (art nouveau, medieval) or functional (shopping, gastronomy) purpose. A number of hotels has a central position in the network, indicating the importance of multiple-day city trips. The pattern of restaurants shows also the presence of 'beer and fries', but also has a number of other type of accessible and quick options (lunchrooms, sandwich shops, burger shops) as well as Italian restaurants, Belgian brasseries and a limited number of exclusive restaurants	This network has the highest degree of geographic clustering of well-connected nodes. Very dominant position of the historic city centre, with incidental connections to peripheral attractions and hotels. Well-connected restaurants are exclusively located in the direct proximity of the main tourist attractions

Table 3 (continued)

	Relational structure of the network	Narrative of the network	Geography of the network
Belgian Nodes: 582 Ties: 22,664 Centrality: 0.089	Moderately dense central cluster, relatively far away from the main hub surrounded by a ring of nodes connected to different sections of the cluster. Ties of the nodes in the cluster with the central hub are very strong, but intra-cluster ties are recurrent but less strong. In this sense the Belgian reviewer network is more to the international tourist networks than the local network. The periphery consists of a relatively modest number of nodes with single connections to the main hub	The main cluster consists of a small number of highly popular museums/attractions (MAS museum, Rubens House, Zoo, Cathedral) surrounded by open, public space with a historical (art nouveau, medieval) or functional (shopping, gastronomy) use. A number of restaurants is connected to the main cluster, consisting of famous but not necessarily centrally located restaurants. These restaurants have in common that they offer a dining experience in a unique setting due to architecture and design. The other, less connected restaurants have distinct cuisines, styles price levels and locations, but have a high rating on TripAdvisor in common	Secondary tourist products surrounding the central station are less well connected, while the historic city centre and direct surroundings has a dominant position. Secondary tourist products in the southern and northern fringe of the city centre boast high degrees and are better connected compared to the international reviewer networks

Table 3 (continued)

	Relational structure of the network	Narrative of the network	Geography of the network
Local Nodes: 908 Ties: 60,680 Centrality: 0.106	Dense and extensive central cluster, strongly connected to the main hub. The periphery consists of a large number of nodes, with different levels of connectivity to the central hub	The museums that portray the different eras and narratives of the history of Antwerp (Rubens House – Baroque, MAS Museum – wide array of expositions and Red Star Line museum – naval history) together with the Cathedral, Zoo and Grote Markt, as well as a selection of streets and squares with a (historic) significance form the backbone of the local narrative of Antwerp as a tourist destination. A cluster of parks can be found at the fringe of the core cluster, as well as other types of public space and churches. The core is surrounded by the different neighbourhoods and districts, which are not connected to each other, but are connected to the core. The pattern of 'beer and fries' is absent in the local network, which better represents the diverse and cosmopolitan gastronomic offer of the city	As in the Belgian review network, secondary tourist products surrounding the central station are not strongly connected. There is a dominant position of the extended historic centre. Secondary tourist products in neighbourhoods in South and North of the city, as some peripheral attractions (mainly parks) show a relatively high degree, indicating proximity plays less of a role compared to the other networks

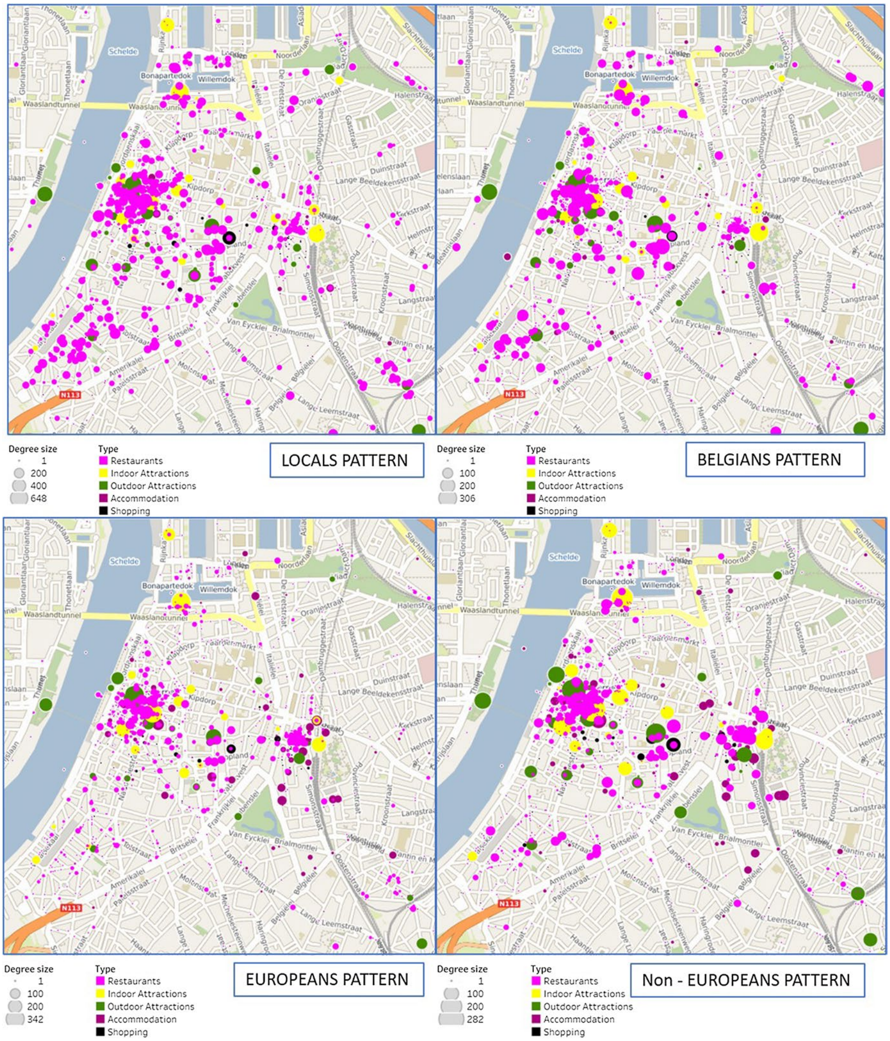

Fig. 3 Maps showing the geographic location and degree of the reviewed places in the different review networks

an interesting finding. Where in general restaurants in de vicinity of the destinations core attractions are also better connected in the review network (in other words, receive a high 'degree'), the maps of the local and Belgian reviewers show that also less proximate restaurants receive high degrees. A signification number of restaurants in the South, South-east and North of the destination are relationally well connected but geographically less proximate to the destination's core attractions. This pattern is however not found for the other groups, for European and non-European reviewers a geographical marginal location often also means a relationally marginal position in the network. In other words, restaurant visitation outside the direct proximity of the destination's core attractions does happen, seems to be happening more incidentally by international tourists.

Even though for all groups the historic city centre is the most important location of the main cluster of reviewed places, the configuration of the main clusters are different. The non-European tourists show the strongest stereotypical review behaviour (both in narrative as in geographical distribution), while the local reviewers highlight the different representations of the history of the city of Antwerp, as well as open and public spaces and a wide representation of the gastronomic offer of the cosmopolitan city. However, for the majority of the reviews by all different groups can de stated they are clustered in a very limited geographical space. This space, the historical city centre, is a place where different types of tourists, excursionists and local inhabitant physically come together.

5 Discussion and conclusion

Applying a relational approach towards the analysis of tourist behaviour, in this case a social network analysis of TripAdvisor reviews proofs to be valuable for two reasons. Both the geographic location (in space) and the relative location (within the review network) can be mapped and analysed relative to other places. Clustering occurs both through review behaviour as through spatial behaviour by tourists, as often reviewed places in the review networks of international tourists are also in proximate location of each other. By regarding a destination as a network, both actual and potential flows can be uncovered, which are different for different groups of reviewers.

Attractions and sites depicting medieval Antwerp and a gastronomic interest in restaurants and bars offering the stereotypical Belgian 'beer and fries', spatially clustered around the central station and historic city centre are dominantly present in the non-European review network. This pattern reflects the tourist bubble thesis coined by Cohen (1972) and found present in the work by Lew and McKercher (2006) and the social network analysis of UGC by Leung et al. (2012). For international reviewers, geographic distance and relational distance seemed to be more similar compared to domestic reviewers of the destination. International reviewers in general were found to mainly stick to the area surrounding the main tourist hubs, connecting mainly places within or between these zones. The distribution of degree among the nodes in the network indicates that an extension of the tourism product offering by the entering of new restaurants, hotels or attractions would even strengthen the existing pattern. The opening of the popular MAS museum in 2011 on the Northern fringe of the city centre can be seen as an example of this claim. Even though this attraction is featured centrally in the different review networks, it is mainly connected to other already popular attractions in the historic city centre. While one of the goals of the policymakers was to use the museum as a magnet for visitors into a recently redeveloped part of the city, this analysis shows, despite efforts of local businesses (Fig. 4) that international tourists tend to visit the MAS museum as a satellite-visit from the historic city centre, and not stick around to use the ancillary services like bars and restaurants. Local reviewers, however, were found to review

Fig. 4 Secondary tourist products try to benefit from proximity to core tourist cluster, picture taken from the roof of the MAS Museum (personal archive, 2016)

places in this area more often, causing a trickling down of economic benefits to the neighbourhood. This also indicates that temporarily closing a core attraction, especially when its not directly surrounded by other core attractions, will affect proximate services such as restaurants due to a likely decrease in international visitors.

The analysis of UGC in Antwerp and the example of the MAS museum highlight difficulties associated with spreading tourists over space in historic cities. However, this study shows it is not only tourists who rely heavily on the historic city centre, also Belgian and local review behaviour cluster in these areas. The limited geographical scale in which the majority of the well-connected primary and secondary touristic products and services are located can cause overcrowding and tourismification

(Russo 2002), but also makes a certain level of mixing between different groups of tourists as well as between tourists and local inhabitants possible.

This paper shows geography matters in tourism behaviour. Strong geographical clustering of reviews were found in the historic city centre and surrounding the central station, especially for international reviewers. However, review patterns are strongly influenced by the type of places, connecting mainly the most famous attractions with each other. A clear core-periphery structure, both in geography as well as in relational and thematic manner, is present (Peng et al. 2016). The core-attractions are surrounded by incidental reviews of other, relationally peripheral but geographically central service providers and less popular attractions. While these are clustered in space, they are mainly related to the nearby core attractions. Our analysis shows almost no relational clustering including secondary tourist products and ancillary services. There are thematic and geographic clusters consisting of core attractions that can be distinguished though, corresponding to the conclusion by Lew and McKercher (2006, p. 410) that "each tourist has a distinct set of motivations, resources, accommodations, services, attractions, and movements, even though they may visit many of the same attractions during a trip".

In their social network analysis of TripAdvisor reviews, Hernández et al. (2018) segment tourists a posteriori based on their review behaviour as well as a priori based on age and self-chosen traveller profile. The differences in review networks and core attractions distinguished by the found tourist segments are subsequently explored and clusters of reviewed attractions are generated from the results. This allows to learn more about different, previously unknown, segments of the tourist population and the authors argue the a priori segmentation produced the best opportunities for marketing and offering tailor-made products. Hernández et al. (2018) also argue that for core attractions, relational proximity and geographic proximity not necessarily overlap; attractions reviewed by the same tourist segments are not necessarily located close together. In our paper, we chose to segment tourists a priori based on their listed place of residents and show differences in relational and geographical clustering of reviewed places. Another difference compared to the study by Hernandez et al. was that we focus on intra-destination relations within a city and that we take into account all reviewed places, instead of just attractions. Our findings therefore provide another perspective on the possibility of applying social network analysis on UGC.

Our a priori segmentation showed tourists listing different places of residence have both different geographic as well as relational review pattern, and notable differences exist between local, domestic and international tourists. The patterns of clustering found through the spatial and relational analysis brings recommendations for destination managers as well as attraction managers and tourism SMEs. The clustering shows a pattern of complementarity and competition to be present in Antwerp. The relational analysis shows a potential for complementarity for the core attractions, but a situation of competition for peripheral nodes, consisting mainly of hotels and restaurants. The relational and geographical analysis indicate international reviewers tend to review a limited number of these service providers (e.g. the non-European review network showing a preference for places offering a 'beer and fries' type of cuisine), geographically proximate to the destination's core attractions. The autonomous forces present indicate the presence of an information cascade, which acts as a self-fulfilling prophecy guiding tourist and keeping them on the beaten tracks in a limited geographic space. Destination managers should actively engage in managing tourist flows and offering the possibility to look beyond the dominant tourism cluster and see what the destination has to offer. A way to do this is offering tourist products based on the interests of visitors as indicated by the present analysis (Hernández et al. 2018). Promoting a combination of visits to the core cluster (e.g. Medieval and Baroque heritage) with a typical Belgian gastronomic experience in a restaurant area outside the historic city centre popular by domestic reviewer could for example appeal to non–European tourists. The clustering of attractions and relations between the central clusters and ring of connected nodes do give an indication to stimulate collaboration between these nodes to come up with routes, joint marketing or joint product development which suits the needs and behaviour of tourists.

While the analysis of UGC using TripAdvisor reviews within a selected destination gives a helpful insight into tourist behaviour (Hernández et al. 2018), there are some weaknesses associated with the presented type of research. First, the size and richness of this type of UGC data is one of the main strengths, but also a threshold for transferring it into concrete policy recommendation. "The distribution and sheer amount of UGC data available present methodological challenges in collecting, organizing, and analysing the bulk of this material in a quantifiable, time-efficient, and ethical manner" (Lu and Stepchenkova 2015, p. 121). Collecting, extracting, analysing and representing UGC is necessary before it can be translated into policy measures or assist product development (Chareyron et al. 2014; Hernández et al. 2018; van der Zee et al. 2018). The bigger and richer the UGC data, the more technically and methodologically complex these steps get. In the case of the present study, the collecting of data appeared to be time-consuming, and the practically motivated choice to use reviewers of the central station brings some issues. Furthermore, there are a number of concerns with the representativeness of UGC. It is not entirely known *who* are the creators of UGC, and whether this group represents the tourist population in a destination (Johnson et al. 2012). Also, as it is not clear whether UGC reflects tourist behaviour, or limits itself to a biased snapshot of the tourist experience (Akehurst 2009; Carson 2008). Therefore, it is highly recommended to triangulate this type of research, both with other sources of UGC, as well as with studies into actual tourist behaviour.

References

Akehurst G (2009) User generated content: the use of blogs for tourism organisations and tourism consumers. Serv Bus 3(1):51–61

Alvin Chua A, Servillo LA, Marcheggiani E, Vande Moere A (2016) Mapping cilento: using geotagged social media data to characterise tourist flows in southern Italy. Tour Manag 57:295–310

Ashworth G, Page SJ (2011) Urban tourism research: Recent progress and current paradoxes. Touri Manage 32(1):1–15

Baggio R (2011) Collaboration and cooperation in a tourism destination: a network science approach. Curr Issues Tour 14(2):183–189

Baggio R, Scott N, Cooper C (2010) Network science: a review focused on tourism. Ann Tour Res 37(3):802–827
Banerjee S, Chua AY (2016) In search of patterns among travellers' hotel ratings in TripAdvisor. Tour Manag 53:125–131
Batty M (2013) The new science of cities. Mit Press, Boston
Bendle LJ (2015) The structures and flows of a large tourist itinerancy network. Curr Issues Tour 21(1):103–122. https://doi.org/10.1080/13683500.2015.1092948
Borgatti SP, Everett MG, Freeman LC (2002) Ucinet for windows: software for social network analysis. https://pages.uoregon.edu/vburris/hc431/Ucinet_Guide.pdf
Boy JD, Uitermark J (2016) How to study the city on instagram. PLoS One 11(6):e0158161
Buhalis D, Amaranggana A (2015) Smart tourism destinations enhancing tourism experience through personalisation of services. In: Tussyadiah I, Inversini A (eds) Information and communication technologies in tourism 2015, vol 1. Proceedings of the International Conference in Lugano, Switzerland, February 3–6, 2015. Springer, Cham, pp 377–389
Carson D (2008) The blogosphere' as a market research tool for tourism destinations: a case study of Australia's Northern Territory. J Vacat Market 14(2):111–119
Chareyron G, Da Rugna J, Raimbault T (2014) Big data: a new challenge for tourism. In: IEEE international conference on Big data (Big data), 2014. IEEE, pp 5–7
Cohen E (1972) Toward a sociology of international tourism. Soc Res 39(1):164–182
Edwards D, Cheng M, Wong IA, Zhang J, Wu Q (2017) Ambassadors of knowledge sharing: co-produced travel information through tourist-local social media exchange. Int J Contemp Hosp Manag 29(2):690–708
Fuchs M, Höpken W, Lexhagen M (2014) Big data analytics for knowledge generation in tourism destinations—a case from Sweden. J Destin Market Manag 3(4):198–209
Govers R, Go FM, Kumar K (2007) Promoting tourism destination image. J Travel Res 46(1):15–23
Gretzel U, Sigala M, Xiang Z, Koo C (2015a) Smart tourism: foundations and developments. Electron Mark 25(3):179–188
Gretzel U, Werthner H, Koo C, Lamsfus C (2015b) Conceptual foundations for understanding smart tourism ecosystems. Comput Hum Behav 50:558–563
Gursoy D, Del Chiappa G, Zhang Y (2017) Preferences regarding external information sources: a conjoint analysis of visitors to Sardinia, Italy. J Travel Tour Mark 34(6):806–820. https://doi.org/10.1080/10548408.2016.1237405
Hawelka B, Sitko I, Beinat E, Sobolevsky S, Kazakopoulos P, Ratti C (2014) Geo-located Twitter as proxy for global mobility patterns. Cartogr Geogr Inf Sci 41(3):260–271
Hays S, Page SJ, Buhalis D (2013) Social media as a destination marketing tool: its use by national tourism organisations. Curr Issues Tour 16(3):211–239
Hernández JM, Kirilenko AP, Stepchenkova S (2018) Network approach to tourist segmentation via user generated content. Ann Tour Res 73:35–47
Hudson S, Thal K (2013) The impact of social media on the consumer decision process: implications for tourism marketing. J Travel Tour Market 30(1–2):156–160
Hwang YH, Gretzel U, Fesenmaier DR (2006) Multicity trip patterns: tourists to the United States. Ann Tour Res 33(4):1057–1078
Johnson PA, Sieber RE, Magnien N, Ariwi J (2012) Automated web harvesting to collect and analyse user-generated content for tourism. Curr Issues Tour 15(3):293–299
Kádár B (2014) Measuring tourist activities in cities using geotagged photography. Tour Geogr 16(1):88–104
Kennell S, Rushton AM (2015) User-generated content (UGC) in pleasure travel decision-making
Kwok L, Yu B (2013) Spreading social media messages on Facebook: an analysis of restaurant business-to-consumer communications. Cornell Hospitality Q 54(1):84–94
Law R, Qi S, Buhalis D (2010) Progress in tourism management: a review of website evaluation in tourism research. Tour Manage 31(3):297–313
Leung XY, Wang F, Wu B, Bai B, Stahura KA, Xie Z (2012) A social network analysis of overseas tourist movement patterns in Beijing: the impact of the Olympic games. Int J Tour Res 14(5):469–484
Lew A, McKercher B (2006) Modeling tourist movements: a local destination analysis. Ann Tour Res 33(2):403–423
Liu FJ, Zhang J, Zhang JH, Chen DD (2010) Roles and functions of provincial destinations in Chinese inbound tourist flow network. Geogr Res 6:1141–1152

Liu S, Law R, Rong J, Li G, Hall J (2013) Analyzing changes in hotel customers' expectations by trip mode. Int J Hosp Manag 34:359–371
Lu W, Stepchenkova S (2015) User-generated content as a research mode in tourism and hospitality applications: topics, methods, and software. J Hosp Market Manag 24(2):119–154
Marchiori E, Cantoni L (2015) The role of prior experience in the perception of a tourism destination in user-generated content. J Destin Market Manag 4(3):194–201
Marine-Roig E, Clavé SA (2015) Tourism analytics with massive user-generated content: a case study of Barcelona. J Destin Market Manag 4(3):162–172
Marine-Roig E, Clavé SA (2016) A detailed method for destination image analysis using user-generated content. Inf Technol Tour 15(4):341–364
McKercher B, Shoval N, Ng E, Birenboim A (2012) First and repeat visitor behaviour: GPS tracking and GIS analysis in Hong Kong. Tour Geogr 14(1):147–161
Miguéns J, Baggio R, Costa C (2008) Social media and tourism destinations: TripAdvisor case study. Adv Tour Res 26(28):1–6
Mkono M, Tribe J (2017) Beyond reviewing: uncovering the multiple roles of tourism social media users. J Travel Res 56(3):287–298
Munar AM, Jacobsen JKS (2014) Motivations for sharing tourism experiences through social media. Tour Manag 43:46–54
O'Reilly T (2005) What Is Web 2.0, design patterns and business models for the next generation of software. Online available at: http://www.oreilly.com/lpt/a/6228.
Peng H, Zhang J, Liu Z, Lu L, Yang L (2016) Network analysis of tourist flows: a cross-provincial boundary perspective. Tourism Geographies 18(5):561–586
Russo AP (2002) The "vicious circle" of tourism development in heritage cities. Ann Touris Res 29(1):165–182
Russo AP, Clave SA, Shoval N (2010) Advanced visitor tracking analysis in practice: explorations in the PortAventura theme park and insights for a future research agenda. Inf Commun Technol Tour 2010:159–170
Shih HY (2006) Network characteristics of drive tourism destinations: an application of network analysis in tourism. Tour Manag 27(5):1029–1039
Shoval N (2018) Urban planning and tourism in European cities. Tourism Geographies 20(3):371–376
Shoval N, Isaacson M (2007) Tracking tourists in the digital age. Ann Tour Res 34(1):141–159
Tilly R, Fischbach K, Schoder D (2015) Mineable or messy? Assessing the quality of macro-level tourism information derived from social media. Electron Markets 25(3):227–241
Toerisme Vlaanderen (2018) Toerisme in cijfers XL 2017. Brussels: Toerisme Vlaanderen. https://www.toerismevlaanderen.be/toerisme-cijfers-2017-xl. Accessed 29 Nov 2018
Tripadvisor (2016) Tripadvisor website. https://www.tripadvisor.com/. Accessed 13 Oct 2016
van der Zee E, Bertocchi D, Vanneste D (2018) Distribution of tourists within urban heritage destinations: a hot spot/cold spot analysis of TripAdvisor data as support for destination management. Curr Issues Tour. https://doi.org/10.1080/13683500.2018.1491955
Versichele M, Neutens T, Delafontaine M, Van de Weghe N (2012) The use of bluetooth for analysing spatiotemporal dynamics of human movement at mass events: a case study of the Ghent festivities. Appl Geogr 32(2):208–220
Wyss R, Luthe T, Abegg B (2015) Building resilience to climate change—the role of cooperation in alpine tourism networks. Local Environ 20(8):908–922
Xiang Z, Gretzel U (2010) Role of social media in online travel information search. Tour Manag 31(2):179–188
Ye Q, Law R, Gu B, Chen W (2011) The influence of user-generated content on traveler behavior: an empirical investigation on the effects of e-word-of-mouth to hotel online bookings. Comput Hum Behav 27(2):634–639
Zeng B, Gerritsen R (2014) What do we know about social media in tourism? A review. Tour Manag Perspect 10:27–36