THE HONG KONG POLYTECHNIC UNIVERSITY
香港理工大學

Press Release
新聞稿

26 October 2018

PolyU Study Finds Photos Tell More Than A Thousand Words

A new data mining technique developed by Professor Rob Law of the School of Hotel and Tourism Management (SHTM) at The Hong Kong Polytechnic University and his co-researchers is able to capture extensive information on travellers' behaviour and travel patterns from huge photo datasets. In a recently published paper, the researchers apply the method to almost a million photos uploaded by Australian travellers to the photo-sharing site Flickr to discover their sequential travel destinations and itineraries, thus producing a wealth of information that could "support destination marketing organizations (DMOs) in promoting appropriate destinations to travellers".

Although vast amounts of tourism data are collected, few techniques are available for analysing them, preventing tourism managers from gaining useful insights into tourists' behaviour and preferences. According to the researchers, a better understanding of travel behaviour would allow tourism practitioners to "formulate more appropriate business strategies and travel service/products". Information about movement patterns, for instance, could help in identifying bottlenecks and unnecessary barriers in the flow between destinations and in "segmenting the tourism market to identify suitable travel packages".

Information about tourists' travel patterns is usually gathered from the travel history recorded by travellers during their trips, referred to as a "travel diary". Travel diaries may be recorded on paper, video and online blogs, and more recently using GPS loggers that provide travellers' locations. These diaries provide spatial and temporal information about the sequential associations between visited locations; for instance, many travellers who visit France also visit Italy, and visitors to North America also tend to visit Canada. The availability of such information allows tourism practitioners to create what the researchers call "appropriate and promising travel packages".

However, it is generally difficult to uncover such associations using traditional methods of analysis. Still, newly emergent data mining techniques are beginning to allow the extraction of more complex information from the large volumes of user-generated data on social media platforms.

As part of such advances, the researchers developed a novel data mining technique, referred to as "sequential rules mining" to "extract the sequential patterns from travel diaries" using geotagged photos taken by GPS-enabled devices and uploaded by travellers to the photo-sharing website Flickr. They focused on the international travel behaviour of Australian travellers and aimed to capture sequential travel patterns to destinations in Asia, Europe and America.

The first step was to retrieve photos from Flickr, along with information about the users that allowed the researchers to identify the target group, Australian travellers, and categorise them according to their home locations, such as those who were residents of Melbourne. They then

sorted the photos into domestic and international trips, so that they could analyse "outbound trips to other countries", and sorted them from oldest to newest. The final dataset included 809,313 photos taken by 3,623 users between 2001 and 2015.

In the next step, users' photo collections were converted into "sequences of visited destinations", called outbound travel diaries, by mapping the GPS information from the photos to their corresponding locations. The researchers point out that information can be retrieved at various levels, including the countries, cities and particular attractions that tourists visit. For simplicity, however, they focused on the country and city levels and did not attempt to differentiate between different travel purposes, such as business, holiday and family trips. In the final step, they applied their sequential rules mining method, a set of mathematical rules, to extract the patterns in the large dataset.

Once the photos had been sorted and processed according to location and time, the researchers obtained 17,188 travel diaries, a far higher number than has been analysed before. Of these, they note, "12,819 corresponded to a single country and 4,369 involved two or more countries". More than a third of trips within a single country were in Asia, whereas almost half of the trips to Europe involved visits to more than one country.

The most popular destinations were the US, the UK and New Zealand, and many travellers visited these countries several times, perhaps because these are the home countries of many Australian residents. Most of the top 20 destinations were the same as those identified in outbound travel surveys, with the exception of Fiji, which surprisingly did not appear in the top-20.

The 4,639 diaries involving travel to multiple countries were analysed to discover patterns of behaviour such as the likelihood of visiting one country after visiting another. For instance, Australian travellers who visited Canada or Mexico had an almost 75% chance of also visiting the US and those who landed in Bolivia had a greater than 87% chance of also visiting Peru. In Asia, the only strong association was between Laos and Thailand, mainly because most trips in this region were to a single destination. However, there were several strong sequential associations between European destinations, with high numbers of travellers to the Czech Republic, France and Austria also likely to visit Germany, while Italy was a common destination after visiting Austria, France or Greece.

There were many sequential associations with the UK, which was frequently the last destination after visiting various combinations of other European countries. The researchers suggest that a likely explanation for these patterns is that the UK is the home country of many Australian residents, and they are probably "taking advantage of their trips back home to visit other European countries on the way".

Next, the researchers examined the travel patterns "at the micro level between cities". They point out that much of the information here is redundant, because, for instance, the likelihood of travelling from Dublin to London is the same as travelling from Ireland to the UK. The same is

true for many of the most popular tourist destinations. Nevertheless, some interesting patterns emerged among second- and third-tier destinations. In the UK, for instance, travellers were likely to visit Oxford after Cowley, Edinburgh or Killington.

Most crucially, the researchers demonstrate how DMOs can use such information to identify travel patterns and itineraries. Australian travellers "often travel to Europe or America via Asian cities", they argue, because of the options provided by airlines. The itineraries extracted from the travel diaries show that "Dubai, Hong Kong and Singapore are the most popular destination for Australians traveling to London". Hong Kong is also a popular destination for those travelling to Paris, while Shanghai is more popular for those going on to Berlin. Among those travelling to Los Angeles, the most popular path is through Tokyo, and more than 70% of travellers who took this route spent more than one day there, suggesting that "Tokyo is usually visited for other purposes rather than simply for connecting flights".

DMOs can use such information to promote package trips with multiple destinations and thus "encourage travellers to travel to more destinations and purchase higher-value travel packages", according to the researchers. This applies at both the country and domestic levels. The strong associations between travel to Chicago, Denver and Los Angeles, for instance, suggest that DMOs could offer packages that target these cities, perhaps offering alternative modes of transport other than flying.

Ultimately, the researchers show how emergent data-mining techniques can be applied to offer useful insights into tourists' behaviour. Further application of their sequential rules mining method in the future could also offer the possibility of analysing data at more micro-levels, such as the sites visited at a destination, and other factors such as "travel styles, preferences, and travel purposes", which would provide even richer information for DMOs to develop more attractive packages.

Vu, Huy Quan, Li, Gang, Law, Rob and Zhang, Yanchun. (2018). Travel Diaries Analysis by Sequential Rule Mining. *Journal of Travel Research*, 57(3), 399-413.

*****

Press contact : Ms Pauline Ngan, Senior Marketing Manager
School of Hotel and Tourism Management

Telephone : (852) 3400 2634

E-mail : pauline.ngan@polyu.edu.hk

Website : www.polyu.edu.hk/htm