Tourism Management xxx (2016) 1–6

Contents lists available at ScienceDirect

Tourism Management

journal homepage: www.elsevier.com/locate/tourman

The Baidu Index: Uses in predicting tourism flows –A case study of the Forbidden City

Xiankai Huang, Lifeng Zhang*, Yusi Ding

Tourism College of Beijing Union University, Beijing 100101, China

HIGHLIGHTS

- Indicates role of Baidu analytics in forecasting future domestic tourist flows in China.
- Uses Baidu analytics with econometrics to forecast flows for the Forbidden Palace.
- Provides example of application during a "Golden Week".

ARTICLE INFO

Article history:
Received 7 March 2016
Accepted 7 March 2016
Available online xxx

Keywords:
Baidu Index
Tourist attractions
Co-integration
Autoregressive moving average model
Autoregressive distributed lag model

ABSTRACT

Tourist overcrowding of sites during the 'Golden Week' is a not an uncommon situation in China today. Consequently the prediction of tourist numbers is important for tourist attractions management and planning. Most existing methods rely on well-structured statistical data published by the government. However, this approach is limited in two aspects: 1) there may be significant delays in the publication of such data and 2) the sample size can be small, leading to inaccurate predictions. This paper proposes a novel approach for predicting tourist flows based on the Baidu Index. The Index provides search history containing different keywords on a daily basis dating back to 2006. The approach uses co-integration theory and Granger causality analysis to find the relationship between the internet search data and the actual tourist flow. The paper compares analysis results obtained by two kinds of predictive models, with or without considering Baidu Index. The study shows that there is a long-term equilibrium relationship and Granger causal relation between the observed number of tourists and a set of related keywords in the Baidu Index. It indicated a positive correlation between the increasing Baidu keyword search index and the increasing observed tourist flow.

© 2016 Published by Elsevier Ltd.

1. Introduction

With the rapid development of China's tourism industry, the number of tourists has significantly increased at famous tourism city and destinations, especially in May Day and National Day holidays, and many destinations are now over-crowded far beyond the destinations' carrying capacities. In the National Day Holiday in 2012, even ShanXi Huashan tourism destinations were significantly adversely affected, an event that made headlines throughout the country and attracted wide public concern. How to predict tourist flows, connect these predictions to a destination's carrying capacity, and make reasonable and timely arrangements has become important to avoid damage and improve sustainable development.

Traditional data forecasting methods come from government statistical reports and those of relevant statistical departments, but the collection and publication of data often lags well behind events and be based on insufficient samples, thereby rendering its usefulness to be insufficient for model building and effective prediction. Therefore, finding higher quality data resources has important significance for improving predictive results. The development of research engines has opened new areas for the prediction of economics, management and other disciplines. Based on internet search data, influenza forecasts, unemployment rates, automotive and real estate industries sales and other industries all have been shown to create higher forecast accuracy. The internet search data's real-time performance overcomes the lagging of traditional forecasting methods and is more timely.

* Corresponding author.
E-mail addresses: xiankai@buu.edu.cn (X. Huang), zhanglifeng999@163.com (L. Zhang), lytyusi@buu.edu.cn (Y. Ding).

http://dx.doi.org/10.1016/j.tourman.2016.03.015
0261-5177/© 2016 Published by Elsevier Ltd.

Please cite this article in press as: Huang, X., et al., The Baidu Index: Uses in predicting tourism flows –A case study of the Forbidden City, *Tourism Management* (2016), http://dx.doi.org/10.1016/j.tourman.2016.03.015

Consequently internet search data can be used to monitor and forecast socio-economic phenomena and it is possible to build composite indices through data analysis from complex data and find effective ways to solve problems. In China, use of Internet technology is growing in popularity, which means tourism information dissemination is not limited by time and space. Many tour operators and tourism enterprises as well as government departments publish tourism information through internet platforms, and websites are increasingly becoming the main information source for most tourists. Baidu and Google have launched the Baidu Index and Google search analysis functions, and through these functions one can obtain a keyword's attention frequency in Baidu and Google respectively for a given period. These functions could objectively reflect the tourist hot spots and users' interests and needs in a particular period. Therefore, based on the timeliness of internet search data, analysing the relationship between user's degree of attention to tourism websites and real tourism needs can well guide tourism management to better decision taking, as well as promoting sustainable tourism development.

2. Research condition at home and abroad

Past research studies have showed that correlations exist between these massive search data and the reality of social behavior. This research method was applied to epidemiological surveillance when Ginsberg et al (2009) used search analytic functions provided by Google and found a high correlation between the degree of internet presence of some keywords related to influenza virus and patients. By constructing a monitoring data model based on Google search data, the model could predict the trends in influenza outbreak two weeks earlier than traditional monitoring methods (Ginsberg et al., 2009). After this, this method quickly spread from infectious diseases to economics, social behavior, retail products sales forecasting (Choi & Varian, 2012), stock market forecasting (Da, Engellberg, & Gao, 2011; Liu, Lv, & Peng, 2011), unemployment rate forecasting (Askitas & Zimmermann, 2009), consumer confidence index forecasting (Schmidt & Vosen, 2009), and others – each gaining significant research achievements.

Currently, traditional tourist flow prediction research methods include time series, linear regression, exponential smoothing, gray forecasting, and artificial neural network models, etc. These methods mainly use historical data to predict, possess greater latency because of the need for historical data, and it is difficult to achieve high predictive accuracy. Along with the in-depth study of social behavior based on internet technology, tourist behavior prediction has also been studied. Tourism behavior prediction research based on internet search technology is mainly focused on the relationship between the degree of internet usage and tourist flows. Lu & Liu. (2007) analyzed the correlation between searches of relevant tourism websites and actual number of visitors, and internet information's guiding effect on tourist flow. They also researched Australia tourism websites, and confirmed the virtual network information flow on the role of the real tourist flow (Lu, Zhao, Wu, & Han, 2007). Li, Qui & Chen (2008) used Baidu Index for a statistical analysis of the degree of daily attention given to China's first 5A-class tourism destinations. Ma, Sun, Huang, and Zhou (2011) analyzed the main factors influencing tourist flow and temporal and spatial variation with the degree of internet usage and the marginal effect, and constructed relevant models of national tourist flow and tourist attention for both temporal and spatial variables. Long, Sun, Ma, and Wang (2011) found positive correlations between regional tourism Baidu usage and the actual tourist flow, and the former's change in advance to the latter's changes. Davidson and Yu 2005 used empirical comparative analysis between tourism websites in the case of Taiwan as a tourism destination and found that website information flow has a potential role in predicting tourists' travel behavior.

In summary, foreign studies use internet search data mainly for the forecasting of micro, meso and macro-economic phenomenon, but only a few have related to tourism forecasting by using of website search data. China has a literature about tourism website analysis, or just internet search data, but few have combined internet search data and real data together, and analyzed the correlation to establish predictive models. Therefore, this article takes The Palace Museum as an example, and uses Baidu search index, cointegration and Granger causality theory to find the relationship between internet search data and the actual number of tourists in The Palace Museum. It also uses the actual number of tourists in The Palace Museum to establish an auto-regressive moving average model and then by adding Baidu search index and an interpretation model to compare the predictions and so seeks to overcome the dependency on historical data of traditional forecasting methods, thereby gained better predictions.

3. Empirical analysis

3.1. Baidu keywords and data selection

Baidu (www.baidu.com) is the largest Chinese search engine. Baidu Company launched Baidu Index function based on massive data, and has provided different keyword's internet daily search frequency data since 2006. The Baidu Index is a free massive data analysis service based on Baidu web search and Baidu news, which used to reflect the different keyword's 'user awareness' and 'media attention' during the past period. From the Baidu Index, one can find, share, and mine information to reflect social hot spots, users' interests and needs. The Baidu Index is very easy to use, and involves going to the Baidu Index page, entering such keywords in the Baidu Index search bar, selecting the searched time, range and region, click on the 'Baidu', and then obtain the trend line of the keyword for the corresponding period and region. By moving the mouse over the trend line, the trend map can also show the searched keyword's daily search amount within the region indicated by the mouse.

The selection of search keywords can be seen as the core step for internet searching and relevant research about economic behavior. With reference to keyword selection methods, current studies have not yet achieved consensus. Currently, keyword selection methods take the following forms: technological keyword selection, keyword selection directly and keyword selection scope. Technological keyword selection uses high-performance, large-case scale computing facilities to place all possible keywords into a classification pertaining to a research area, and selects core keywords through the relevant statistical models compiled as procedures for calculation. Keyword selection directly determines keyword by using subjective assessment of a set of text or data. Keyword selection scope uses predetermined selection firstly, and then carefully selects text pertinent to keywords. Even though technological keyword selection has high accuracy, it is limited by the quality of algorithms and a need for powerful computing capable of a larger workload. Keyword selection directly and keyword selection scope reduce the workload significantly, but there exists the risk of core keyword omission (Sun & Lv, 2011).

Tourism activity is a mass behavior, and with the rapid development of internet search technology, people can obtain more information through the internet. Taking Beijing's Palace Museum as the subject of the case study, using the Baidu Index, and analyzing the relationship between internet search data and actual data through keyword selection, provided an example of how use of Internet search data might improve forecasting visitation to the

Please cite this article in press as: Huang, X., et al., The Baidu Index: Uses in predicting tourism flows –A case study of the Forbidden City, *Tourism Management* (2016), http://dx.doi.org/10.1016/j.tourman.2016.03.015

Download English Version:

https://daneshyari.com/en/article/5108735

Download Persian Version:

https://daneshyari.com/article/5108735

Daneshyari.com