Classifying multi-destination trips in Austria with big data

Irem Önder, Dr.

MODUL University Vienna, Department of Tourism and Service Management, Am Kahlenberg 1, A-1190 Vienna, Austria

ABSTRACT

Multi-destination trips are interesting for research in order to see which destinations are combined into one leisure trip. The aim of this study is to classify multi-destination trips in Austria based on geotagged photos on Flickr. The study sample includes tourists in Austria who visited at least two different cities based on the geolocations of their photos. The results revealed three types of multi-destination trips: (1) single destination trips (57%); (2) base camp trips (30%); and (3) regional tour trips (13%). Furthermore, cluster analysis was conducted to categorize the cities. The first cluster covers the eastern part of the country, which includes larger cities such as Vienna and Graz, and the second cluster refers to the western part of Austria. Practical implications include creating joint marketing campaigns and new tourism products such as hiking trails between cities in the same cluster.

1. Introduction

Planning a trip involves many decisions such as where to go, how to get there and what to do at the destination. Among all these decisions, destination choice is one of the first to be decided by travellers. According to Crompton (1992), selection of a vacation destination occurs in three stages. The first stage is having an awareness set, which comprises all of the destinations a person knows. In the second stage, a consideration/evoked/relevant set is chosen from the awareness set, and represents the destinations the person is able to visit based on their circumstances. The final stage involves the selection of destinations that are deemed worth visiting and getting information about them, resulting in one destination being selected as the chosen vacation destination (Crompton, 1992). However, previous research has shown that 30% to 50% of all pleasure trips are multi-destination trips (Hanson, 1980; O’Kelly, 1982). These multi-destination trips can take various forms, such as visiting towns that are on the way to the main destination, or visiting nearby regions during the stay at the main destination.

Multi-destination trips are interesting for research in order to see which destinations are combined together in one leisure trip. Some examples of previous research concerning travel patterns and spatial movement of travellers include Lue, Crompton, and Fesenmaier (1993), Stewart and Vogt (1997), Tideswell and Faulkner (1999), and Hwang and Fesenmaier (2003). Knowing which destinations are visited during a trip is invaluable information for destination marketing organizations (DMOs), whether a city tourism organization or a regional tourist office. This knowledge enables the identification of potential marketing synergies, especially for destinations that do not have enough attractions to draw visitors on their own (Tideswell & Faulkner, 1999). For instance, destinations can pursue joint marketing campaigns by creating hiking trails spanning two or more destinations, or advertise on each other’s destination websites.

The purpose of this study is identifying and classifying multi-destination trips in Austria based on data retrieved from Flickr. Surveys are conducted by destinations to collect data from travellers; however, there is no data collected in Austria that shows which combination of destinations tourists visit if they visit multiple destinations in the country during their stay. This data can be retrieved from Flickr by using the geolocations of the photos taken. Using this information, multi-destination trips were identified, multi-destination tourists were categorized, and destinations were clustered according to the trip patterns. This study builds on previous research by Önder, Koebritz, and Hubmann-Heidvogel (2016) by extending the focus to multi-destination trips.

2. Literature review

2.1. Multi-destination trips

Multi-destination trips are rational for tourists because they are cost and time effective. However, these are not among the five main reasons for engaging in multi-destination trips as identified by Lue et al. (1993). First, multi-destination trips better enable the satisfaction of heterogeneous travel preferences within travel groups even as small as two people. For instance, when one person from a given travelling group wants to climb a mountain while another wants to relax on the beach, a single destination may not be able to fulfil both of these needs. Second, the fact that nearly half of all US tourists stay with friends and relatives during their trip may lead to additional trips around the primary travel destination. Third, considering a multi- rather than a single-destination trip
may increase the travel needs of potential travellers and encourage them to seek more variety in their travel. Fourth, combining multiple attractions or destinations into a single trip may increase travel satisfaction by diversifying the experiences and thereby reduce the level of risk. Fifth, spatial, temporal, and personal constraints in leisure travel may result in different destinations being part of one trip in order to satisfy the multiple needs of travellers. In sum, the many different needs of individual travellers and the varying forms of travel groups often make multi-destination trips a logical choice.

According to the LCF model by Lue et al. (1993), there are five multi-destination travel patterns: (1) single destination, in which only one destination is visited; (2) en route travel, includes trips on the way to the main destination or on the way back home, excluding side trips; (3) base camp trip, involves side trips around the main destination; (4) regional tour, includes trips in a region with stops in the smaller destinations in the area; and (5) trip chaining pattern, includes multiple destinations in a region and between regions (Lue et al., 1993).

Stewart and Vogt (1997) utilized the multi-destination traveller categorization by Lue et al. (1993) in order to identify multi-destination trip patterns in Branson, USA, and their results indicate a slightly different traveller classification than the LCF model. They differentiated trip chaining from regional tours, in that trip chaining describes more extensive tours which include different regions in North America, whereas regional tours capture multi-destination travel within smaller regions.

Investigating domestic travel patterns in the USA, Hwang and Fesenmaier (2003) find that multi-destination trips can be categorized as en route travel, in which individuals stop over at other destinations on the way to their main destination, rather than taking sub-trips around the main destination. The type of destination, whether en route or base camp, can also influence the length of stay, while prior experience at the main destination can influence the bundling of additional destinations (Hwang & Fesenmaier, 2003).

Overall, multi-destination trips are driven by cumulative attraction, which states that “a given number of attractions whose primary target is tourists will do more business if they are located en route, in proximity, or in a logical sequence to each other than if they are widely scattered” (Lue et al., 1993, p. 297). In the same line, if one destination can identify the other destinations that are close by and have something to offer to tourists, these two destinations can be combined in a multi-destination trip. Such synergies can result in economic benefits for both locations and may also increase the duration of stay in the area. Moreover, cumulative attraction indicates that tourism business is shared. According to Lue et al. (1993, p. 297) “an attraction secures its visitation not only as a result of its own generative power, but also as a result of the generative power of other attractions”. Thus, it is crucial for especially smaller or rural destinations to determine which other destinations they can partner with to increase demand for both destinations.

It is also important to know about multi-destination trips for the following reasons: (1) most destinations are not stand-alone cases, but form part of a product which includes the surrounding destinations; (2) identifying the destinations that are combined into single trips and learning the motives behind visitation of each destination enhance understanding of the destination and the surrounding destinations; (3) a good understanding of multi-destination trips improves the accuracy of tourism demand forecasts; (4) understanding the linkages between destinations can help inform joint marketing efforts by the destinations; (5) multi-destination models can help us to better understand the economic impact of tourism in the region or country (Lue et al., 1993).

Surveys are the most common method used to identify the destinations included in a multi-destination trip, despite being both time consuming and expensive. On the other hand, big data such as the traces individuals leave on the internet in the form of geotagged photos can show where individuals have been, which can be used to identify the destinations that individuals bundle during multi-destination trips.

2.2. Big data

Big data is described as “data sets and analytical techniques in applications that are so large (from terabytes to exabytes) and complex (from sensor to social media data) that they require advanced and unique data storage, management, analysis, and visualization technologies” (Chen, Chiang, & Storey, 2012, p. 1166). In order to call data big data, it needs to satisfy these three categories (McAfee & Brynjolfsson, 2012): volume, velocity, and variety. Volume refers to the quantity of data, velocity refers to the speed of information retrieval, and variety refers to different types of data such as text messages, photos on social networks, and GPS signals from mobile phones. Some examples of big data are credit card transactions, search engine trends (e.g. Google Trends), social media data from Facebook messages to twitter posts, and photos shared on social media such as on Flickr. As a data-driven methodology, big data is used to gain an understanding from the data which can then be used to enhance business intelligence. According to Dolnicar and Ring (2014), big data has the potential to change knowledge generation in terms of speed and quantity. Thus, big data and data-driven approaches have been applied in previous research across a variety of fields such as retail (Brown, Chui, & Manyika, 2011; Lee, Lee, & Sohn, 2013), healthcare (Brinkmann, Bower, Stengel, Worrell, & Stead, 2009), security and safety (Chen et al., 2012), education (Siemens & Long, 2011), government (Mervis, 2012), services (Acker, Gröne, Blockus, & Bange, 2011, Demirkan & Delen, 2013; Kauffman, Srivastava, & Vayghan, 2012), technology (Bradbury, 2011), and fraud detection (Abbasi, Albrecht, Vance, & Hansen, 2012).

Tourism research utilizes big data as well: for instance in research streams focused on forecasting tourism demand. Some examples include predicting hotel demand based on website traffic data (Yang, Pan, & Song, 2014); predicting tourism demand to Caribbean islands using Google Trends data (Bangwayo-Skeete & Skeete, 2015); and predicting actual tourist arrivals to Vienna by using Google Analytics website traffic indicators from the Viennese DMO website (Gunter & Önder, 2016). In all of the aforementioned studies, the use of big data is demonstrated to enhance forecasts. Another research stream focuses on the use of big data for recommendation systems. For instance, Fuchs, Hoepekken, and Lexhagen (2014) propose a framework for a destination management information system which combines big data with other traditional tourism data. Each of these studies utilizes a different form of big data, another manifestation of which – geotagged photos on the internet – has been the subject of tourism research in the past years.

2.3. Geotagged photos

Geotagged photos such as the ones uploaded on Flickr (www.flickr.com), are a type of big data which can be used to identify travel patterns. According to Lo, McKercher, Lo, Cheung, and Law (2011), 89% of Hong Kong residents who take leisure trips take photographs and 41% of those post them online. Moreover, 40% of these travellers use Flickr and other similar types of photo sharing websites (Lo et al., 2011). Flickr data can be used to identify the exact locations where tourists have been, including points of interests, and can reveal new trends among tourists such as visiting lesser known attractions. Analyzing Flickr data that covers a bigger region, such as a country, also gives the opportunity to see the domestic multi-destination travel patterns. For instance, Koerbiz and Önder (2014) show how geotagged photos can be used for destination benchmarking, and Vu, Leung, Rong, and Miao (2016) investigate park visitor behavior in Hong Kong based on geotagged photos.

Previous research has used Flickr to identify user movements for purposes such as creating a recommendation system which suggest places to visit for first time visitors to a destination based on their previous behavior (Mamei, Rosi, & Zambonelli, 2010); creating automated travel itineraries (De Choudhury et al., 2010), identifying places visited, duration of stay and panoramic spots of the destination (Popescu,
3. Methodology

3.1. Data collection

Flickr data is publicly available on the internet via Application Programming Interface (API). In order to collect the metadata of the photos on Flickr, an application was developed which retrieves the meta-data for the photos of a given destination taken within a specified time frame. The meta-data collected for this study includes the following: a) user defined textual information such as title, description and tags of the image; b) geographical information such as longitude, latitude and a plain-text name of the location, e.g. ‘Vienna/Vienna/Austria’; and c) date including the date when the photo was taken or uploaded. Moreover, user specific information such as name, current location, and current occupation were also retrieved. The longitude and latitude of the photo is defined in two different ways. The first geolocation is assigned automatically by Flickr when the user uploads the photo on a map. The second one comes directly from Flickr users’ cameras which feature built-in GPS systems and assign geolocations to photos automatically (Pereira, Vaccari, Giardin, Chiu, & Ratti, 2011).

The data collection was carried out between March and July of 2012. The data included photos tagged by Flickr users within Austria and all of its regions (Vienna, Burgenland, Carinthia, Styria, Upper Austria, Lower Austria, Salzburg, Tirol and Vorarlberg) which were taken between 01.01.2007 and 31.12.2011. The data collection included the name of the destinations both in German and in English (e.g. Wien and Vienna), which were retrieved based on the tags included in the photos on Flickr. The total number of photos collected for the study was 1,183,889.

3.2. Data cleaning

The photos collected may have been taken by residents of the destination or by tourists; thus the next step is to separate the two groups. Previous research used heuristics to categorize residents and tourists based on the time span between the geotagged photos. For instance, De Choudhury et al. (2010) used a 21 day time span between the first and the last photo taken and at least 2 points of interest visited in the same city as an indication of a tourist, whereas Girardin et al. (2008) opted for a 30 day time span to identify tourists. Since there is not a rigorous way to identify tourists, Girardin et al.’s (2008) procedure is followed in this study.

First, the individual users were identified, which totalled 27,901 individual users in the sample. Then for each individual, the dates on which their first and last online photos were taken were compared and those users for whom the time span was less than 30 days were classified as tourists. The final study sample includes 1,183,889 photos from Austria of which 883,465 photos were from residents and 300,424 photos were from tourists. The total number of tourists is 20,067. A modified version of the trip pattern classification by Lue et al. (1993) was used in this study to classify the tourists based on their trip patterns. To be considered in a multi-destination trip in this study, each tourist must have stayed at least one night in two different destinations in the study sample. The trips are classified as follows: (1) single destination trips include tourists who visit one destination and spend multiple nights; (2) base camp trips include tourists who spend multiple nights at one destination and spend single nights in other destinations; and (3) regional tours include tourists who spend more than one night in more than one destination. Finally, cluster analysis was conducted to categorize the similar regions and cities according to tourists who had been to the same destinations.

4. Results

The results of multi-trip patterns indicate that 57% of tourists are on a single destination trip and they spend on average 5.08 days in Austria, 30% of tourists are base camp travellers who stay 5.56 days on average, while the remaining 13% of the study sample are regional tourists who stay 8.11 days in Austria on average.

Fig. 1 is a map of multi-destination travel patterns based on the study sample. The size of the bubbles indicates the number of tourists who visited the destination. The arrows show the direction and percentage of tourists who originated from that destination. For instance, Vienna has the highest number of multi-destination tourists and the arrows indicate that tourists who are on a trip in Austria visit Vienna from Melk, Salzburg, Graz, Innsbruck, and Hallstadt. However, only a small percentage of those tourists visit other cities after their stay in Vienna. Travellers who had been to Vienna account for only 4% of overnight arrivals in Salzburg, and only 1% of arrivals to Innsbruck, Graz, and Melk.

The difference between cities based on different tourist types can be seen from the results as well. For example, Vienna is primarily a single destination city, as the majority of tourists stay only in Vienna (67%), while a further 21% of travellers use Vienna as a base camp. On the other hand, Melk is a regional tour city, as most of the tourists who visit Melk also spend multiple nights in other cities (58%), whereas only 1% of Melk tourists are single destination travellers. These characteristics are also reflected in the time spent at each destination. For instance, compared to Vienna, where tourists spend 2.55 days on average, Melk tourists stay only 1.08 days on average. On the other hand, Salzburg attracts predominantly base camp travellers (47%), who stay mainly in Salzburg but also spend a few nights in other cities, followed by regional tour travellers (32%), and single destination travellers
visitors to Salzburg stay an average of 1.63 days. Hallstatt, a small city located near Salzburg, is visited mainly by regional tour travellers (49%) and base camp travellers (47%).

Additionally, cluster analysis was conducted to group the cities that were visited by the same tourists, to find out whether the mapping of the cities and the trip categories were in line. Thus, the cities which were visited by the same tourists had to be in the same clusters. Since the focus of this study is multi-destination trips, only the 7920 tourists who had visited at least two different cities were included in this analysis. In addition, only the cities that had received more than ten tourists were included in the cluster analysis to ensure a more generalized data set. After computing the Yule coefficient for each pair of cities, the similarity matrix is transformed into a distance matrix to run a hierarchical cluster analysis using Ward’s method. The results of the cluster analysis support previous results of trip categorization and mapping of the multi-destination trips and are shown in Fig. 2.

Fig. 1. Multi destination travel patterns of tourists in Austria.

Fig. 2. City and region clusters*. *All the names of destinations are written in German as this was the case in metadata of the photos. For instance, in Group 3, Wien is the German name of Vienna.

The cluster analysis results indicate that there are two main clusters based on multi-destination trips taken in Austria. The first cluster encompasses the eastern part of the country, which includes larger cities such as Vienna, Graz, Linz, and Salzburg, and the second cluster refers to the western part of Austria’s Alpine area. In the case of Vienna and Melk, for instance, which were visited by the same tourists during their stay in Austria, most of the tourists stayed in Vienna according to the trip mapping figure which shows the destination from which the trips originated. One reason for this can be the close proximity (87 km) and good transportation connections between these destinations.

5. Conclusion and implications

The results of this study show that tourists tend to visit places which are in close proximity when they are on a multi-destination pleasure trip in Austria. This study has shown how data regarding multi-destination trips in Austria, which is not collected elsewhere, can be retrieved from Flickr and used in destination marketing. Multi-destination trip data is especially important for smaller destinations, whose marketing budgets are substantially smaller than larger cities such as Vienna, as these smaller destinations can join forces to attract travellers they know will be interested in visiting both destinations. For instance, destinations can place online advertisements on each other’s DMO websites. In this case, Hallstatt would be advised to advertise on Salzburg’s website since Salzburg is a base camp city from which many visitors also visit Hallstatt. Depending on the features of specific regions, attractions such as hiking trails or wine trails may be created to further bond proximate destinations. There are different ways to market destinations and knowing which destinations are visited by the same tourists during their vacation is important information for
national destination marketing as well. The Austrian National Tourist Office can use this information to create bundles of smaller and larger destinations according to the results of the cluster analysis. Future research may include seasonality aspects into such cluster analyses in order to classify winter tourists and summer tourists and identify the regions visited by each group. Also, multi-destination travel patterns can be identified at a broader scale, including multinational trips. In addition, conducting cluster analysis to find subcategories within each traveller type, such as subcategories of base camp travellers, would be interesting. This study can be replicated in different regions of the world to see if cultural differences have an impact on the duration of stay in different regions or cities.

This study is meant to show how to use big data in classifying multi-destination trips and how destinations can benefit from these trips. One limitation of the study stems from the assumption that all the travellers who upload photos on Flickr take and upload photos in all the cities or regions they visit, when in reality this may not be the case. Additionally, the travellers who upload photos on Flickr may not be representative of the whole traveller sample. A further limitation is related to the age of the data, which was collected in 2012 and may no longer be representative of this rapidly growing online data source. However, the scope and quantity of the data collected at that time is sufficient to ensure that the results are representative for Austria. Moreover, statistical evidence shows that in the last 10 years the number of bednights in the study sample cities increased by approximately between 1 and 5% (www.tourmis.info), which indicates that the classification of the destinations in Austria would not have changed over the intervening years.

Multi-destination trip information and knowledge on which other destinations tourists visit may be used to enhance the marketing efforts of individual DMOs, and inform the collaborative marketing efforts of associated destinations. Smaller destinations can use this information to create joint travel products with other destinations, or complementary products to attract more tourists from the bigger cities. Transportation authorities may also increase their ability of provide sufficient local transportation options such as buses and trains between popular destinations, especially during the more crowded seasons.

References