International Journal of Contemporary Hospitality Management

Periodicity analysis and a model structure for consumer behavior on hotel online search interest in the U.S.

Juan Liu, Xue Li, Ya Guo,

Periodicity analysis and a model structure for consumer behavior on hotel online search interest in the U.S.

Submitted: 05 June 2015

1st Revision: 25 December 2015

2nd Revision: 10 April 2016

3rd Revision: 10 July 2016

Accepted: 14 September 2016

Abstract

Purpose

This work aims to analyze and model consumer behavior on hotel online search interest in the United States.

Design/methodology/approach

Discrete Fourier Transform was used to analyze the periodicity of hotel search behavior in the United States by using Google Trends data. Based on the obtained frequency components, a model structure was proposed to describe the search interest. A separable nonlinear least-squares algorithm was developed to fit the data.

Findings

It was found that the major dynamics of the search interest was composed of nine frequency components. The developed separable nonlinear least-squares algorithm significantly reduced the number of model parameters that needed to be estimated. The fitting results indicated that the model structure could fit the data well (average error 0.575%).

Theoretical implication

The reported model structure may be incorporated into a hotel demand prediction model in future research.

Practical implications

Knowledge of consumer behavior on online search is critical to marketing decision because search engine has become an important tool for customers to find hotels. This work is thus very useful to marketing strategy.

Originality/value

This research is the first work on analyzing and modeling consumer behavior on hotel online search interest.

Keywords: Hotel Online Search, Discrete Fourier Transform, Periodicity, Google Trends, Model Structure

1. Introduction

Consumers heavily rely on search engine to find hotel information when they are planning for a travel (Xiang and Pan, 2011). Search engine data have been found very useful for hotel demand prediction in literature (Pan et al., 2010; Pan et al., 2012; Kadir et al, 2014; Bangwayo-Skeete and Skeete, 2015) but there is still a lack of a general model structure with good performance for all applications (Song and Li, 2008; Pan et al., 2012). The methods or model structures being used are often purely numeric or empirical without examining the hidden mechanism of the data or the regularities of consumer behavior. This might contribute to the issue that there is no single method consistently superior to others.

As online search may lead to potential hotel demand, it can be treated as an input variable if hotel demand is the output variable. Better understanding the regularities of consumer search behavior and expressing the search behavior with an explicit analytical function will provide information for hotel demand model structure selection and results analysis and interpretation. It is thus meaningful to investigate the regularities of consumer search behavior and further model the consumer search behavior with an explicit analytical function.

Internet marketing is very important to the tourism and hospitality industry (Werthner and Ricci, 2004; Wang and Fesenmaier, 2006). The moment of consumer online search is the moment that online advertisements potentially take effect. It has been well recognized that an optimal marketing strategy should be based on the understanding of consumer behavior (El-Ansary, 2006; Laroche, 2010; Nica, 2013). From this point of view, it is also meaningful to study and model the regularities of consumer online search behavior.

Periodicity is an important feature of regularities. However, there is no research on the periodicity of hotel online search in the literature. It is not clear whether hotel online search includes some cycles and what the periods for the major cycles are if any, and whether the information of the cycles can lead to an effective analytical expression (model structure, or explicit analytical function) to describe the online search behavior. The present study aims to fill the gaps. In this work, the periodicity of hotel search behavior in the United States is analyzed by Discrete Fourier Transform using Google Trends data. It is found that the major dynamics of the search interest is composed of nine harmonic components. The finding fills voids in literature. A model structure is then successfully developed to describe the daily search interest over the past 11 years based on the periodicity analysis. The cyclic nature of consumer hotel online search behavior obtained from this study may be incorporated into a model structure of hotel demand prediction based on search engine data in future research if hotel demand is treated as the output and the online search behavior is treated as an input. The periodicity may also be used to interpret hotel selling variations with time and employed to plan internet marketing strategy.

2. Literature Review

In this section, recent relevant literature on (1) traditional methods for forecasting hotel room demand, (2) forecasting with search engine data and its application in hospitality and tourism, and (3) periodicity analysis in business research is briefly reviewed.

2.1 Traditional Methods for Forecasting Hotel Room Demand

Previous researchers have used different methods to forecast hotel room demand or hotel occupancy rates. For example, Andrew et al. (1990) forecast the monthly occupancy of a hotel by using the Box-Jenkins method and the exponential smoothing method. By using previous booking curves, Schwartz and Hiemstra (1997) built a model structure to forecast daily occupancy rates. Law (1998) adopted a neural net work approach to forecast the hotel occupancy in Hong Kong. Zakhary et al. (2009) used Monte Carlo simulation method to forecast hotel occupancy. Lim, Chang and McAleer (2009) used Holt-Winters and Box-Jenkins ARMA models to forecast guest night demand in New Zealand. All of these studies have achieved certain accuracy on forecasting hotel room demand or occupancy rates; but they all used historical sale or booking data of individual hotels, which may lead to two limitations. First of all, forecasting based on historical performance assumes a consistent pattern in a stable economic structure, dramatic structural changes in economy or any one-off events may decrease their forecasting accuracy (Yang et al., 2014; Yang et al., 2015). Second, forecasting demand of individual hotels is not robust enough to present the hidden mechanism of hotel consumers' purchasing behavior in general. In order to build a forecasting model that could reveal the hidden mechanism of hotel consumers' purchasing behavior, research with a large amount of consumer purchasing behavior data is needed. While a large amount of consumer purchasing behavior data is not easy to access, consumers' online search behavior could be used as a proxy; therefore, investigating the regularities of online search behavior and expressing it with an explicit analytical function is necessary.

2.2 Forecasting with Search Engine Data and Its Application in Tourism and Hospitality

Because of increasing available data related to online activities, there is research on economic behavior prediction in literature by using online data. This kind of data is traces of consumer behavior, could unveil consumers' interests and purchases in real time; therefore, they are thought as indicators of health status of an industry and are named as "online pulse data" (Hubbard, 2011). Online pulse data which generally include search engine query volumes and social media postings are real-time, high-frequency, and sensitive to slight changes in consumer behavior. There have been a number of studies using search engine data for prediction in various fields. For example, Ginsberg et al. (2009) successfully built a model and predicted influenza outbreaks from the search volumes of certain keywords on Google. Their model could detect flu outbreaks two weeks ahead of CDC. Choi and Varian (2009) built an autoregressive integrated moving average (ARIMA) model with search engine data to predict unemployment claims. Askitas and Zimmermann (2009)'s simple error-correction model indicated a strong correlation between keyword searches and unemployment rates. Zhang et al. (2009)'s study suggests that search engine data could detect users' behavior change across different time periods. Choi and Varian (2012) showed the usefulness of search engine data for forecasting automobile sales.

There is limited number of research taking advantage of search engine data in the field of hospitality and tourism, but the value of search engine data has been recently recognized by researchers in the field. Gawlik et al. (2011) used query-specific search data to predict tourism rates. Feature selection and *k*-fold cross validation were employed to identify the most relevant queries and test the performance of the algorithm. The developed method has successfully predicted the tourism rates in Hong Kong. In Xiang and Zhang (2011)'s research, the patterns of online travel queries across tourist destinations was investigated by using transaction log files from search engines. It was concluded in the work that search volume might serve as a direct indicator of tourism industry. Choi and Varian (2012) emphasized the usefulness of search engine data for forecasting travel destination demand. Yang et al. (2015) demonstrated the value of search engine data for forecasting Chinese tourist volume and proposed a method for selecting predictive queries. Search engine data have also been used for predicting hotel demand by two groups of researchers. In Pan et al. (2012)'s study, the usefulness of search query volume for hotel room demand prediction was investigated. They also studied the performance of various model structures for predicting hotel demand using search engine data, including ARMA (Autoregressive–moving-average), ARMAX (Autoregressive–moving-average with exogenous inputs), ADL (Autoregressive distributed lag), TVP (Time-varying parameter), and VAR (Vector autoregressive). In Yang et al. (2014), the web traffic data of the destination organization has also been used to predict hotel demand. Recent studies of forecasting tourist inflow or tourism demand by using Google trends data can be found in Park et al. (2016) and Önder and Gunter (2016), respectively. Although there are attempts for predicting hotel demand with search engine data, there is still no single general method with good performance for all applications (Song and Li, 2008; Pan et al., 2012). In these model structures, search engine data are often used numeric without examining the regularities of online search behavior, which might reduce the performance of the models (Wong, 1997).

2.2 Periodicity Analysis in Business Research

Periodicity analysis could unveil the hidden relationship between seasonality and consumer behavior, and that has aroused strong interest from business practitioners (Nieves-Rodríguez et al., 2015). Many attempts in the fields of business can be found in literature. For example, Bjørnland (2000) studied the business cycles in Norway. Different detrending methods were applied and the detrended data were analyzed in both time and frequency domains. Broszkiewicz-Suwaj et al. (2004) showed that electricity price returns have daily and weekly pattern and recommended that the periodic nature of the process should be considered for modeling seasonal decisions of consumers and reveal business cycle. Optimized marketing strategy is critical in fashion retailing because of the short selling window. In the work of Choi et al. (2014), it is shown that the features of seasonal cycles have significant impact on the performance of the developed forecasting algorithm for fast fashion demand. Unexpected weather change causes the loss of retail sales and profits. Bertrand et al. (2015) studied the influence of unexpected temperature change on sales by detrending seasonal weather influence pattern. Pricing cycles were also investigated in literature (Bronnenberg, 2006).

As for in the field of hospitality and tourism, our literature search show that only four studies on periodic analysis were found in tourism literature, but none was in hospitality. In one work, the periodic nature of outbound United Kingdom (UK) passengers to several international destinations including transatlantic and northwest Europe was examined (Coshall, 2000a). The result displayed only one significant cycle which was the seasonal cycle. In another research carried by Coshall (2000b), spectral analysis was used to study overseas tourist's expenditures in the UK. Periodic behavior within and between series data was revealed and the annual cycle was the only significant one. Spectral analysis was used by Latzko (2015) to investigate the influence of fluctuations in the Japanese economy on the Hawaiian tourist industry. The periodic behavior between quarterly time series for the number of visitors to Hawaii from Japan and several economic or financial indexes, such as gross domestic expenditures, currency exchange rate, tourism prices, was analyzed. It is found that the cycles of tourism prices were strongly correlated with tourist arrival cycles. Spectral analysis was also conducted to reveal the seasonality in New Zealand tourism demand from Australia and the Unites States (Chan and Lim, 2011). These findings indicated that different travel purposes have similar periodic components but they have different weights in the total variation of tourism demand, demonstrating the necessity of periodicity analysis on demand for different hospitality and tourism products.

However, the above literature review for periodicity analysis studies also revealed that there was a big void in periodicity analysis research in the field of hospitality and tourism. While with its seasonality trait of hospitality and tourism, periodicity analysis research is undoubtedly important and necessary for demand prediction and marketing strategy selection. The present study aims to contribute to periodicity analysis research in hospitality by building a model structure with search engine data.

3. Data Description

The data of consumer's hotel search interest were downloaded from Google Trends (http://www.google.com/trends/hottrends). "Hotel" was used as the search term and the location was set as "United States" (U.S.). The query for the keyword of "Hotel" reflects the number of all the search records that include "Hotel", which is much higher than the number for searching a specific hotel. This can be easily verified from Google Trends. The query for "Hotel" thus can reflect the trend of overall interest in hotels except for differing by a scaling factor, and can be considered as a sign of the potential hotel demand. Fourier transform is a linear transform and the scaling factor does not affect the analyzed periods. For the more specific periodicity analysis for a specific hotel and/or a specific location, corresponding data can be downloaded from Google Trends and analyzed by following the framework proposed in this research. Google Trends provides data on search interest after 2004, but only provides daily search interest when the query duration is set not longer than three months; otherwise, data are provided at weekly or monthly base. In order to obtain the daily search interest over years, the data were downloaded for three consecutive months each time with one month overlapping, for example, one query is from January to March, the next one is from March to May. Because the Google Trends data are relative data for each query, the overlapping data were used to normalize all the downloaded data. In this way, the relative daily search interest from 2004 to 2014 was obtained.

4. Periodicity Analysis and Development of a Model Structure for the Online Search Interest

4.1 Periodicity Analysis

Discrete Fourier Transform (DFT) was employed to analyze signal periodicities or frequency components (A sine or a cosine function is a frequency component. A signal usually contains multiple frequency components, which have different frequencies.) in this work, which is the discrete form of Fourier Transform and the Fast Fourier Transform is a fast algorithm realization of DFT. Fourier Transform is a linear transform as mentioned, has been widely used to analyze signal frequency components in many fields including business and tourism (Chan and Lim, 2011; Fumi et al., 2013; Wong, 1997), it is also the most commonly used method for periodicity analysis. The obtained sine or consine components by Fourier Transform were incorporated into linear trend predicting models, which significantly increased the accuracy of the models in previous studies (Fumi et al., 2013; Wong, 1997). However, to the authors' knowledge, no study was reported in the literature on applying DFT in the hospitality industry. The Fourier Transform of a signal $x(t)$ is defined as (Harris and Stocker, 1998):

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi t}dt \tag{1}$$

where f is frequency, j is the imaginary unit, and $X(f)$ is the spectrum at frequency f. DFT is the discrete version of Fourier Transform, which can be represented as:

$$X(m\Delta f) = \sum_{p=0}^{P-1} x(p\Delta t)e^{-j2\pi mk/P} \tag{2}$$

where P is the total number of sampling data points, Δt is sampling interval, Δf is frequency resolution and it is equal to $1/(P\Delta t)$, $X(m\Delta f)$ is the m^{th} spectral line (m=0, 1, …, P-1).

The period and frequency of a signal component are reciprocal to each other. According to this, the spectrum can be represented as a function of period as:

$$X(T_m) = X(\frac{1}{m\Delta f}) = \sum_{p=0}^{P-1} x(p\Delta t)e^{-j2\pi mp/P} \tag{3}$$

where T_m is the period of the m^{th} signal component. When the spectrum is expressed as a function of period, m cannot take the value of 0 as its reciprocal is infinite large. This will not affect the analysis result, because the 0^{th} frequency component corresponds to the constant shift, which is the commonly called DC component.

$X(T_m)$ is a complex number, which can be written as:

$$X(T_m) = X(T_m)_R + jX(T_m)_I \tag{4}$$

where $X(T_m)_R$ and $X(T_m)_I$ are the real part and imaginary part of $X(T_m)$, respectively.

The amplitude spectrum of $X(T_m)$ can be expressed as:

$$\|X(T_m)\| = \sqrt{X(T_m)_R^2 + X(T_m)_I^2} \tag{5}$$

Matlab was used to conduct the DFT analysis. The data from the past 11 years were used.

4.2 Development of a Model Structure for the Online Search Interest

According to the results of DFT analysis, the major dynamics of the search interest (y) over the past 11 years can be represented by the summation of some sine or cosine functions, which is

$$y = \sum_{i=1}^{N} [g_i \cos(2\pi \frac{1}{T_i} t + \varphi_i)] \tag{6}$$

where g_i is the spectral magnitude of the i^{th} frequency component, T_i is the period of the i^{th} frequency component, φ_i is the initial phase of the i^{th} frequency component, t is time, and N is the number of selected frequency components. In this work, the unit for t is day and it is counted as the number of days after January 1, 2004 (January 1, 2004 is time 0). A higher N value implies including more frequency components, which enables the model to cover more details of the data, but it is not necessary because the major dynamics is the most important and more details bring complexity to the model structure and reduces computation speed. Eqn. (6) can be expanded as:

$$
\begin{aligned}
y &= \sum_{i=1}^{N}[g_i \cos(2\pi \frac{1}{T_i} t)\cos(\varphi_i) - g_i \sin(2\pi \frac{1}{T_i} t)\sin(\varphi_i)] \\
&= \sum_{i=1}^{N}[c_i \cos(2\pi \frac{1}{T_i} t) + d_i \sin(2\pi \frac{1}{T_i} t)]
\end{aligned} \tag{7}
$$

where $c_i = g_i \cos(\varphi_i)$, $d_i = -g_i \sin(\varphi_i)$.

Because the search interest over the past 11 years has a slow exponential decreasing trend and the amplitude of the data looks smaller at the later stage, the trend is modeled as an exponential function and the model structure in Eqn. (6) or (7) an be further written as:

$$
y = (a + be^{-kt})(c_0 + \sum_{i=1}^{N}[c_i \cos(2\pi \frac{1}{T_i} t) + d_i \sin(2\pi \frac{1}{T_i} t)]) \tag{8}
$$

where c_0 is a constant value.

Assume that the data of Google search interest are denoted as $\hat{y} = [\hat{y}_1, \hat{y}_2, \hat{y}_3, \cdots]^T$ and the search interest predicted by Eqn. (8) is denoted as $y = [y_1, y_2, y_3, \cdots]^T$.

The sum of squared errors between $\hat{y}$ and y is:

$$
e = \left\| \hat{y} - y \right\|^2 \tag{9}
$$

Minimization of error e in Eqn. (9) is equivalent to finding an optimal solution of η for

$$
\Psi \eta = \hat{y} / (a + be^{-kt}) \tag{10}
$$

where

$$
\Psi = \begin{bmatrix}
1 & 1 & 0 & \cdots & 1 & 0 \\
1 & \cos(2\pi \frac{1}{T_1}) & \sin(2\pi \frac{1}{T_1}) & \cdots & \cos(2\pi \frac{1}{T_N}) & \sin(2\pi \frac{1}{T_N}) \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
1 & \cos(2\pi \frac{M-1}{T_1}) & \sin(2\pi \frac{M-1}{T_1}) & \cdots & \cos(2\pi \frac{M-1}{T_N}) & \sin(2\pi \frac{1}{T_N})
\end{bmatrix} \tag{11}
$$

and

$$\eta = \begin{bmatrix} c_0 \; c_1 & d_1 & \cdots & c_N & d_N \end{bmatrix}^T \tag{12}$$

The least-squares solution for η is:

$$\eta = (\Psi^T \Psi)^{-1} \Psi^T \bar{y} \tag{13}$$

where $\bar{y} = [\frac{\hat{y}_1}{a+b}, \frac{\hat{y}_2}{a+be^{-k}}, \frac{\hat{y}_3}{a+be^{-2k}}, \cdots]^T$

The sum of squared errors e in Eqn. (9) can thus be expressed as:

$$e = \left\| \hat{y} - \Psi \eta y * \right\|^2 = \left\| \hat{y} - \Psi (\Psi^T \Psi)^{-1} \Psi^T y^* \right\|^2 \tag{14}$$

where $y^* = [\hat{y}_1(a+b), \hat{y}_2(a+be^{-k}), \hat{y}_3(a+be^{-2k}), \cdots]^T$

The total squared error expressed by Eqn. (14) only depends on a, b, and k, which significantly reduces the variables that need to be estimated, shortens computational time, and improves algorithm convergence. The Levenberg-Marquardt method (Marquardt, 1963; Levenberg, 1944; Constantinides and Mostoufi,1999) can be used to optimize the three parameters and fit the data from Google search interest by updating parameters at each iteration as:.

$$\Delta \boldsymbol{u} = (\lambda \mathrm{I} + J^T J)^{-1} J^T (\hat{y} - \Psi (\Psi^T \Psi)^{-1} \Psi^T y^*) \tag{15}$$

where $\boldsymbol{u}= [u_1, u_2, u_3] = [a, b, k]$, I is the identity matrix, λ is a coefficient for improving the convergence of the algorithm, J is the Jacobian matrix of partial derivatives of y^* with respect to $\boldsymbol{u}$ evaluated at all P data points given as:

$$J = \begin{bmatrix} \frac{\partial y_1^*}{\partial u_1} & \frac{\partial y_1^*}{\partial u_2} & \frac{\partial y_1^*}{\partial u_3} \\ \vdots & \vdots & \vdots \\ \frac{\partial y_P^*}{\partial u_1} & \frac{\partial y_P^*}{\partial u_2} & \frac{\partial y_P^*}{\partial u_3} \end{bmatrix} \tag{16}$$

5. Results

5.1 Results of Periodicity Analysis

Google search interest on “Hotel” in the U.S. from 2004 to 2014 is shown in Figure 1 in the time domain. In Figure 1(a), the search interest from 2004 to 2014 is shown as a 1D diagram, which clearly indicates that the data have an annual cyclic pattern and a decreasing trend. The annual cyclic pattern is expected because human activities have a lot of similarities each year. The decreasing trend may be caused by fewer consumers using Google to search for hotels because other tools for hotel information finding are available or become popular. It looks in an exponential manner, which may affect the magnitude of frequency components, but should not affect the estimated component frequencies as the cyclic pattern is still dominant. This has been verified by the Fourier spectrum in the later part of this section and the model fitting in section 5.2. Figure 1b shows the search interest for each year and each month in 3D manner. It reveals that there are at least two monthly search peaks in a year, one is in the middle February and the other is in the middle June. In order to show consumer weekly search behavior, a 3D diagram for 4-week average of each year is shown in Figure 1c, which clearly reveals the weekly pattern. Although Figures 1 (a), (b), and (c) show clear cyclic pattern, it is unclear what the exact periods are and whether there are some other cyclic components.

<Figure 1>

Fourier amplitude spectra for hotel search interest in the U.S. from 2004 to 2014 are shown in Figures 2(a), (b), and (c) with periods in different scales. Figure 2(a) shows the overall picture of the spectra. It clearly reveals that the search interest has five major frequency components with periods ranging from 60 days to 420 days. Figure 2 (b) reveals that the spectra have several frequency components with the periods ranging from 30 days to 60 days. The magnitudes of these components are smaller than those in the 60 days to 420 days range. There is no significant frequency component with the period between 10 days and 30 days, so a separate figure is not used exclusively to show the spectrum in this range; but it can be noticed from the combination of Figure 2(a) and 2(c). Figure 2 (c) shows the spectra with short periods or high frequencies. As indicated in the figure, there are three major peaks and the highest one has a period of 7 days. The magnitudes and frequencies of major frequency components are listed in Table 1.

As listed in Table 1, the periods of the nine major components of the search interest are 365.2727 days, 182.6364 days, 7.0000 days, 121.7576 days, 3.5000 days, 91.3182 days, 73.0546 days, 60.8788 days, and 52.1818 days. It is very interesting because the 365.2727-day cycle corresponds to a year, 182.6364-day cycle corresponds to half a year, 7.0000-day cycle corresponds to a week, 121.7576-day cycle corresponds to four months, 3.5000-day cycle corresponds to half a week, 91.3182-day cycle corresponds to three months, 60.8788-day cycle corresponds to two months, 73.0546-day cycle corresponds to 1/5 year, and 52.1818-day cycle correspond to 1/7 year. The spectrum of Figure 2(b) does indicate that there is a cycle corresponds to a month (30 days), but the amplitude is much smaller, which implies that hotel search does not change significantly and regularly in a monthly manner. The 3.5000-day cycle implies that consumers have different hotel search tendencies even within a week, which is witnessed by the 3D diagram for 4-week average of each year in Figure 2(c). The frequency analysis of the supermarket visits conducted by Mackay (1973) also showed half-a-week cycle.

<Figure 2>

<Table 1>

5.2 Model Fitting Performance

According to the major frequency components as listed in Table 1, N is set as 9 in this work. Figure 3 shows the model fitting for the data over the 11 years and the corresponding residue based on the error function of Eqn. (14). The relative error for Figure 3 is 0.575%, which is computed as $\sqrt{\frac{1}{M}\sum_{i=1}^{M} e_i^2}$.

<Figure 3>

The function for the trend in Eqn. (8) can be changed to other functions by slightly modifying the model structure and the algorithm according to the specific data in other applications or even does not include it if there is no significant trend.

6. Discussion and Conclusion

6.1 Theoretical Implications

The results of this work clearly reveal that consumer behavior on hotel search has cyclic patterns. The top 3 frequency components have the periods of 365.2727 days, 182.6364 days, and 7.0000 days, which means that the behavior heavily change annually, half-a-yearly, and weekly. The behavior has a monthly changing cycle, but the cycle is not as significant as the weekly or annual cycle because its magnitude is relatively very small. The proposed model structure can describe hotel search behavior at a daily base over past 11 years with a low relative error. It proves the usefulness of the model structure and the correctness of consumer search periodicities obtained in this work.

Our literature search indicates that this work is the first one on the periodicity of consumer hotel search or plan behavior. The method proposed in this work does not depend on the hospitality industry, because there is no specific assumption for the hospitality industry is made in the equations. Therefore, any processes that include cyclic components can be approximated by the summation of sine or cosine functions and the model parameters can be estimated by the separable nonlinear least-squares algorithm. Actually, many processes are not totally random but contain some periodic components. For example, the business of restaurants, flights, tourism and etc all appear high demand and low demand patterns. The periodicities of consumer search or plan behaviors in these fields can thus be studied by the proposed framework. The algorithm of this study could be embedded in professional softwares. Application of the proposed method by using professional softwares is not as technical as it appears. It can be operated just like ANOVA without knowing too much about the coding.

Understanding the periodicities of consumer search behavior is helpful to hotel market demand prediction. The hotel market demand and the online search interest can be both described by the model structure of a summation of multiple sine or cosine functions as Eqn. (6) or Eqn. (8) (The difference between Eqn. (6) and (8) is with or without a trend term.). The model parameters G_i, Ψ_i, g_i, and φ_i can be estimated from historical data, where G_i is the magnitude of the i^{th} frequency component of the hotel market demand, Ψ_i is the initial phase of the i^{th} frequency component of the hotel market demand, g_i is the spectral magnitude of the i^{th}

frequency component of the online search interest, φ_i is the initial phase of the i^{th} frequency component of the online search interest for historical data, respectively. The relationship between historical search interest and historical hotel market demand can be presented by the magnitude ratio (G_i / g_i) of each frequency component and the initial phase difference ($\Psi_i - \varphi_i$) of the corresponding frequency component. The magnitude ratio (G_i / g_i) of each frequency component of the hotel market demand and the online search interest implies the magnitude change of the hotel market demand caused by unit change of the online search interest at the specific frequency or period, which can be called as conversion rate. The initial phase difference ($\Psi_i - \varphi_i$) of the corresponding frequency component implies how long it will take for a change of the online search interest to cause the hotel market demand change at the specific frequency or period, which can be called as time lag or conversion delay. If the conversion rate and time lag do not change significantly in short time, the hotel demand for the near future can be predicted easily by using the most recent online search interest data, which obviously considers both the long term periodic changes and the real short term online search volume dynamics. This may lead to an alternative and more accurate model structure for hotel market demand prediction (Wong, 1997) and the model parameters have clear physical meaning.

6.2 Practical Implications

Consumer search periodicities can be incorporated into the marketing strategy to plan promotion events and optimize the profits of the company. For example, the online advertisement, marketing emails, and pricing can be released or changed in a manner to compensate or make use of the dynamics of online search periodicities (yearly, half-a-yearly, and weekly, instead of monthly). In this way, maximum sales or more constant daily sales might be achieved. As explained above, the proposed concept can be used to predict market demand, which is useful to marketing strategy optimization quantitatively.

6.3 Limitations and Future Research

The obtained periodic features in this work were only analyzed by using the data of hotel online search behavior in the U.S. It may not be generalized to the hotel online search behavior in other countries because people in different cultures have different habits. For example, people in high uncertainty avoidance culture like to make plans, but people in low uncertainty avoidance culture do not (Hofstede, 1980). The habit of making plans or not may affect people's online search behavior. Therefore, it is valuable for future research to examine the regularities of people's hotel online search behavior in other cultures with higher or lower uncertainty avoidance than the U.S.

6.4 Conclusion

In this work, DFT was employed to analyze the periodicity of hotel search behavior in the United States using Google Trends data. The major dynamics of the online search interest is found with nine cyclic components. The top 3 frequency components reveal that the online search behavior heavily changes annually, half-a-yearly, and weekly. These immediately lead to a model structure on hotel search interest and a separable nonlinear least-squares algorithm is designed to fit the data. The model can fit the data with a daily relative error of 0.58% over the past 11 years. Along with the developed model structure, the identified periodicities can be incorporated into hotel market demand prediction model and marketing strategy in future research, which allows changing online advertisement, marketing emails, and pricing plans to achieve maximum sales or more constant daily sales. The proposed concept and algorithm does not depend on specific business, therefore any process in business that includes cyclic components can be analyzed by the framework.

References

Andrew, W. P., D. A. Cranage, and C. K. Lee. (1990), "Forecasting Hotel Occupancy Rates with Time Series Models: An Empirical Analysis", *Journal of Hospitality & Tourism Research*, Vol. 14 No. 2, pp. 173-82.

Askitas, N., and Zimmermann, K. F. (2009), "Google Econometrics and Unemployment Forecasting." *Applied Economics Quarterly*, Vol. 55 No. 2, pp.107-20.

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

Bertrand, J. L., Brusset, X.and Fortin, M. (2015),"Assessing and hedging the cost of unseasonal weather: case of the apparel sector", *European Journal of Operational Research*, Vol. 244 No.1, pp. 261-276.

Bjørnland, H. C. (2000),"Detrending methods and stylized facts of business cycles in Norway–an international comparison", *Empirical Economics*, Vol. 25 No. 3, pp. 369-392.

Bronnenberg, B. J., Mela, C. F.andBoulding, W. (2006), "The periodicity of pricing", *Journal of Marketing Research*, Vol. 43 No. 3, pp. 477-493.

Broszkiewicz-Suwaj, E., Makagon, A., Weron, R. and Wyłomańska, A. (2004),"On detecting and modeling periodic correlation in financial data", *Physica A: Statistical Mechanics and its Applications*, Vol. 336 No.1, pp. 196-205.

Chan, F. and Lim, C. (2011), “Spectral analysis of seasonality in tourism demand”, *Mathematics and Computers in Simulation*, Vol. 81 No.7, pp. 1409-1418.

Chiu, Y. J., Chen, H. C., Tzeng, G. H.andShyu, J. Z. (2006),“Marketing strategy based on consumer behaviour for the LCD-TV”, *International Journal of Management and Decision Making*, Vol. 7 No.2, pp. 143-165.

Choi, T. M., Hui, C. L., Liu, N., Ng, S. F.and Yu, Y. (2014), “Fast fashion sales forecasting with limited data and time”, *Decision Support Systems*, Vol. 59, pp. 84-92.

Choi, H., and Varian, H. (2009). “*Predicting Initial Claims for Unemployment Benefits*.” In Google Technical Report.

Choi, H.and Varian, H. (2012), “Predicting the present with google trends”, *Economic Record*, Vol. 88 No.s1, pp. 2-9.

Chan, F.and Lim, C. (2011),“Spectral analysis of seasonality in tourism demand”, *Mathematics and Computers in Simulation*, Vol. 81 No.7, pp. 1409-1418.

Clemons, E. K. (2008), “How information changes consumer behavior and how consumer behavior determines corporate strategy”, *Journal of Management Information Systems*, Vol. 25 No.2, pp. 13-40.

Constantinides, A. andMostoufi, N. (1999), *Numerical methods for chemical engineers with MATLAB applications*(Vol. 443), Prentice Hall PTR, Upper Saddle River, NJ.

Coshall, J. (2000a), “Spectral analysis of international tourism flows”, *Annals of Tourism Research*, Vol. 27 No.3, pp. 577-589.

Coshall, J. (2000b), “Spectral analysis of overseas tourists’ expenditures in the United Kingdom”, *Journal of Travel Research*, Vol. 38 No.3, pp. 292-298.

El-Ansary, A. I. (2006), “Marketing strategy: taxonomy and frameworks”, *European Business Review*, Vol. 18 No.4, pp. 266-293.

Fumi, A., Pepe, A., Scarabotti, L., & Schiraldi, M. M. (2013), “Fourier analysis for demand forecasting in fashion company”, *International Journal of Engineering Business Management*, Vol. 5 No. 30, pp. 1-10.

Gawlik, E., Kabaria, H.andKaur, S. (2011), “Predicting tourism trends with Google Insights”, available at http://mfile.narotama.ac.id/files/Umum/JURNAR%20STANFORD/Predicting%20tourism%20trends%20with%20Google%20Insights.pdf/(accessed 10 February 2015).

Ginsberg, J., Mohebbi, M. H., Patel,R. S., Brammer, L., Smolinski,M. S., and Brilliant,L. (2009).“Detecting Influenza Epidemics Using Search Engine Query Data.” *Nature*, Vol. 457 No. 7232, pp. 1012-14.

Harris, J. W.andStöcker, H. (1998), *Handbook of mathematics and computational science*.Springer Science & Business Media.

Hofstede, G. (1980), “Culture’s Consequences: International Differences in Work-Related Value”. Beverly Hills, CA: Sage Publications.

Huang, P., Lurie, N. H. andMitra, S. (2009), “Searching for experience on the web: an empirical examination of consumer behavior for search and experience goods”, *Journal of Marketing*, Vol.73 No.2, pp. 55-69.

Hubbard, D. W. (2011). *Pulse: The New Science of Harnessing Internet Buzz to Track Threats and Opportunities*, Wiley, Hoboken, NJ.

Kadir, S. N., Tahir, N. M., Yassin, I. M. and Zabidi, A. (2014), Malaysian tourism interest forecasting using Nonlinear Auto-Regressive Moving Average (NARMA) model, in *Wireless Technology and Applications (ISWTA), 2014 IEEE Symposium,* pp. 193-198.

Laroche, M. (2010), “Advances in internet consumer behavior and marketing strategy: Introduction to the special issue”, *Journal of Business Research*, Vol. 63 No.9, pp. 1015-1017.

Latzko, D. A. (2005), “Economic conditions and Japanese tourism to Hawaii”, *Asia Pacific Journal of Tourism Research*, Vol. 10 No.2, pp. 151-156.

Law, R. (1998), “Room Occupancy Rate Forecasting: A Neural Network Approach”, *International Journal of Contemporary Hospitality Management*, Vol. 10 No. 6, pp. 234-39.

Levenberg, K. (1944), “A method for the solution of certain problems in least squares”, *Quarterly of applied mathematics*, Vol. 2, pp. 164-168.

Lim, C., Chang, C., & McAleer, M. (2009), “Forecasting h(m)otel guest nights in New Zealand”, *International Journal of Hospitality Management*, Vol. 28, pp.228-235.

MacKay, D. B. (1973), “A spectral analysis of the frequency of supermarket visits”, *Journal of Marketing Research*, pp. 84-90.

Marquardt, D. W. (1963), “An algorithm for least-squares estimation of nonlinear parameters”, *Journal of the Society for Industrial & Applied Mathematics*, Vol. 11 No.2, pp. 431-441.

Nica, E. (2013), “Marketing implications of consumer behavior”, *Economics, Management, and Financial Markets*, Vol.1, pp. 124-129.

Nieves-Rodríguez, E. B., Cao-Alvira, J. J. and Pérez, M. M. (2015), “The Influence of Special Occasions on the Retail Sales of Women’s Apparel”, in *Ideas in Marketing: Finding the New and Polishing the Old,* Springer International Publishing, pp. 213-221.

Önder, I. and Gunter, U. (2016) “Forecasting Tourism Demand with Google Trends For a Major European City Destination” *Tourism Analysis*, Vol. 21 No 2-3, pp.203-220.

Pan, B., Xiang, Z., Tierney, H., Fesenmaier, D. R. and Law, R. (2010), “Assessing the dynamics of search results in Google”, *Information and Communication Technologies in Tourism* Vol. 2010, pp. 405-416.

Pan, B., Chenguang Wu, D. and Song, H. (2012),“Forecasting hotel room demand using search engine data”, *Journal of Hospitality and Tourism Technology*, Vol. 3 No.3, pp. 196-210.

Park, S., Lee, J., and Song, W. (2016), “Short-term forecasting of Japanese tourist inflow to South Korea using Google trends data”, *Journal of Travel & Tourism Marketing*, DOI: 10.1080/10548408.2016.1170651.

Richard, M. O., Chebat, J. C., Yang, Z. and Putrevu, S. (2010), “A proposed model of online consumer behavior: Assessing the role of gender”, *Journal of Business Research*, Vol. 63 No.9, pp. 926-934.

Song, H. and Li, G. (2008), “Tourism demand modelling and forecasting—A review of recent research”, *Tourism Management*, Vol. 29 No.2, pp. 203-220.

Schwartz, Z., and S. Hiemstra. (1997). “Improving the Accuracy of Hotel Reservations Forecasting: Curves Similarity Approach”, Journal of Travel Research, Vol. 36 No. 1, pp. 3-14.

Wang, Y. and Fesenmaier, D. R. (2006), “Identifying the success factors of web-based marketing strategy: An investigation of convention and visitors bureaus in the United States”, *Journal of Travel Research*,Vol. 44 No.3, pp. 239-249.

Werthner, H. and Ricci, F. (2004), “E-commerce and tourism”, *Communications of the ACM*,Vol. 47 No.12, pp. 101-105.

Wong, K. K. (1997). The relevance of business cycles in forecasting international tourist arrivals. *Tourism Management*, Vol.18 No.8, pp.581-586.

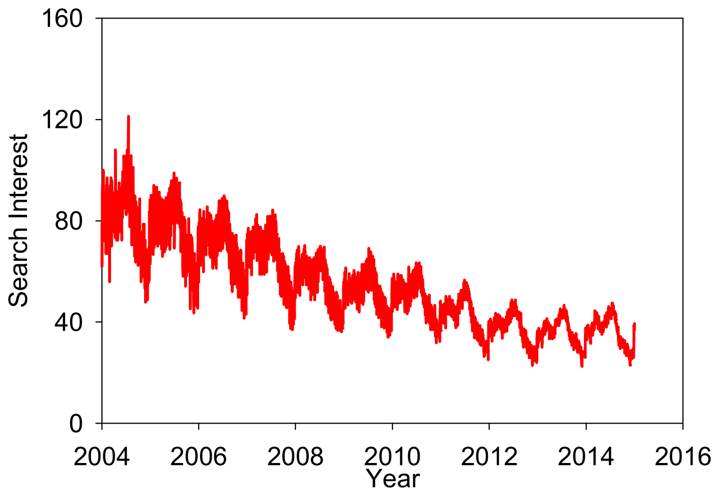
Xiang, Z.and Pan, B. (2011), “Travel queries on cities in the United States: Implications for search engine marketing for tourist destinations”, *Tourism Management*,Vol. 32 No.1, pp. 88-97.

Yang, X., Pan, B., Evans, J.A., &Lv, B. (2015), “Forecasting Chinese tourist volume with search enginedata”, *Tourism Management*, Vol. 46, pp. 386-397.

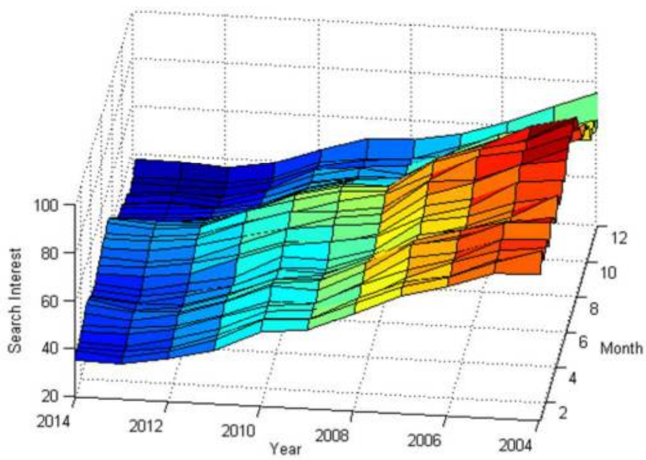
Yang, Y., Pan, B. and Song, H. (2014),“Predicting hotel demand using destination marketing organization’s web traffic data”, *Journal of Travel Research*, Vol. 53 No.4, pp. 433-447.

Zhang, Y., Jansen ,B. J., and Spink , A. (2009). “Time Series Analysisof a Web Search Engine Transaction Log.” *Information Processing & Management*, Vol. 45 No. 2, pp. 230-245.

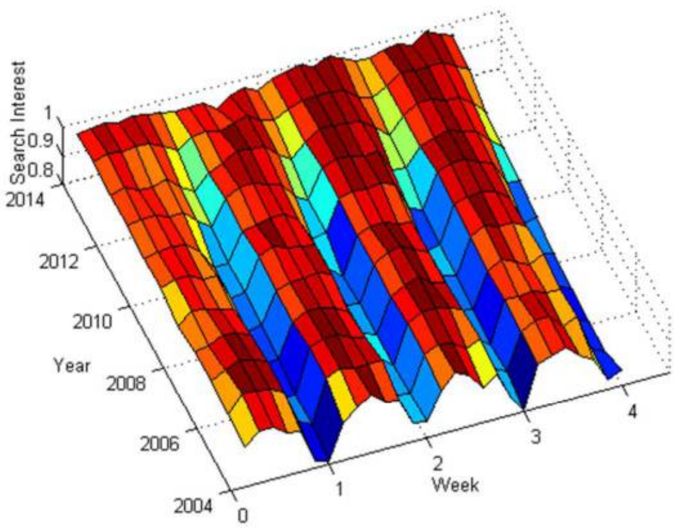
Zakhary, A., A. F. Atiya, H. El-Shishiny, and N. E. Gayar. (2009), “Forecasting Hotel Arrivals and Occupancy Using Monte Carlo Simulation”, *Journal of Revenue & Pricing Management*, Vol. 10 No. 4, pp. 344-366.


160
120
80
40
0
Search Interest
2004
2006
2008
2010
2012
2014
2016
Year

(a)

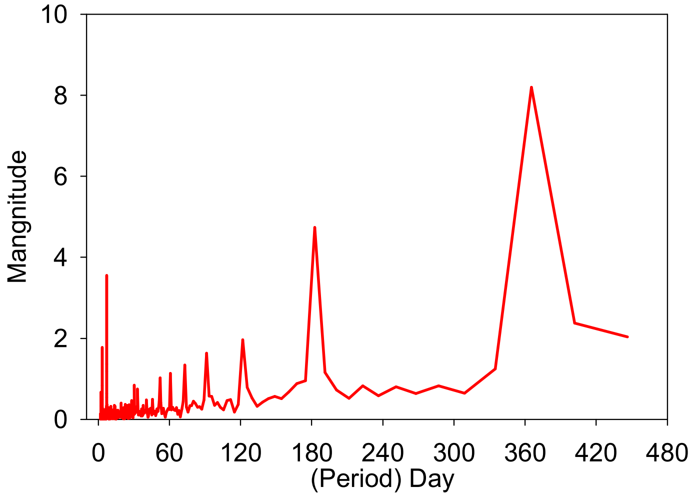

100
80
60
40
20
Search Interest
2014
2012
2010
2008
2006
2004
Year
12
10
8
6
4
2
Month

(b)

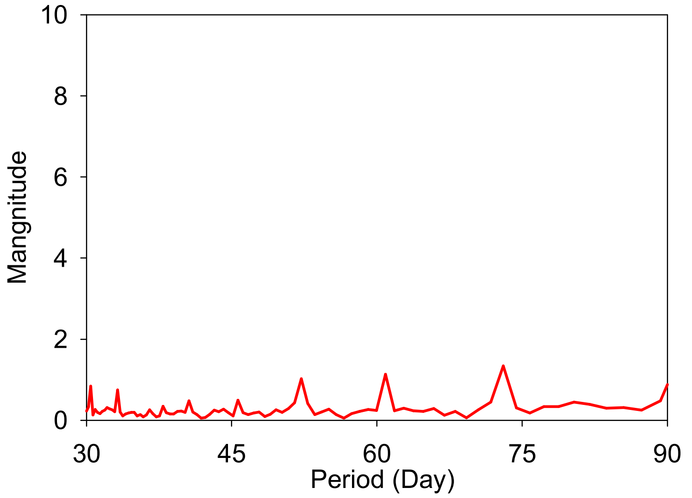

1
0.9
0.8
Search Interest
2014
2012
2010
2008
2006
2004
Year
0
1
2
3
4
Week

(c)

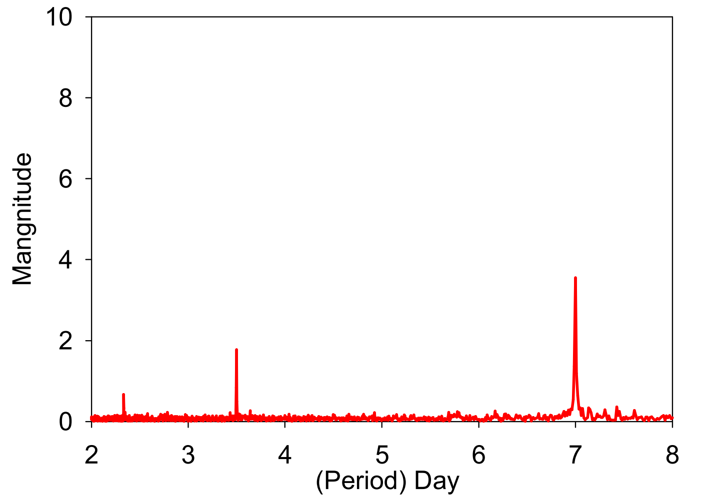
Figure 1. Google search interest on “Hotel” in the U.S. from 2004 to 2014 (a) 1D diagram for all years (b) 3D diagram for each year and each month (c) 3D diagram for 4-week average value of each year



(a)



(b)



(c)

Figure 2. Fourier amplitude spectra for hotel search interest in the U.S. from 2004 to 2014 (a) Period from 0 to 480 days (b) Period from 30 to 90 days (c) Period from 2 to 8 days

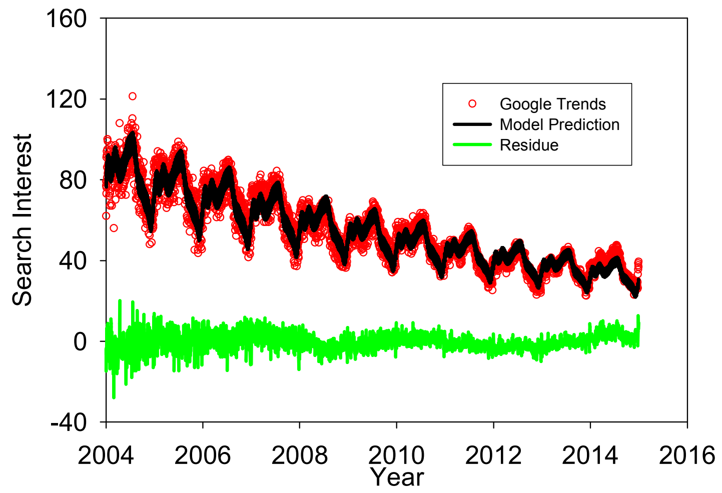


Figure 3. Model fitting for the data over the 11 years

Table 1 Major frequency components of the hotel search interest in the US from 2004 to 2014

Period	Magnitude	Rank
365.2727	8.1998	1
182.6364	4.7394	2
7.0000	3.5531	3
121.7576	1.9642	4
3.5000	1.7741	5
91.3182	1.6368	6
73.0546	1.3398	7
60.8788	1.1342	8
52.1818	1.0251	9