Annals of Tourism Research 87 (2021) 103113

Contents lists available at ScienceDirect

Annals of Tourism Research

journal homepage: https://www.journals.elsevier.com/annals-of-tourism-research

Research Article

Tourism flows in large-scale destination systems

Bálint Kádár [a,*], Mátyás Gede [b]

[a] Budapest University of Technology and Economics, Department of Urban Planning and Design, 1111 Budapest, Műegyetem rkp 3, Hungary
[b] Eötvös Loránd University, Department of Cartography and Geoinformatics, 1117 Budapest, Pázmány Péter stny. 1/A, Hungary

ARTICLE INFO

Article history:
Received 22 January 2019
Received in revised form 15 November 2020
Accepted 28 November 2020
Available online xxxx

Associate editor: Shoval Noam

Keywords:
Tourism networks
Network analysis
Tourist flows
Large-scale destinations
Multi-destination trips
Danube
Flickr analysis

ABSTRACT

Large-scale destination systems, especially cross-border regions are less studied in literature as their size and transnational nature makes these hard to analyse with traditional methods. Tourism systems like the Danube Region are composed of several local and regional destinations, and even when these are branded together for tourists the integration of these into one system is often compromised by national boundaries and socio-economic differences. This study shows how the Danube region is composed of different clusters of destinations, and how national boundaries have a strong shielding effect in the interregional movements of tourists. A methodology based on network analysis with efficient clustering algorithms applied on large geotagged datasets from User Generated Content is proposed. Flickr data was used to map short time-interval visitor flows along the linear system of the river Danube. 18 regional clusters integrated into 3 strong, but separated destination systems were identified by modularity analysis. The central integrating effect of the large capital cities and the boundary-shielding effect impeding the total integration of this large-scale system were made measurable.

Introduction

The conceptualization and analysis of large-scale destination systems are emerging fields in tourism research, helped by Big Data analysis and network analytic tools. On the other hand the definition of destination itself is problematic in such scale (Jovicic, 2016; Pearce, 2014). Multi-destination trips (Lue et al., 1993) have been studied in various scales, but most of the studies for larger systems reaches still only a regional scale (Asero et al., 2016), except from those which analyse international traffic between separate destinations (Novotna, 2018). Large scale tourism regions usually comprehend many destinations, having complex interrelationships between visitors and the industry inside the spatial system. The largest integrated tourism systems in the literature are entire nations usually visited as complex travel destinations, e.g. New Zealand (Balli et al., 2015; Pearce & Heike, 2015), or even transnational destinations, like large mountain areas e.g. the Pyrenees (Balli et al., 2015), or seaside destination regions like the Caribbean (Bangwayo-Skeete & Skeete, 2015; Lorde et al., 2016; McLeod et al., 2017). A higher level of complexity characterizes transnational destinations where different socio-political systems are present often divided by physical (and intangible) boundaries. This paper aims to explore the possibilities to define geographically large-scale destination systems based on network and cluster analysis of tourism flows mapped from user generated Big Data. The levels of integration between places visited in large-scale transnational systems depend on their thematic and infrastructural connections, distance, and the effect of boundaries inside the system. Therefore, there are two main research questions to answer in this paper. First, on which levels of integration

* Corresponding author.
E-mail addresses: kadarb@urb.bme.hu (B. Kádár), saman@map.elte.hu (M. Gede).

https://doi.org/10.1016/j.annals.2020.103113

can we call a large-scale multi-destination system an integrated destination? Second, how much does the boundary-shielding effect impedes the integral development of such a large-scale destination system?

Defining and mapping large-scale tourism regions

The analysis of large-scale tourism destinations is problematic in part for the conceptualization of destinations themselves. A number of authors highlight the inconsistencies of the literature in using this term (Framke, 2002; Jovicic, 2016; Pearce, 2014). Large-scale tourism regions can be analysed in the integrative conceptual framework of tourism destinations proposed by Pearce (2014): these can be defined geographically, but also as complex networks of industrial organization, or as dynamic adaptive systems. The later are the most recent conceptualizations (Baggio & Sainaghi, 2011), where the supply and demand sides both can change in time, forming a space of "variable geometry" (Beritelli et al., 2014). However, the geographical attributes of tourism destinations still represent the key component of their resource base (Jovicic, 2016). Pearce and Heike (2015) analysed destinations from tourists' perspectives; most respondents traveling in New Zealand felt a destination is "a geographical area within which the tourist enjoys various types of tourism experiences", and many agreed that it is "a bundle of tangible and intangible attributes, and it can potentially be seen as a product and a brand", but almost no one responded that it is "the location of a cluster of attractions and related tourist facilities which a tourist or tour group selects to visit or which providers choose to promote" (p 9.). Their results highlight the importance to form a network of smaller tourism clusters in geographical, organizational and branding sense as well, especially in larger-scale destination systems subject to multi-destination trips.

Multi-destination trips occur between places of attraction, mostly inside a larger-scale tourism region. The patterns of such trips define the structure of the tourism region itself. McKercher and Lew (2004) conceptualized different touring components of itineraries, among them multiple transit and destination components, also stating that distance decay affects tourist behaviour. Later they defined 4 territorial models and 3 linear path models of tourist behaviour in local destinations, among them the touring point-to-point pattern applicable to linear destinations (Lew & McKercher, 2006). While these widely used categories can help to understand visitor behaviour also in more complex tours, the five patterns identified by Lue et al. (1993) are still the most accurate for describing multi-destination trips: single destination, en route, base camp, regional tour, and trip chaining. In their paper they theorized cumulative attractions, stating that the sum of attractions completing each other creates a more attractive destination than single site attractions.

The understanding of tourism flows in large-scale tourism regions through complex system analysis is an important area today in tourism research. Apart from urban areas (Grinberger et al., 2014; Kádár, 2013), the theories of multi-destination trips were applied to wine tourism regions (Ferreira & Hunter, 2017), national parks (Connell & Page, 2008), and for entire countries (Balli et al., 2015) as well as to smaller tourism regions. Chua et al. (2016) used user generated content and advanced visualization techniques to analyse tourist flows based on 72,031 geotagged tweets posted by 3135 unique individuals from a tourism region in Southern Italy. They separated domestic and foreign tourists, tracing their movements in time separately, identifying popular locations. Nine of such interconnected tourism areas were identified in the Pyrenees mountain region by Blasco et al. (2014); they searched for hub-and-spoke tourist consumption patterns using hierarchical cluster analysis based on drive-time between the 321 attractions. They found that all of the tourism clusters were cross-border areas, corresponding to historical regions rather than to administrative regions in the Pyrenees. In fact, in order to define large-scale destination systems such as the Pyrenees, the clusters of well-integrated tourism areas and their basic components must be analysed, and the inter-connectedness of elements inside the clusters and between the clusters must be considered. Such methods were used by Peng et al. (2016) studying tourist flow networks of cross-boundary tourism between Chinese regions. They used network-structure analysis to find out how cross-boundary tourist flows are significantly influenced by the boundary-shielding effect, stating that such analysis is useful also in cross-national scales. This paper builds on the theoretical findings of the abovementioned studies, analysing a linear, but even larger scale tourism network from the demand side of visitors, finding tourism clusters at different scales and analysing the boundary-shielding effect separating these.

Analysing tourism destinations as networks

Network analysis in tourism literature boomed in the past 10 years; three different streams of research could be identified (Liu et al., 2017 p.133). First, the analysis on tourism research collaboration and knowledge creation, not related to tourism destinations. Second, the network analysis based on tourist movements and behavioural patterns, which is the field of this paper. And third, the analysis on the tourism supply, destination, and policy systems, contributing to the understanding of destination systems just as much as the second stream.

Research on the tourism supply often uses Social Network Analysis (SNA) to explore the properties of the organization of the tourism industry in a destination (Baggio, 2011; Baggio et al., 2010; Nicolosi et al., 2018; Scott et al., 2008; Valeri & Baggio, 2020). Stakeholders' collaboration and other interconnections between tourism businesses are generally analysed in such research. Scott et al. (2008) compared the structural properties of interorganizational networks within destinations between four different tourism regions in Australia; they stated that the analysis of the structure of the network allows strategic weaknesses in the cohesiveness of the destination to be addressed by policy and management. Networks revealed in this and other studies were reanalysed by Baggio (2020) in order to prove that different destinations hold similar characteristics regarding the supply side, therefore similar managing strategies could be used. While studies on visitor behaviour in space can lead to better attraction design and spatial planning implications, and reveal the demand side of tourism, on the supply side SNA can give powerful feedbacks to Destination

Management Organizations (DMO) for better management and policy planning, especially when research takes a step forward from descriptive analytics to comparative or inferential methods to understand the structure of complex tourism systems (Williams & Hristov, 2018), or to mixed methods of quantitative and qualitative analysis (Mariani & Baggio, 2020). As SNA data is not confined by the geographic relations of the analytic of tourist movements, there are much more possibilities to apply complex network analytics, or even to use statistical approaches like the probabilistic Exponential Random Graph Models (ERGM), frequently used in social sciences, but only recently introduced to tourism research (Khalilzadeh, 2018; Lyócsa et al., 2019; Williams & Hristov, 2018).

This paper analyses a geographically determined system of tourist flows. Most of the research involving network analytics on the demand side of tourism constructs descriptive graph models, analysing the network's clustering patterns, its connectedness and centrality measures and structural holes (Shih, 2006). Centrality measures show how important are certain nodes in a network (Scott & Carrington, 2011). In a tourism system where nodes are the attractions and edges are the connections weighted by the number of visitors moving between nodes degree centrality shows how many other attractions are directly connected to a node, closeness centrality shows how easily all other nodes can be reached, and betweenness centrality shows the frequency a tourist would stop at a node between two other nodes in the system. Structural holes theory (Burt, 1992) helps in identifying nodes that control without alternative linkages subgroups of a network, therefore attractions in a stronger position and isolated clusters within the network can be identified.

Hwang et al. (2006) examined international tourists' multicity trip patterns, defining multi-destination trip patterns as network structures, proposing centrality and connectedness measures to describe such travel systems, finding main hubs and less connected sub-systems. They emphasized the role of hub cities and the role of collective advertising and increasing connectivity between destination cities to form multi-trip destination systems. Smallwood et al. (2012) used the ArcGIS environment and its Network Analyst to draw the network of 1208 tourists giving travel information in face-to-face interviews from a 300 km long coastal Marine Park in north-western Australia, showing decisive space usage patterns. Grinberger et al. (2014) applied a clustering algorithm in the same GIS environment to group tourist behaviour mapped by GPS trackers, while Shoval et al. (2015) applied a sequence alignment method to find spatio-temporal categories. Baggio and Scaglione (2017) analysed the travel patterns of 18,138 anonymized mobile users of a major Swiss mobile provider, defining clusters from the modularity analysis. Asero et al. (2016) in a study based on 3182 face-to-face interviews with self-organized tourists visiting Sicily used CONvergence of iterated CORrelation (CONCOR) procedure to define clusters. They demonstrated that tourist choice defined the role of a destination as "central" or "peripheral" within a network, where tourists build their own networks around nodal destinations. Consequently, cluster detection is one of the commonly used result of network analysis in tourism studies.

Geographical network analysis in tourism studies often use weighted networks. Tourism flows are aggregated from direct or indirect sources, and the weight of connections between nodes defines structural proximity of the nodes. Shih (2006) evaluated travel data from a survey of 2142 respondents in Taiwan, drawing a weighted network of a tourism region among the firsts. More recently Lee et al. (2013) identified the centrality of different villages in a Korean tourism region, where the weight of connections between villages were estimated applying a gravity model using path length, population and rural amenities. González-Díaz et al. (2015) used network analysis to compare the most relevant structural changes in the configuration of the accommodations network in Spanish regions. Zach and Gretzel (2012) collected 1009 surveys from visitor centres on places visited, demonstrating that the tourism network in Northern Indiana has an extremely low density. The abovementioned analyses relied on surveys to get data on spatio-temporal tourist behaviour. Surveys are a valid method to construct the analytical framework of a tourist network in space, but they lack the geographic accuracy to unfold all spatial connections between nodes, and they lack the quantity of samples to describe exact correlations with the actual number of tourists moving in the system (Kádár, 2014). Accurate tracking technologies have been widely used in tourism research in the past decade (Shoval & Ahas, 2016). Recently researches used such technologies to arrive to extremely detailed descriptions of complex tourism destinations, defining exact weight of all connections between destinations. Some of these studies draw conclusions without network analysis (Chua et al., 2016; Connell & Page, 2008), others use network analysis without considering the actual weight of connections in high-resolution systems, e.g. using Space Syntax methodology (Kádár, 2013; Y. Li et al., 2016), but there are very few studies similar to present research using accurate weighted network analyses based on visitor flows, allowing the description of the exact structure of the tourism area (Peng et al., 2016; Shih, 2006; Taczanowska et al., 2014).

Big data in the research of tourism flows

Depending on the input data, tourist flows in different scales can be analysed. This ranges from movements of extreme geographic precision inside a destination (Grinberger et al., 2014) to movements between cities and regions worldwide based on statistical data on international tourism arrivals (Lozano & Gutiérrez, 2018). Data quality is influenced by the quantity of tracks recorded. Mobile cell data (Ahas et al., 2007; Baggio & Scaglione, 2017) allows the analysis of large samples of population in large areas, even at a national scale, making it an ideal tool for large-scale tourism system analysis, if the scale is inter-urban, as the resolution of the geographic positioning is low. User Generated Content (UGC) from online services have often accurate GPS coordinates associated, and quantities of such data allowing Big Data analysis can be acquired from some of the services (Birenboim & Shoval, 2016; Li et al., 2018). Most large social networks today allow no access to their databases, therefore researchers in tourism geography use Twitter (Chua et al., 2016; Liu et al., 2018) or Flickr, as these first-generation social-web services still have open access APIs, unlike the defunct image sharing service Panoramio (Encalada et al., 2017; Orsi & Geneletti, 2013), or Instagram, where API access is discontinued since 2016. The most popular geo-spatial data source, Flickr also discontinued its free image hosting service from February 2019, closing a flourishing decade of geo-tag based research. The many publications from this decade indicate how Flickr geo-tagged images can measure tourism demand from an interregional scale down to the urban scale. At the urban scale Kádár (2014) described the tourist and local space usages of 3 cities, comparing these based on 150 thousand geo-tagged images, while Straumann et al. (2014) collected 81,194 images from Zürich, applying network analysis to visualize different space usage of domestic and foreign photographers. At a regional scale Girardin et al. (2008) analysed the trajectories of 2880 users visiting Central Italy, comparing tourist itineraries of users of different nationalities. More recent research using even larger datasets showed the potential of Flickr data in large-scale tourism systems. Cai et al. (2014) analysed 383,335 entries from Australia showing cross-country trajectories based on temporal data. Önder et al. (2014) mapped tourism demand for the entire country of Austria based on 1,183,889 photos. Finally, Paldino et al. (2015) use a dataset of approximately 70 million photographs from the 10 most photographed cities of the world, measuring attractivity through mobility network analysis, but at the same time demonstrating the street-level accuracy of the method.

Abovementioned works show that Flickr is a valid research database for measuring tourism demand, even if not all social groups or travellers use such platform at a same extent, Lo et al. (2011) show that image hosting services like Flickr.com are more widely used by middle-aged and older social groups than all other social media, therefore these sources represent better the overall travel patterns than any other type of UGC. The ultimate advantage of the Flickr database is that it can deliver comparable data on visitor flows from different parts of the world. Its availability across regions and countries can constitute the basis of network analysis in large-scale destinations, delivering evidence on the interconnectedness or fragmentation of a tourism region large as the one of the Danube river.

The Danube as a tourism region

The Danube is the most international river, flowing through 10 countries for 2860 km, connecting east with west through Central Europe. This is one of the most international spatial systems branded as an integrated tourism destination, still, its large-scale linear geography and trans-national cultural diversity poses the question whether it can be or not considered a unified tourism system. As its spatial system passes through many national boundaries, and no previous studies analysed the interconnectedness of the well-known destinations along it, the study of tourism along this river has the potential to answer the research questions posed in this paper.

A unique aspect of the Danube as a tourism region lies in its linearity, typical to historic routes, e.g. the U.S. Route 66 (Caton & Santos, 2007), or pilgrimage routes, e.g. the Camino de Santiago (Amaro et al., 2018; Lois-González & Santos, 2015). However, it is a more complex system, because of the complexity of the tourism offers at one side, and the fragmentation of its nationalities, socio-cultural and economical status and division by national boundaries on the other.

Regarding the tourism offers there are evenly distributed cultural and natural destinations along the river besides all forms of river related tourism collected by Steinbach (1995), and several tourists visit more types of these in one travel. The Wachau Cultural Landscape, the historic cities of Budapest, Vienna and Regensburg, monastery of Ivanovo and the natural landscapes of the Srebarna Natural reserve and of the Danube Delta are part of the UNESCO World Heritage. Other sites, like the Djerdap National Park of the Iron Gates or the Danube Limes and many unique historical places in Esztergom, Komarno-Komárom and Smederevo are on the tentative list. There most common subjects of tourism analytics in the region are the main capital cities on the river: Budapest (Puczkó et al., 2007; Rátz et al., 2008; Smith et al., 2018), Vienna (De Frantz, 2018; Kádár, 2013, 2018) and Belgrade (Joksimović et al., 2014). Similarly studied by tourism scholars are the Danube Delta region (Damian & Dumitrescu, 2009; Hall, 1993), the Wachau region in Austria (Meschik, 2012; Ploner, 2009) and the Iron Gate area (Boengiu, 2012; Mazilu, 2011).All the above-mentioned destinations brand themselves with the Danube river; on the other hand, still few are the multi-destination tourism products along the river. Exceptions are the river cruise ship tours (Dragin et al., 2007; Dragin et al., 2010), the transnational Eurovelo 6 bicycle route (Meschik, 2012; Vujko et al., 2013), and some cultural routes promoted by the Council of Europe, like the Roman Emperors & Danube Wine Route (Terzić & Bjeljac, 2016). In Europe the highest number of cruise ships operates on the Danube and the situation is similar in bicycle tourism, as the Danube bike trail from Donaueschingen to Vienna, and in many cases until Budapest is the most trafficked bike trail in the continent (Steinbach, 1995). The transport methods between singular destinations along the Danube are therefore the cruise ship, the bicycle, and the car, except between capital cities, well connected by train and by air carriers, enabling multiple city visit trips between cities like Vienna, Bratislava and Budapest (Kádár, 2014). There is a strong professional effort to reveal the joint potentials of the Danube as a European tourism destination. The Danube region was acknowledged by the European Union as a standalone region in 2009, and since then there are many programmes following a common strategy to unite the different regions along the river (Busek & Gjoreska, 2010). Projects funded by the European Commission such as DATOURWAY (Talabos, 2014) or DANUrB (Kádár & Vitkova, 2019) work in a transnational cooperation to create a unified Danube tourism brand.

However, the tourism geography of the Danube today is still fragmented, mostly analysed in national perspectives (Widawski & Wyrzykowski, 2017). In countries like Bulgaria and Croatia there are no connecting tourism regions important at a national level, but also in Slovakia (Kasagranda et al., 2016) or even in Germany (Oppermann, 1996) only cruise ship and bicycle tourism is relevant along the river. Successful tourism destinations are the Wachau region in Lower Austria (Meschik, 2012), but also the region around Linz in Upper Austria (Iordanova, 2017), the Hungarian Danube Bend in Hungary, and the developing region of the Danube Delta in Romania (Hall, 1993). The border regions between Slovakia and Hungary, Croatian and Serbia, Serbia and Romania, and Bulgaria and Romania could not produce popular destinations for tourism, the Iron Gates

Fig. 1. Economical differences and separative borders along the Danube river.

region is the only one where measurable cross-border tourism is present (Mazilu, 2011). It must be also noted, that Upper-Danube regions have much better economic outcomes than others on the Lower Danube (Fig. 1). The main reasons of the fragmentation of this region are historical: the Iron Curtain separated the socialist block from western countries until 1989, while the Yugoslavian war and the break of the federal state first impeded all traffic along the Danube, and after created new national borders dividing a previously functional tourism region (Lagiewski & Revelas, 2004). Today this is a border region of the European Union, with the Danube constituting interior borders between member states inside and out the Schengen area, and crossing the non-EU member Serbia. Therefore, it is not surprising, that tourism by the Danube developed in an uneven pace until recently, and the complete integration of the area will be possible only after the foreseen integration of the whole region into the EU.

Because of the still existing borders and uneven development, the Danube as a tourism region is an ideal field of analysis to demonstrate the real boundary-shielding effects inside a tourism system. It is supposed, that there is a certain level of integration of the tourism system all along the river, but the question is how to measure this and is it enough to form one large-scale tourism destination or do destinations along the Danube remain isolated. The supply side of tourism by the Danube should also be analysed with SNA methods to fully answer this question, while this paper analyses the demand side, applying a geographical approach to the question of a destination (Pearce, 2014). The basic theorization of MacCannell (1976) is still valid: in order to have tourism in a destination a Site, a Marker which makes the site significant, and the Tourist itself who visits the site is all needed at once. The Danube as a large-scale tourism region has two of the three criteria: a comprehensive geographical Site (river) with a set of attractions and services well Marked as historically and geographically integer and interesting. This paper focuses only on the demand side: are there tourists who take multi-destination trips connecting the different attractions of the Danube in numbers to call the whole region one large-scale destination?

Methods and results

In this paper a methodology is presented to draw a map of tourism demand and visitor flows using User Generated Content, converting Flickr datasets to weighted graphs of connectivity in order to define the structural integrity of the Danube as a large-scale tourism destination system. The spatial system to be analysed comprehends all administrative territories at the community level that have boundaries with the Danube. The initial phase of this study showed that there is an imbalance in the entry-points of the tourism system of the Danube: Vienna, Budapest, Bratislava and Belgrade are large capital cities with international airports, while other larger cities with good transport connections are Novi Sad, Linz, Regensburg, Ulm. In the Lower-Danube region there are no such hubs on the river from where tourists can start their journeys. In fact, visitors arriving to the Danube Delta or to other Lower-Danube destinations usually arrive by air via Bucharest, the Rumanian capital, situated only 60 km from the Danube. Accordingly, Bucharest was added to the spatial system, cross-checking that no other transport hubs can be found in such range along the river.

A Flickr database of the Danube

A recursive mechanism for downloading large amounts of metadata of geotagged photographs within a geographic quadrangle was used (Kádár & Gede, 2013). In this case bounding boxes of every settlement along the Danube were created; as these boxes overlap each other, in a pre-processing step they were transformed to a set of adjacent, non-overlapping rectangles (Fig. 2). The data acquiring programme was run on these rectangles and resulted in a database of 2,187,243 geo-tagged photos uploaded until December 31. 2017.

In the following step, each photograph's metadata was completed with the name of the settlement of whose area its position belongs to, as well as the river kilometre of the closest point along the Danube and whether it is on the left or the right riverbank (river kms are calculated from the delta of the Danube). These additions made it possible to create various statistics and data visualizations of the spatial distribution of photographs. To validate the method, the correlation have been calculated between existing statistical data on the bednights spent in cities by the Danube (TourMIS, 2018) and the number of geotagged images downloaded from those urban areas from where the statistical data could have been retrieved for the year 2017. A significant positive correlation of 0.9878 was found, the diagram on Fig. 2 shows little deviation for any of the destinations with existing statistical data.

An interactive application was developed to visualize the data (Fig. 3, http://mercator.elte.hu/~saman/de/). The number of photographs on each side of the river is marked by scaled red (left) and green (right) bands. Hovering the mouse over the bands reveal the exact number of photographs belonging to that specific section (sections are one or ten km long, depending on the zoom level).

To be able to visualize the popularity of all destinations along the Danube at once, a linear diagram was created, representing the 2800 km on one line, still distinguishing photos taken in the left and right banks (Fig. 4). In this diagram the photos taken less than 1 km from the middle line of the river were represented as blue lines, while the photos taken in farther areas of riverside municipalities were represented by yellow lines. The diagram gives an instant overview of the unbalanced attractivity of the different sections of the river, highlighting the hubs – mainly the capital cities along the river, and showing how the Lower-Danube region is much less used for photographic and tourism related activities than the Upper-Danube.

The network of multi-destination visits along the Danube

To reconstruct the movements of the visitors, the river was cut into 10 km long segments. As some larger cities (e.g. Budapest, Vienna or Belgrade) stretch beyond one single segment, such adjacent segments were united into one node, as this scale was not meant to represent the movements inside cities, rather than between cities. Flickr data was grouped and aggregated by different users, by river segments (left and right bank data united), and by the time it was taken (with one-hour granularity) – defining a spatio-temporal sequence for each user. The spatio-temporal sequences were not used in this case to define different travel patterns, but to create the weighted network of all multi-destination visits. Aggregated data was transformed into a graph where nodes are the river segments and edges between them indicate the number of cases when a user's time series include a transition from one segment to another within two days; directionality of the travel was not considered. The limitation of 48 h made possible to show accurately which destinations were part of one journey. In this network the "weight" of edges equals to the number of users moving between the nodes; a total of 75,958 such movements were mapped on 2293 connections between 270 nodes.

The graphs were elaborated in the graph analytic software Gephi v0.9.2 (Bastian et al., 2009), capable of different calculations and visualizations of networks. Gephi is one of the most widely used network analytic tools in social and geographical network analysis, having integrated algorithms to calculate centrality measures, structural holes, and clustering calculations, having very powerful visualization options to show the results of such calculations. Its only limitation lies in the handling of very large datasets over 100,000 connections, because it uses the resource demanding Java platform requiring large amounts of RAM.

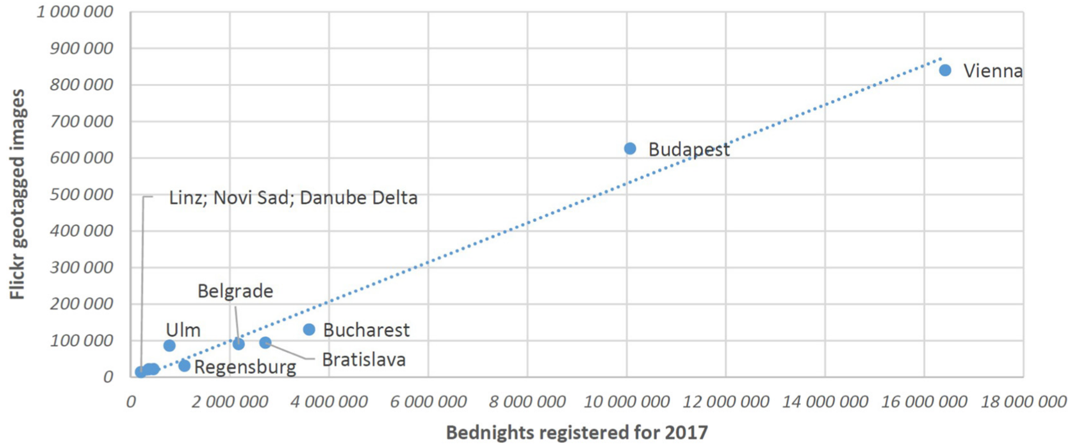

Fig. 2. Correlation between Flickr geotagged images and number of overnight stays in selected destinations by the Danube.

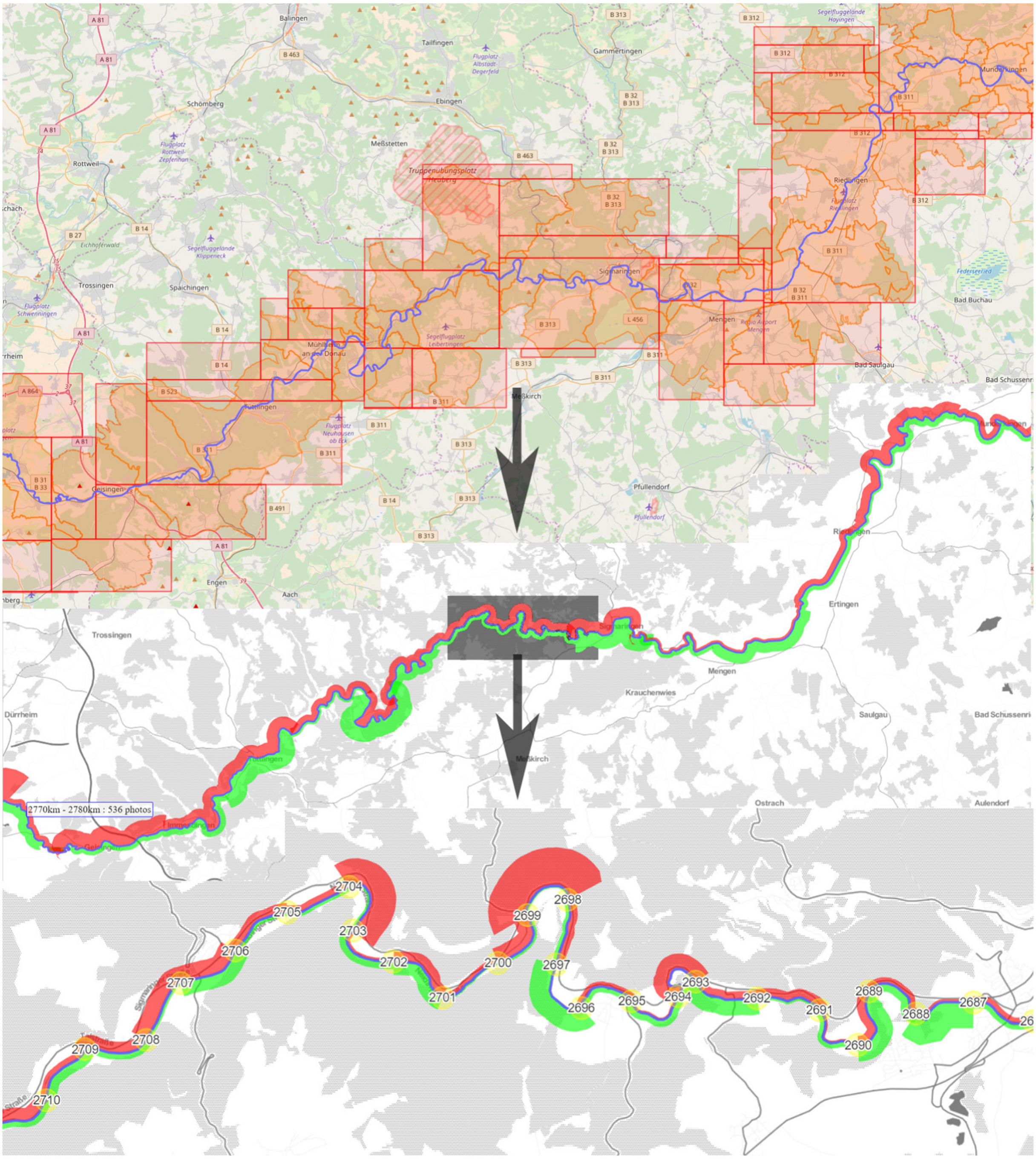

Fig. 3. Creating the dataset of Flickr photo distribution along the Danube. Top: the bounding boxes of settlements from where metadata was downloaded. Middle: data visualized on the zoomable web-map tool "Danube Explorer". Bottom: zoomed in section showing a scaled diagram of the number of photographs between each river km separated by right and left bank.

However, it is an ideal tool for datasets having the scale of present study for the 270 nodes of Danube's network. The Geo Layout visualizes all nodes in their exact geographical positions according to their coordinates (Fig. 5), showing tourist movements between all 10 km long sections of the Danube, with the added node of Bucharest. The weight of the edges shows how many users of Flickr made journeys between different sections of the Danube, while the size of the nodes is proportional to the number of photographs taken in a given section (Fig. 6). The connections between capital cities are explicitly larger than connections between any neighbouring sections.

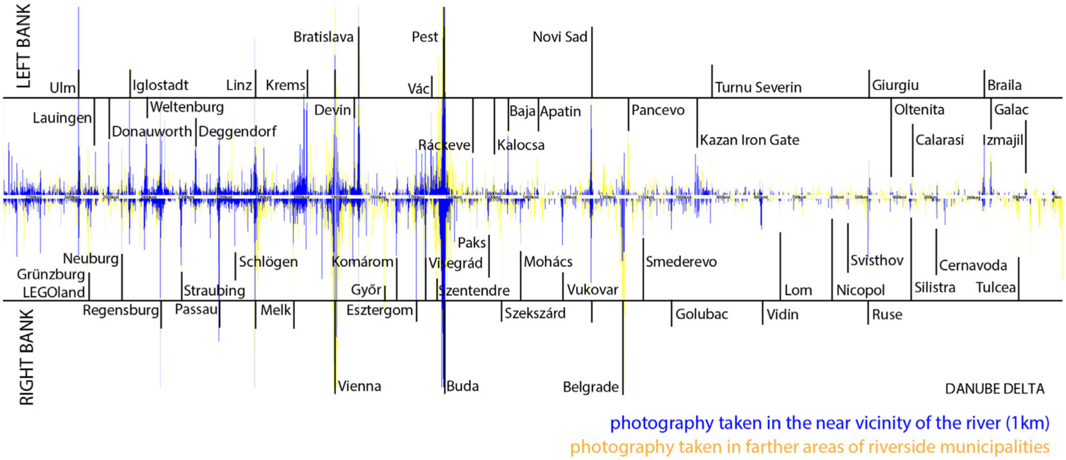

Fig. 4. Flickr photo distribution along the Danube, measured in the administrative areas of municipalities adjacent to the river, separated to right and left bank areas.

Network analysis and cluster detection

The visualization of networks based on the geographical location (Geo Layout) of nodes is a special feature of Gephi, but its force-directed layouts highlight best the main properties of weighted networks (Novotna, 2018). Force Atlas 2 (Jacomy et al., 2011) algorithm shows the clusters of places most connected to each other, while it has extended behaviour alternatives to fine-tune the representation, therefore the effect of weighted edges can be better visualized. Gephi also calculates the network properties of all nodes, in this research their closeness and betweenness centrality values were considered. The overall network of tourism flows along the Danube has an average degree of 16.985 and a weighted degree of 562.652. In Fig. 7 the size of the nodes is proportional to their betweenness centrality, indicating how often that node appears on the shortest path between all other nodes in the network. Larger nodes therefore represent locations which are entry hubs to specific sections of the Danube, usually these are cities with good transport connections, serving as centres of a smaller tourism region along the river. The network analysis of all journeys along the Danube clearly shows the clustering of the different sections of the river, describing the functional structure of tourism (Fig. 7). Vienna and Budapest are the most central hubs of the system, well connected to Belgrade, serving as central hub to all Serbian and Croatian sections of the river. Surprisingly Bucharest is only connected to Vienna, and most Romanian sections are loosely tied to it. All capital cities are in the middle of the network, their connections to each other are one order of magnitude stronger than their connections to sub-centres, and these are still one order of magnitude stronger than all other connections in the system. This network gives an accurate picture of the hierarchy of tourism hubs along the Danube, but the centrality of capital cities hides the linear connectivity of the subsequent sections of the river.

In the second step all capital cities were excluded from the network in order to analyse the continuity of the tourism system along the Danube. Connections between capital cities are in larger part not related to other destinations along the Danube. Each capital is at once an international destination for cultural tourism per se, and a national hub with the major airport and train stations to send outbound tourists to other destinations (Novotna, 2018). The Danube can be considered a destination system if there is multi-destination travel along the river, and the data of capital cities compromises the analysis of this aspect. Compared to Fig. 7 the new network of Fig. 8 shows much better the linearity of this large-scale system. Without the extreme travels to and from capital cities the network is much more balanced. The difference in the betweenness centrality of nodes comes out more evidently; mostly the larger cities (Ulm, Regensburg, Linz, Novi Sad, Ruse) or important historic centres (Melk and Krems, Szentendre, Vidin) are the hubs of the different regional tourism clusters.

Tourism clusters were defined by the modularity analysis of the network. To determine regional tourism clusters, travel between and inside clusters had to be separated. A tour to a cluster begins at an entry point and finishes at an exit point, and it is explored inside the cluster on a multi-destination itinerary. Using clustering algorithms on time-space sequences of users (Grinberger et al., 2014; Shoval et al., 2015) would lead directly to relevant clustering results, but in order to analyse the overall space system the sequences were discarded, and only the number of visitors moving from one node to another were kept. This way connections between popular clusters have similar weights as connections inside the clusters, and to compensate this travel distances had to be considered. The weight of connections between nodes was divided by the geographical distance of the nodes in order to have a weighted network representing travels in a regional scale. The "Modularity" function of Gephi with resolution set to 0.9 was used to find 18 strongly connected clusters. This algorithm uses the Louvain method (Blondel et al., 2008) which optimizes the partition of the graph to get the maximum possible modularity value (a scalar between −1 and 1 measuring the density of links inside communities as compared to links between communities). The modularity index is $Q = 0.866$, which shows a set of well-defined clusters ($Q = 0$ means no separation, $Q = 1$ total separation).

Fig. 8 shows the network excluding the five capital cities, with network connections weighted only by the number of movements, but also showing the 18 regional clusters defined with the weights divided with distance. This network has an average

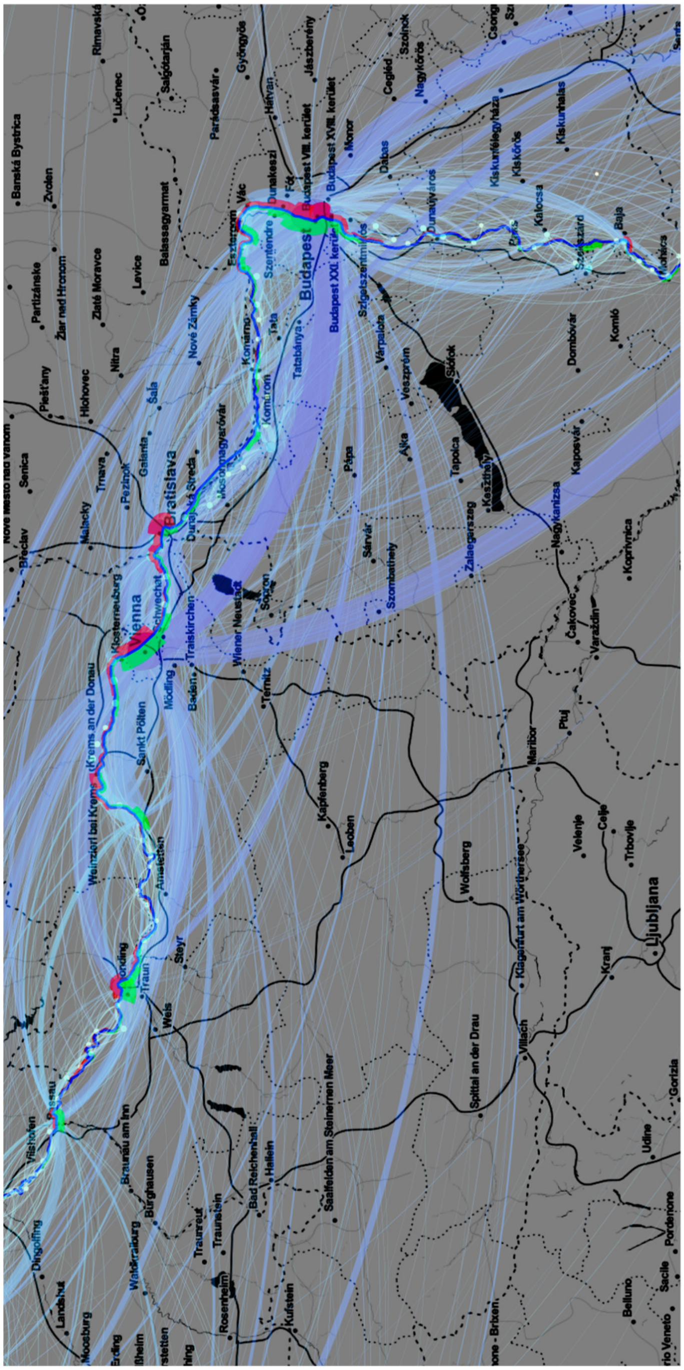

Fig. 5. Visitor flows between destinations along the Danube superposed on the "Danube Explorer" map; thickness of connections is proportional with the number of users moving between the places (detail of map).

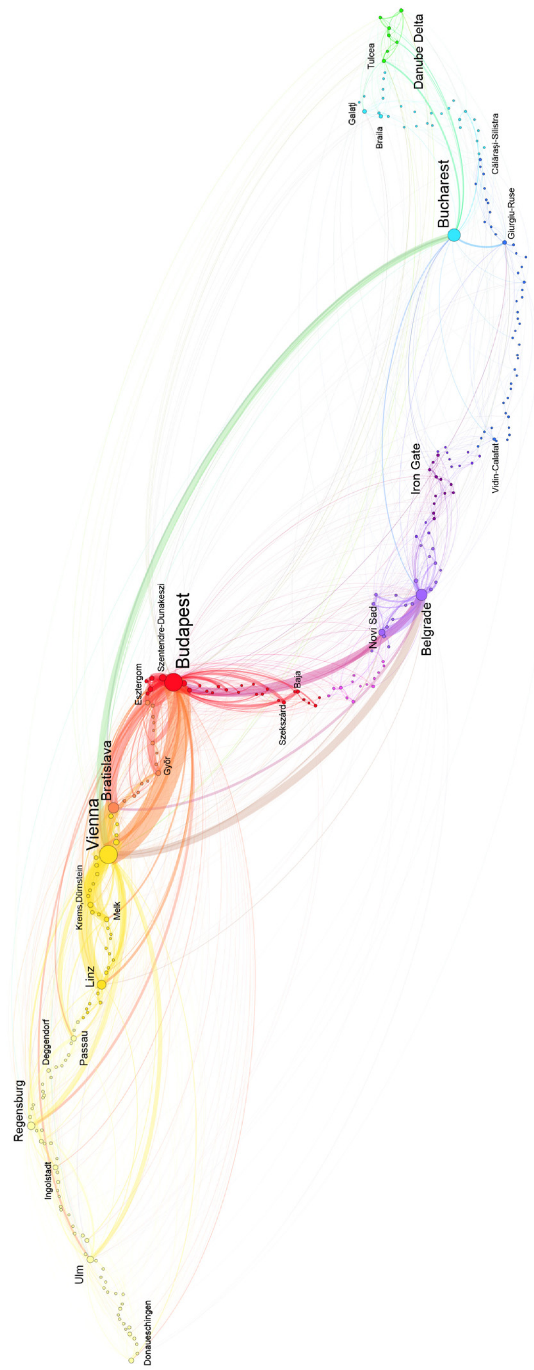

Fig. 6. Visitor flows between destinations along the Danube; colours indicate the different countries, thickness of connections is proportional with the number of users moving between the places (complete graph). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

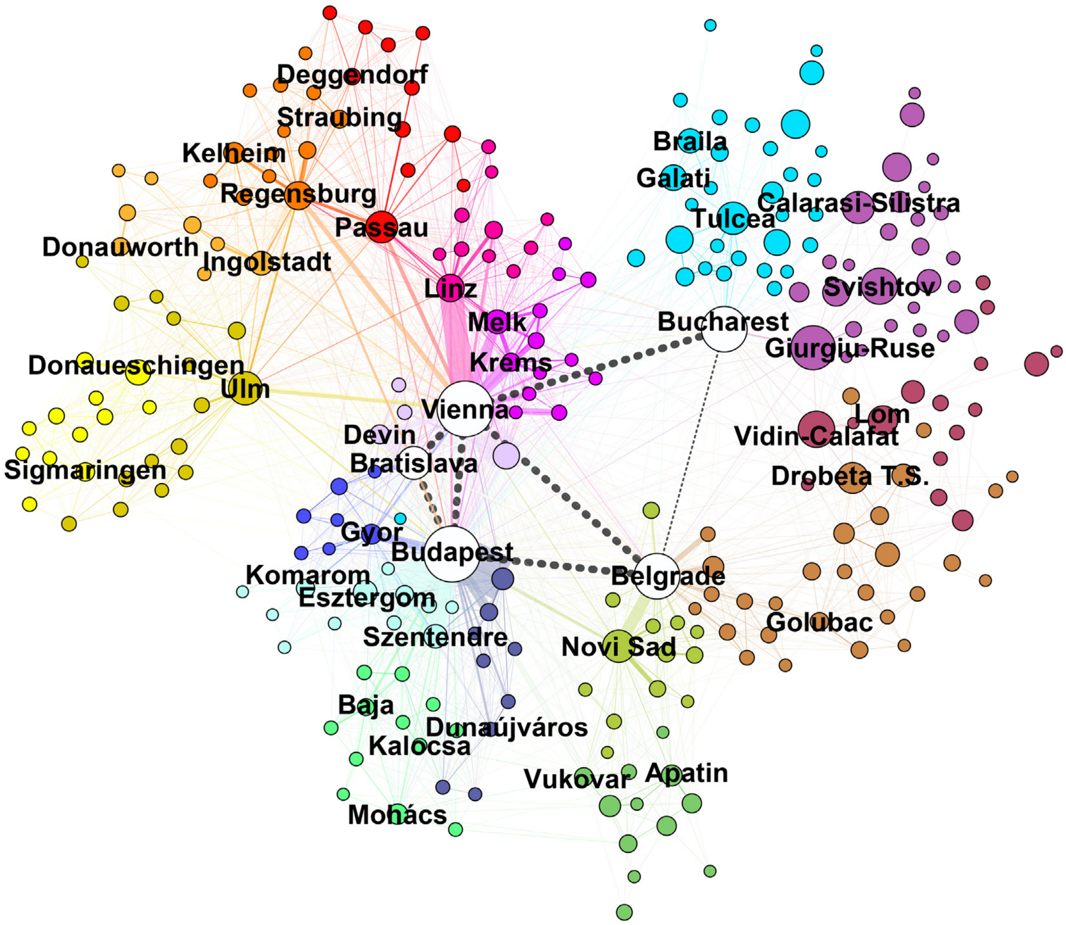

Fig. 7. Network analysis of the visitor flows along the Danube done with the graph layout algorithm Force Atlas 2. The size of the nodes is proportional to their betweenness centrality in the network. Connections between capital cities are much stronger than other edges, these are shown with dotted lines.

degree of 13.067 and a weighted degree of 186.126. National borders are also indicated, showing how such boundaries divide the network in clusters. The level of connectivity between the 18 strongly connected clusters was defined with the implementation of a third step of analysis.

These regional tourism clusters (without capital cities) were united into a new, much simpler graph where the 18 clusters became the nodes and the sum of edge weights connecting all of the original sections resulted in the weight of the new edges (Fig. 9). The analysis of this weighted graph gave a clear overview of the interconnectedness of the interregional system of tourism along the Danube. This simplified network has an average degree of 11.556 and a weighted degree of 659.556, very similar to the original extended networks. The 18 nodes form 3 separated communities with $Q > 0.721$, 4 with $Q > 0.232$ and the maximum number of clusters is 5 (above that $Q < 0$). These clusters clearly separate along national borders. There is a clear separation into 3 (Germany-Austria, Hungary-Slovakia, Lower-Danube), but even by clustering into 5 communities the new separations appear at the Austrian-German border and at the Serbian-Romanian-Bulgarian border (Fig. 9). The weight of connections in the Romanian-Bulgarian section are definitely lower than in the Upper-Danube regions, nodes here are completely cut-off from the system (the 5th cluster cannot be called an integrated tourism region). The resulting destination system visualized in Fig. 9 is the condensed synthesis of all visitor flows along the Danube.

Discussion and conclusions

In this paper we aimed to define the levels of interconnectedness in large-scale destination systems from the demand side of tourism and the boundary-shielding effect in such system. For this purpose we proposed a method based on the network visualization and cluster analysis of the weighted graphs of all movements within the different geographic points inside such destinations. We measured and visualized tourist consumption for the large-scale linear system of the Danube based on the analysis of the geotags of more than two million images uploaded to Flickr.com. All methods used in this paper have become emerging tools for tourism research: User Generated Content (UGC) was used successfully to measure tourism flows (Chua et al., 2016; Girardin et al., 2008; Kádár, 2014; Önder et al., 2014; Paldino et al., 2015); network analysis of such flows helped to define destination

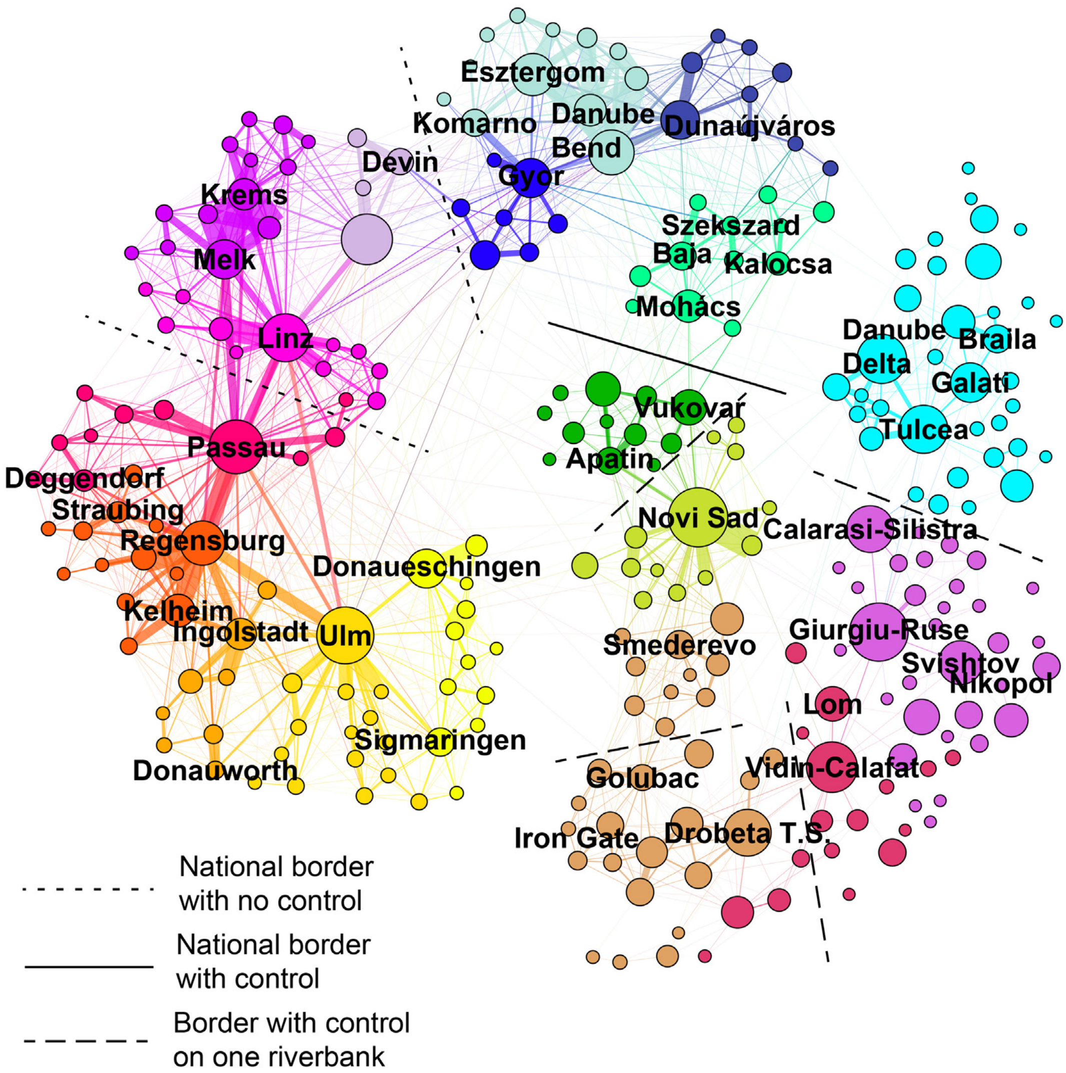

Fig. 8. Network analysis of the visitor flows along the Danube without the five capital cities. The size of the nodes is proportional to their betweenness centrality in the network, their colour represents the 18 destination clusters calculated by the algorithm. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

systems (Hwang et al., 2006; Lee et al., 2013; Shih, 2006; Taczanowska et al., 2014; Zach & Gretzel, 2012); clustering algorithms made visible the different destination clusters (Asero et al., 2016; Baggio & Scaglione, 2017; Grinberger et al., 2014). Only UGC could provide tourism data in such a large-scale interregional system, and only network analysis methods could process such large datasets. The results complement in a larger scale previous studies on interregional cross-boundary tourism systems (Blasco et al., 2014; Peng et al., 2016).

In our case the first research question aimed to define the level of integration between the points visited in a tourist system in order to be able to call it a large-scale destination. We proposed a multiple-step clustering method analysing weighted networks of visitor flows based on Flickr data that demonstrated how the Danube is not an integrated destination, it falls into 3 separate destination systems divided by strong national borders. The Upper-Danube forms an integrated tourism system, connecting all sections of the Danube in Austria and Germany. The Hungarian section of the Danube together with the section where the Danube borders Slovakia and Hungary form another well integrated system, in these sections the regional destinations are connected not only along the river, also the non-subsequent nodes have strong connections. A third system is the Lower-Danube, where a less visited, but linearly integrated Serbian section of the Danube is loosely connected to the underdeveloped sections of Bulgaria and Romania, where only the Delta has sections where travel patterns along the Danube are to be considered relevant.

The second research question aimed to define the role of national boundaries in the integration of large-scale transnational tourism destinations. This study reinforces the observations of Peng et al. (2016) stating that cross-boundary tourist flows are significantly influenced by the boundary-shielding effect. Both the analysis of connections between all 10 km long sections and the

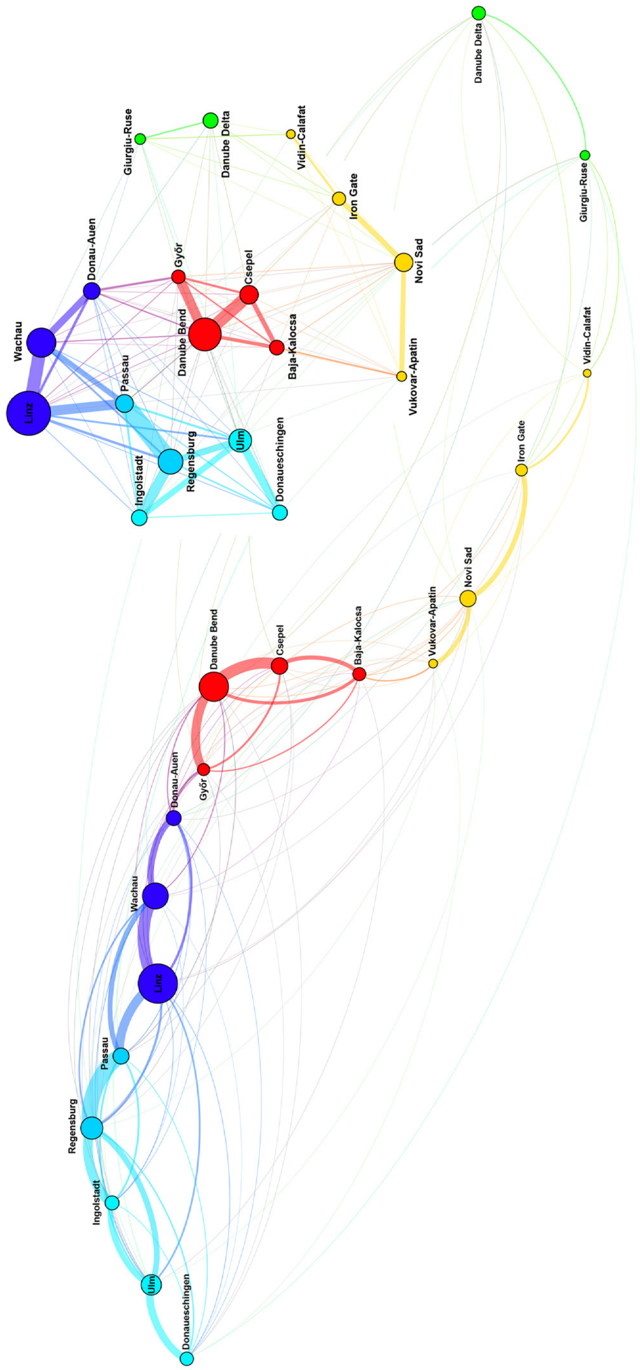

Fig. 9. Visitor flows along the Danube between the 18 identified tourism clusters without the five capital cities in a Geo Layout network visualization and a Force-directed network from Force Atlas 2 algorithm visualization (overlaid graph at top-right). Colours represent the identified cluster groups between tourism clusters. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

analysis of connections between the final 18 tourism clusters showed that transnational connections were much weaker than intra-national ones, and that clusters have clear limits at national borders. Even in the Upper-Danube, where tourists on a cross-border bicycle tour does not have to face any controls between Austria and Germany, there is a significant decrease in connectivity below the border city of Passau (361 connections to Linz region as opposed to the 474 connections between Linz and Wachau or the 642 connections between Passau and Regensburg region). The boundary-shielding effect is even greater at the border between Austria and Slovakia-Hungary, where only 72 connections were measured compared to the subsequent 254 above and 360 below. Quite drastic is the effect of the Schengen and EU borders between Hungary and Serbia-Croatia: 49 connections compared to the 163 above and 175 below. It must be also noted that none of the most photographed destinations stand in cross-border regions. The Danube as a national border with border controls makes the Croatian Vukovar-Apatin region along the Serbian border less attractive, Vidin-Calafat or Ruse-Giurgiu between Romania and Bulgaria almost unmeasurable as destinations, and the Iron Gate region between Romania and Serbia, the most scenic section of the whole river much less visited than regions like the Hungarian Danube Bend, or Wachau in Austria. The results show how “thick” borders impede more the integration of the tourism system than “thin” borders in the Schengen area (Haselsberger, 2014), also questioning the effectiveness of past interregional initiatives of the EU, underlining the observations of O'Dowd (2001). Programmes related to the current EU Strategy for the Danube Region (EUSDR) aim to make the Danube a connective element for Central-Eastern Europe (Busek & Gjoreska, 2010; Talabos, 2014). The underdevelopment of the destinations on the Lower-Danube are well known to local DMOs, but this fact makes them relate their tourism development work even more to the Danube as an integrated tourism brand (Damir, 2012; Matei et al., 2009; Mazilu, 2011). This paper reinforces previous case studies stating that the most important places of intervention for an integrated destination management are the border regions (Bjeljac & Curcic, 2006; Rădoi, 2020; Radu et al., 2010), however, single cross-border tourism projects are not enough to guarantee a destination-wide impact in this scale (Stoffelen & Vanneste, 2017).

Tourism has a great importance in the process of EU integration also because the Danube region suffered from many ethnic and political-historical conflicts in the past (Wingfield, 2003), and literature suggests that live tourism connections help in keeping peace among these nationalities (Becken & Carmignani, 2016; Farmaki, 2017). Historical and ethnic conflicts are probably having a negative effects on the willingness to travel to neighbouring countries also in this region (Khalilzadeh, 2018), while such conditions and the out-of-the-ordinary nature of less visited regions are effecting the destination choice of tourists from other regions as well (Karl, 2018).

The geographic approach of this study couldn't reveal all aspects of a complex tourism destination, as nor the motivations or neither the composition of the demand side were described, and none of the supply side was part of the research. In future research the SNA of the supply side could complement effectively current findings, adding also new tools of network analysis as the connections of stakeholders along the Danube can be analysed without the geographic constraints as opposed to travel patterns (Baggio, 2020). Also, qualitative analyses of travellers' and stakeholders' sides could add to the full understanding of such complex systems.

The spatio-temporal methodology presented is reproducible for any large-scale tourism system, and answers clearly whether destinations form integrated tourism systems at the demand side or not. However, in the case studied we also found some limitations for the methodology. First, no data exists on the penetration of Flickr usage in different countries in Central-Eastern Europe. More work would be needed to see how large portion of local visitors used this service in different countries. Quantification of Flickr data delivers results correlating well to the tourism flows of international visitors from mainly western countries, but local flows of Romanian and Bulgarian tourists might be less represented. The authors made official study trips to all sections of the Danube to verify that it is indeed realistic to see so much lower visitor numbers in Lower-Danube regions respect to destinations up the river, however, other methodologies should also prove the robustness of this method. The second limitation of the study comes from the compromising effect of the four capital cities on the Danube. These cities are the major enter and exit points to the Danube's regional destinations, but they have much higher tourism flows unrelated to the overall system. For this reason, we excluded them completely from the cluster analysis, but this also means we could not count them as intermediate stops in an itinerary that goes along the river. Capital cities make the system corrupt also because of the large-scale visitor movements into destinations of the agglomeration. Finally, we found just enough visitor connections between destination clusters upstream and downstream of these capitals to make the three destination systems involved uninterrupted by the absence of these cities, therefore the exclusion of the capitals from the destinations systems resulted to be a valid method to map this large-scale linear tourism destination system.

Declaration of competing interest

None.

Acknowledgements

This research was completed under the DANUrB interreg project of the EU Danube Transnational Programme co-funded by European Union funds (ERDF and IPA), reference number DTP1-1-249-2.2.

References

Ahas, R., Aasa, A., Mark, Ü., Pae, T., & Kull, A. (2007). Seasonal tourism spaces in Estonia: Case study with mobile positioning data. *Tourism Management*, *28*, 898–910.
Amaro, S., Antunes, A., & Henriques, C. (2018). A closer look at Santiago de Compostela's pilgrims through the lens of motivations. *Tourism Management*, *64*, 271–280.

Asero, V., Gozzo, S., & Tomaselli, V. (2016). Building tourism networks through tourist mobility. *Journal of Travel Research*, *55*(6), 751–763.

Baggio, R. (2011). Collaboration and cooperation in a tourism destination: A network science approach. *Current Issues in Tourism*, *14*(2), 183–189.

Baggio, R. (2020). Tourism destinations: A universality conjecture based on network science. *Annals of Tourism Research*, 82.

Baggio, R., & Sainaghi, R. (2011). Complex and chaotic tourism systems: Towards a quantitative approach. *International Journal of Contemporary Hospitality Management*, *23*(6), 840–861.

Baggio, R., & Scaglione, M. (2017). Strategic Visitor Flows (SVF) analysis using mobile data. In R. Schegg, & B. Stangl (Eds.), *Information and communication technologies in tourism 2017* (pp. 145–157). Springer.

Baggio, R., Scott, N., & Cooper, C. (2010). Network science and socio-economic systems: A review focused on a tourism destination. *Annals of Tourism Research*, *37*(3), 802–827.

Balli, F., Curry, J., & Balli, H. O. (2015). Inter-regional spillover effects in New Zealand international tourism demand. *Tourism Geographies*, *17*(2), 262–278.

Bangwayo-Skeete, P. F., & Skeete, R. W. (2015). Can Google data improve the forecasting performance of tourist arrivals? Mixed-data sampling approach. *Tourism Management*, *46*, 454–464.

Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: An open source software for exploring and manipulating networks. *Third international AAAI conference on weblogs and social media* (pp. 361–362).

Becken, S., & Carmignani, F. (2016). Does tourism lead to peace? *Annals of Tourism Research*, *61*, 63–79.

Beritelli, P., Bieger, T., & Laesser, C. (2014). The new frontiers of destination management: Applying variable geometry as a function-based approach. *Journal of Travel Research*, *53*(4), 403–417.

Birenboim, A., & Shoval, N. (2016). Mobility research in the age of the smartphone. *Annals of the American Association of Geographers*, *106*(2), 283–291.

Bjeljac, Z., & Curcic, N. (2006). Tourism in the Serbian, Romanian and Hungarian borderline area as part of cross-border cooperation. *Geographica Pannonica*, *10*, 73–77.

Blasco, D., Guia, J., & Prats, L. (2014). Tourism destination zoning in mountain regions: A consumer-based approach. *Tourism Geographies*, *16*(3), 512–528.

Blondel, V. D., Guillaume, J. -L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2008*(10), 1–12.

Boengiu, V. (2012). Evaluation of tourism resources in the Iron Gates Natural Park in order to identify the potential of tourism development. *Analele Universităţii Din Oradea, Seria Geografie*, *22*(2).

Burt, R. (1992). *Structural holes: The social structure of competition.* Harvard University Press.

Busek, E., & Gjoreska, A. (2010). The Danube region: Transformation and emergence. *Eastern Journal of European Studies*, *1*(1), 9–20.

Cai, G., Hio, C., Bermingham, L., Lee, K., & Lee, I. (2014). Mining frequent trajectory patterns and regions-of-interest from flickr photos. *Proceedings of the annual Hawaii international conference on system sciences* (pp. 1454–1463).

Caton, K., & Santos, C. A. (2007). Heritage tourism on route 66: Deconstructing nostalgia. *Journal of Travel Research*, *45*(4), 371–386.

Chua, A., Servillo, L., Marcheggiani, E., & Moere, A. V. (2016). Mapping Cilento: Using geotagged social media data to characterize tourist flows in southern Italy. *Tourism Management*, *57*, 295–310.

Connell, J., & Page, S. J. (2008). Exploring the spatial patterns of car-based tourist travel in Loch Lomond and Trossachs National Park, Scotland. *Tourism Management*, *29*(3), 561–580.

Damian, N., & Dumitrescu, B. (2009). Sustainable development prospects for the Danube delta rural communities. *Revue Roumaine de Geographie/Romanian Journal of Geography*, *53*(2), 153–163.

Damir, D. (2012). The importance of the Danube strategy for tourism and culture development of the Croatian Danube region. *Geographica Pannonica*, *16*(3), 112–125.

De Frantz, M. (2018). Tourism marketing and urban politics: Cultural planning in a European capital. *Tourism Geographies*, *20*(3), 481–503.

Dragin, A. S., Dragin, V., Plavša, J., Dragin, A. S., & Plavša, J. (2007). Cruise ship tourism on the Danube in Vojvodina Province as a segment of global tourism. *Geographica Pannonica*, *11*, 59–64.

Dragin, A. S., Jovicic, D., & Boškovic, D. (2010). Economic impact of cruise tourism along the Paneuropean corridor VII. *Ekonomska Istrazivanja*, *23*(4), 127–141.

Encalada, L., Boavida-Portugal, I., Cardoso Ferreira, C., & Rocha, J. (2017). Identifying tourist places of interest based on digital imprints: Towards a sustainable smart city. *Sustainability*, *9*(12), 2317.

Farmaki, A. (2017). The tourism and peace nexus. *Tourism Management*, *59*, 528–540.

Ferreira, S. L. A., & Hunter, C. A. (2017). Wine tourism development in South Africa: a geographical analysis. *Tourism Geographies*, *19*(5), 676–698.

Framke, W. (2002). The destination as a concept: A discussion of the business-related perspective versus the socio-cultural approach in tourism theory. *Scandinavian Journal of Hospitality and Tourism*, *2*(2), 92–108.

Girardin, F., Fiore, F. D., Blat, J., Ratti, C., & Dal Fiore, F. (2008). Understanding of tourist dynamics from explicitly disclosed location information. *Journal of Location Based Services*, *2*(1).

González-Díaz, B., Gómez, M., & Molina, A. (2015). Configuration of the hotel and non-hotel accommodations: An empirical approach using network analysis. *International Journal of Hospitality Management*, *48*, 39–51.

Grinberger, A. Y., Shoval, N., & McKercher, B. (2014). Typologies of tourists' time–space consumption: A new approach using GPS data and GIS tools. *Tourism Geographies*, *16*(1), 105–123.

Hall, D. R. (1993). Ecotourism in the Danube Delta. *Revue de Tourisme - The Tourist Review*, *48*(3), 11–13.

Haselsberger, B. (2014). Decoding borders. Appreciating border impacts on space and people. *Planning theory and practice. Vol. 15, Issue 4*. (pp. 505–526). Taylor & Francis.

Hwang, Y. H., Gretzel, U., & Fesenmaier, D. R. (2006). Multicity trip patterns. Tourists to the United States. *Annals of Tourism Research*, *33*(4), 1057–1078.

Iordanova, E. (2017). Tourism destination image as an antecedent of destination loyalty: The case of Linz, Austria. *European Journal of Tourism Research*, *16*, 214–432.

Jacomy, M., Heymann, S., Venturini, T., & Bastian, M. (2011). *Force Atlas 2, a graph layout algorithm for handy network visualizationhttp://www.Medialab.Sciences-Po.Fr/, 1–21.*

Joksimović, M., Golić, R., Vujadinović, S., Šabić, D., Jovanović Popović, D., & Barnfield, G. (2014). Restoring tourist flows and regenerating city's image: The case of Belgrade. *Current Issues in Tourism*, *17*(3), 220–233.

Jovicic, D. Z. (2016). Key issues in the conceptualization of tourism destinations. *Tourism Geographies*, *18*(4), 445–457.

Kádár, B. (2013). Differences in the spatial patterns of urban tourism in Vienna and Prague. *Urbani Izziv*, *24*(2), 96–111.

Kádár, B. (2014). Measuring tourist activities in cities using geotagged photography. *Tourism Geographies*, *16*(1), 88–104.

Kádár, B. (2018). Hotel development through centralized to liberalized planning procedures: Prague lost in transition. *Tourism Geographies*, *20*(3), 461–480.

Kádár, B., & Gede, M. (2013). Where do tourists go? Visualizing and analysing the spatial distribution of geotagged photography. *Cartographica: The International Journal for Geographic Information and Geovisualization*, *48*(2), 78–88.

Kádár, B., & Vitkova, L. (2019). Sustainability options for tourism development. In M. Benkő, P. Gregor, B. Kádár, & L. Vitkova (Eds.), *Book on unexplored cultural heritage in communities by the Danube* (pp. 140). Gasset.

Karl, M. (2018). Risk and uncertainty in travel decision-making: Tourist and destination perspective. *Journal of Travel Research*, *57*(1), 129–146.

Kasagranda, A., Rajčáková, E., & Vystoupil, J. (2016). Urban tourism in Slovakia - Its quantification, spatial differentiation and typification. *Geographica Pannonica*, *20*(2), 105–113.

Khalilzadeh, J. (2018). Demonstration of exponential random graph models in tourism studies: Is tourism a means of global peace or the bottom line? *Annals of Tourism Research*, *69*(October 2017), 31–41.

Lagiewski, R., & Revelas, D. (2004). Challenges in cross-border tourism regions. *RIT scholar works*.

Lee, S. -H., Choi, J. -Y., Yoo, S. -H., & Oh, Y. -G. (2013). Evaluating spatial centrality for integrated tourism management in rural areas using GIS and network analysis. *Tourism Management*, *34*, 14–24.

Lew, A. A., & McKercher, B. (2006). Modeling tourist movements: A local destination analysis. *Annals of Tourism Research*, *33*(2), 403–423.

Li, J., Xu, L., Tang, L., Wang, S., & Li, L. (2018). Big data in tourism research: A literature review. *Tourism Management*, *68*, 301–323.

Li, Y., Xiao, L., Ye, Y., Xu, W., & Law, A. (2016). Understanding tourist space at a historic site through space syntax analysis: The case of Gulangyu, China. *Tourism Management*, *52*, 30–43.
Liu, B., Huang, S.(. S.)., & Fu, H. (2017). An application of network analysis on tourist attractions: The case of Xinjiang, China. *Tourism Management*, *58*, 132–141.
Liu, Q., Wang, Z., & Ye, X. (2018). *Comparing mobility patterns between residents and visitors using geo-tagged social media dataJuly. (pp. 1–18), 1–18.*
Lo, I. S., McKercher, B., Lo, A., Cheung, C., & Law, R. (2011). Tourism and online photography. *Tourism Management*, *32*(4), 725–731.
Lois-González, R. C., & Santos, X. M. (2015). Tourists and pilgrims on their way to Santiago. Motives, Caminos and final destinations. *Journal of Tourism and Cultural Change*, *13*(2), 149–164.
Lorde, T., Li, G., & Airey, D. (2016). Modeling Caribbean tourism demand: An augmented gravity approach. *Journal of Travel Research*, *55*(7), 946–956.
Lozano, S., & Gutiérrez, E. (2018). A complex network analysis of global tourism flows. *International Journal of Tourism Research*, *20*(5), 588–604.
Lue, C., Crompton, J. L., & Fesenmaier, D. R. (1993). Conceptualization of multi-destination pleasure trips. *Annals of Tourism Research*, *20*(2), 289–301.
Lyócsa, Š., Vašaničová, P., & Litavcová, E. (2019). Interconnectedness of international tourism demand in Europe: A cross-quantilogram network approach. *Physica A*, *526*, 120919.
MacCannell, D. (1976). *The tourist.* Schocken Books.
Mariani, M., & Baggio, R. (2020). The relevance of mixed methods for network analysis in tourism and hospitality research. *International Journal of Contemporary Hospitality Management*, *32*(4), 1643–1673.
Matei, E., Stancioiu, A. F., Pargaru, I., Gebriela, M., & Vlădoi, A. -D. (2009). The Romanian ports on the Danube Valley - An emergent tourism destination. *Recent Researches in Tourism and Economic Development*, 113–118.
Mazilu, M. (2011). Regional tourism from the perspective of the Danube strategy – Iron gates region. *Cactus Toruism Journal*, *2*(2), 44–55.
McKercher, B., & Lew, A. A. (2004). Tourist flows, itineraries and factors affecting the spatial distribution of tourists. In A. A. Lew, C. M. Hall, & A. M. Williams (Eds.), *A tourism companion to geography* (pp. 36–48). Blackwell.
McLeod, M., Lewis, E. H., & Spencer, A. (2017). Re-inventing, revolutionizing and transforming Caribbean tourism: Multi-country regional institutions and a research agenda. *Journal of Destination Marketing and Management*, *6*(1), 1–4.
Meschik, M. (2012). Sustainable cycle tourism along the Danube cycle route in Austria. *Tourism Planning and Development*, *9*(1), 41–56.
Nicolosi, A., Laganà, V., Cortese, L., & Privitera, D. (2018). Using the network and MCA on tourist attractions. The case of Aeolian Islands, Italy. *Sustainability*, *10*(11), 4169.
Novotna, M. (2018). Tourist flows between central European metropolises (in the context of metropolisation processes). *Geographia Technica*, *13*(2), 125–137.
O'Dowd, L. (2001). Analysing Europe's borders. *IBRU Boundary and Security Bulletin*, *9*(2), 67–79.
Önder, I., Koerbitz, W., & Hubmann-Haidvogel, A. (2014). Tracing tourists by their digital footprints: The case of Austria. *Journal of Travel Research*, *55*(5), 566–573.
Oppermann, M. (1996). Rural tourism in southern Germany. *Annals of Tourism Research*, *23*(1), 86–102.
Orsi, F., & Geneletti, D. (2013). Using geotagged photographs and GIS analysis to estimate visitor flows in natural areas. *Journal for Nature Conservation*, *21*(5), 359–368.
Paldino, S., Bojic, I., Sobolevsky, S., Ratti, C., & González, M. C. (2015). Urban magnetism through the lens of geo-tagged photography. *EPJ Data Science*, *4*(1), 1–17.
Pearce, D. G. (2014). Toward an integrative conceptual framework of destinations. *Journal of Travel Research*, *53*(2), 141–153.
Pearce, D. G., & Heike, S. A. (2015). Destinations: Tourists' perspectives from New Zealand. *International Journal of Tourism Research*, *17*, 4–12.
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.
Ploner, J. (2009). Narrating regional identity in tourism - Sketches from the Austrian Danube valley. *Language and Intercultural Communication*, *9*(1), 2–14.
Puczkó, L., Ratz, T., & Smith, M. (2007). Old city, new image: Perception, positioning and promotion of Budapest. *Journal of Travel & Tourism Marketing*, *22*(3–4), 21–34.
Rădoi, I. (2020). European capital of culture, urban tourism and cross-border cooperation between Romania and Serbia. *Journal of Balkan and Near Eastern Studies*, *22*(4), 547–559.
Radu, S. A., Bianca, D. A., & Nicoleta, D. A. (2010). Cross-border cooperation in the Danube-lined Romanian/Bulgarian border-space. Geographical considerations. *Geographica Pannonica*, *14*(2), 67–75.
Rátz, T., Smith, M., & Michalkó, G. (2008). New places in old spaces: Mapping tourism and regeneration in Budapest. *Tourism Geographies*, *10*(4), 429–451.
Scott, J., & Carrington, P. J. (2011). The SAGE handbook of social network analysis. *The SAGE handbook of social network analysis*. SAGE publications Ltd.
Scott, N., Cooper, C., & Baggio, R. (2008). Destination networks. Four Australian cases. *Annals of Tourism Research*, *35*(1), 169–188.
Shih, H. Y. (2006). Network characteristics of drive tourism destinations: An application of network analysis in tourism. *Tourism Management*, *27*(5), 1029–1039.
Shoval, N., & Ahas, R. (2016). The use of tracking technologies in tourism research: The first decade. *Tourism Geographies*, *18*(5), 587–606.
Shoval, N., McKercher, B., Birenboim, A., & Ng, E. (2015). The application of a sequence alignment method to the creation of typologies of tourist activity in time and space. *Environment and Planning B: Planning and Design*, *42*(1), 76–94.
Smallwood, C. B., Beckley, L. E., & Moore, S. A. (2012). An analysis of visitor movement patterns using travel networks in a large marine park, north-western Australia. *Tourism Management*, *33*(3), 517–528.
Smith, M. K., Egedy, T., Csizmady, A., Jancsik, A., Olt, G., & Michalkó, G. (2018). Non-planning and tourism consumption in Budapest's inner city. *Tourism Geographies*, *20*(3), 524–548.
Steinbach, J. (1995). River related tourism in Europe - An overview. *GeoJournal*, *35*(4), 443–458.
Stoffelen, A., & Vanneste, D. (2017). Tourism and cross-border regional development: Insights in European contexts. *European Planning Studies*, *25*(6), 1013–1033.
Straumann, R. K., Çöltekin, A., & Andrienko, G. (2014). Towards (re)constructing narratives from georeferenced photographs through visual analytics. *The Cartographic Journal*, *51*(2), 152–165.
Taczanowska, K., González, L. M., Garcia-Massó, X., Muhar, A., Brandenburg, C., & Toca-Herrera, J. L. (2014). Evaluating the structure and use of hiking trails in recreational areas using a mixed GPS tracking and graph theory approach. *Applied Geography*, *55*, 184–192.
Talabos, I. (2014). Customer habits and the Datourway strategy. *Turističko Poslovanje*, *14*(December), 71–80.
Terzić, A., & Bjeljac, Ž. (2016). Cultural routes - Cross-border tourist destinations within southeastern Europe. *Forum Geografic*, *15*(2), 180–188.
TourMIS (2018). *City tourism in Europe.* (accessed 15 May 2019).
Valeri, M., & Baggio, R. (2020). Italian tourism intermediaries: a social network analysis exploration. *Current Issues in Tourism*, *0*(0), 1–14.
Vujko, A., Plavša, J., & Ostojić, N. (2013). Impact of the "Danube cycling route" on the development of cycling tourism in Serbia. *Polish Journal of Sport and Tourism*, *20*(3), 227–233.
Widawski, K., & Wyrzykowski, J. (2017). *The geography of tourism of central and eastern European countries.* Springer.
Williams, N. L., & Hristov, D. (2018). An examination of DMO network identity using exponential random graph models. *Tourism Management*, *68*, 177–186.
Wingfield, N. M. (Ed.). (2003). *Creating the other: Ethnic conflict & nationalism in Habsburg Central Europe (Austrian h)*. Berghahn Books.
Zach, F., & Gretzel, U. (2012). Tourist-activated networks: Implications for dynamic bundling and EN route recommendations. *Journal of Information Technology and Tourism*, *13*(3), 239–257.

Bálint Kádár is an architect, urban scholar and planner, associate professor at the department of Urban Planning and Design at the Budapest University of Technology and Economics (Műegyetem rkp. 3, 1111. Budapest, HU, Email: kadarb@urb.bme.hu), researching and planning urban tourism development.
Mátyás Gede is associate professor at the Department of Cartography and Geoinformatics at the Eötvös Loránd University. His main fields of interest are: globe and relief model digitizing, digital globes, webcartography and map projections.