2019 4th International Symposium on Management, Economics, E-business and Marketing (ISMEEM 2019)

Research on Dynamic Planning of Amusement Park Based on Time Series Analysis

Ruicong Zhang[1, a], Haodong Ma[2, b]

[1]School of Statistics, Shanxi University of Finance and Economics, Taiyuan, Shanxi, China

[2]School of Finance, Shanxi University of Finance and Economics, Taiyuan, Shanxi, China

[a]1025413815@qq.com, [b]3302699226@qq.com

Keywords: Amusement Park Strategy; Time Series Analysis; Fourier Series; Queuing Theory

Abstract: Nowadays, the amusement park is the central place for lighthearted enjoyment, and it is of practical significance to predict the number of visitors to the amusement park and the corresponding waiting time to plan the tourist routes. Through time series analysis, holiday and special events are modeled while Fourier series are used to study seasonal fluctuations, to obtain the predicted number of people. Based on the queuing theory, the average waiting time and the average number of people are obtained to give the optimal planning with the most fun items, and it has reference significance for the development of amusement parks.

1. Introduction

China's economy is developing, and the government promotes the development of the tertiary industry. In recent years, the development of amusement parks has been accelerated, and many international amusement parks have been in the market in China. Most amusement park ticketing methods are divided into pass tickets and general tickets, while pass tickets can enjoy all the facilities in the amusement park, but general tickets can only enter the gate of the amusement park, and you need to buy a new ticket to enjoy any item. There are two types of rules for in-park items, including reserved and available at any time those. The closing time varies with summer and winter as well as indoor and outdoor places. The data from January 1, 2018 to May 1, 2019 were recorded, and it includes the number of tickets sold for 517 days, the number of receptions for 38 amusement projects, and the charge standard of each item, the number of people accommodate and the time of a single visit.

2. Introduction to Theory

The types of amusement park items, the duration, the way of selling tickets, and the situation of the sale of amusement park tickets and pass tickets are known. Time series is used to predict the future tickets and pass tickets of the amusement park with time as the independent variable. To study the number of visitors queuing and the average waiting time of each project, it is necessary to establish a queuing model of the single queue and single service desk. Finally, the best route for tourists is changed into the shortest path problem with time as the weight and a reasonable route is obtained according to Dijkstra algorithm[1].

2.1 Time Series Analysis and Prophet Algorithm

Time series analysis is a statistical method of dynamic data processing[2]. The Prophet algorithm is adopted in this paper. The prophet algorithm is a prediction tool based on time series decomposition and machine learning fitting and can get the predicted results in a relatively fast time. The open-source algorithm can be used to predict the future trend of time series for general business analysis or data analysis needs. For the input and output of Prophet algorithm, achieve: (1) plugin the timestamp of the known time series and the corresponding value; (2) input the length of time series to be predicted; (3) output future trend of time series; (4) the output results can provide

Copyright © (2019) Francis Academic Press, UK DOI: 10.25236/ismeem.2019.134

necessary statistical indicators, including fitting curve, upper bound and lower bound, etc[3].

Generally speaking, in addition to the seasonal term, trend term and remaining term, there is usually the effect of holidays. Considering the above four terms, we assume that the past of things will be carried forward into the future and calculate the predicted value of time series by addition principle. The predicted value of time series is: $\mathrm{y(t)} = \mathrm{g(t)} + \mathrm{s(t)} + \mathrm{h(t)} + \epsilon_t$

The piecewise logistic regression growth model is:

$$\mathrm{g(t)} = \frac{\mathrm{C}(t)}{1 + \exp\left(-\left(k + \mathbf{a(t)}^T \boldsymbol{\delta}\right) \cdot \left(t - \left(m + \mathbf{a(t)}^T \boldsymbol{\gamma}\right)\right)\right)}$$

Indicator function $\mathbf{a(t)} = \left(a_1(t), a_2(t), \cdots, a_S(t)\right)^T$, $a_j(t) = \begin{cases} 1, & \text{if } t \geq s_j \\ 0, & \text{otherwise} \end{cases}$. δ_j represents the change in the increasing rate at s_j in time; γ_j s the setting of the boundary in the line segment.

And the Fourier series would be $s(t) = \sum_{n=1}^{N} \left(a_n \cos\left(\frac{2\pi nt}{P}\right) + b_n \sin\left(\frac{2\pi nt}{P}\right) \right)$

If the number of holidays is L, the holiday effect is: $\mathrm{h(t)} = \mathrm{Z(t)}\mathbf{k} = \sum_{i=1}^{L} k_i \cdot 1_{\{t \in D_i\}}$ Meanwhile, $Z(t) = \left(1_{\{t \in D_1\},} \cdots 1_{\{t \in D_L\}}\right)$ is the indicator function.

$g(t)$ represents the trend, and the changes in the trend of time series in the aperiodic period; $s(t)$ represents the periodic one, or seasonal one, and observed data are generally in quarterly, monthly and weekly; $h(t)$ represents the holiday, indicating whether the day is a holiday or not; $\varepsilon(t)$ represents the error or the remaining while the Prophet algorithm will fit items and eventually add them up to get the predicted value of the time series, and it is needed to make prediction when the peak occurs.

2.2 Queuing Theory

Queuing theory is a theoretical mathematical method to study the random clustering and dispersion phenomenon of the system and the working process of the random service system [4]. Also known as stochastic service system theory and it is based on studying the probability regularity of various queuing systems and solving the problem of optimal design and optimal control of corresponding queuing systems.

It is assumed that the time interval between the successive arrival of customers and the service time are random in the composition of this arrangement system. There are three parts, as follows:

(1) Input Process: tourists are the customers served, and the total number of customers is limited, not more than the number of each project. The arrival of customers is completely random and has no memory, and it is concluded that the arrival time interval of any two customers obeys the k-order Erlang distribution of order 2 [5], which is denoted as E_k. The distribution function is:

$$f(t) = \frac{k\lambda(k\lambda t)^{\mathrm{k}-1}}{(\mathrm{k}-1)!} e^{-k\lambda t}$$

(2) Queuing rules: first come, first go, and the number of tourists for each amusement facility every day is limited. Therefore, it is the FCFS limited queuing mechanism.

(3) Service organization: Since there is only one amusement facility, the number of service desks is obviously 1, and every time the amusement project runs, tourists are accepted in batches, while each service time is fixed, so it is considered to be a fixed-length distribution. Customer departure rate μ, that is, the number of customers leaving per unit of time, can be regarded as $\mu = E(\mu), \mathrm{D}(\mu) = 0$. So, in this case, it belongs to the $\mathrm{E_k / D / 1}$ queuing model of single queue and single service desk. As tourists enter each project in batches, the visitor arrival rate λ_1 and service

rate μ_1 are :

$$\lambda_1 = \frac{M}{60T_i} \qquad \mu_1 = \frac{M^{'}}{t}.$$

M is the number of receptions for each project on August 5; T_i is the sum of the opening time of each game and the opening time of the amusement park before the game; M' is the maximum number of people to accommodate for each project; t is the time for each round. According to the book "queuing theory and the preparation of the simulation program in its application" written by Mr. Lu fengshan[6], we obtained:

$$W_{team} = \frac{\rho^{\alpha}}{2\mu\left(1-\rho^{\alpha}\right)} \qquad .\alpha=2 \qquad L_{team} = \lambda W_{team}.$$

2.3 Dijkstra Algorithm

Before the beginning of the amusement park, we queued up in advance, assuming that the problem of the play sequence route satisfies the optimization principle and the most effective, so the optimal planning problem becomes the solution to the problem of multi-project planning. Assuming that people have the same preference for each item, however, Dijkstra's algorithm mainly solves the shortest path problem of weighted graphs with two specified vertices, while the core idea of weight graph $G=(V, E, W)$ is to obtain the shortest path and distance according to the distance from the starting point to each vertex from near to far [7].

The walking time from one amusement project to another is t_1, the waiting time is t_2; the playtime t_3 is added as the total time as the weight of the edge $w = t_1 + t_2 + t_3$; where v_{walk} are obtained by conversion of walking speed index[8].

3. Empirical Analysis

Establish a time series analysis model, analyze the relevant data of a specific amusement park and substitute it into the calculation. The prophet algorithm was used to predict the sales of tickets and passes on August 5, 2019, and the peak was found during holidays, summer holidays, and weekends. Based on the queuing theory model, the queuing number and waiting time of each item are calculated. Finally, according to the queuing data obtained, Dijkstra algorithm is used to give the optimal planning with the most items, and 17 items can be used on average in a day.

3.1 Data Introduction and Basic Statistical Analysis

The data in this paper were obtained from the data of question A of the 2019 Shanxi provincial college students' mathematical modeling contest. These data include 517 days of tickets and passes sold from January 1, 2018 to May 1, 2019, the number of receptions for 38 amusement events, the charging standard for each event, the number of people accommodated, and the time for a single visit.

Firstly, data preprocessing is carried out for the given data, and it can be seen that there are missing values. In this paper, the null values generated by usual factors are removed, and the remaining missing values are interpolated, using the sliding average window method, that is, the ith position data in a list a is missing data, then take the average value of a given data in the region before and after interpolation, and obtain the data table after interpolation for analysis, and take the number of pass tickets sold after data cleaning as an example for statistical description:

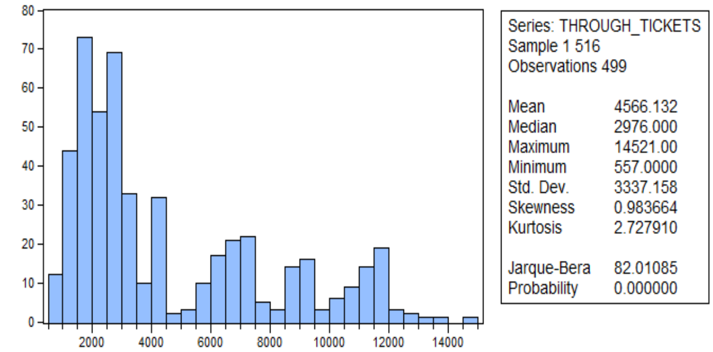

Figure 1 Statistical description of the number of pass tickets sold

FIG.1 shows that the maximum value is 14521 and the minimum value is 557, and the skewness coefficient is greater than 0, indicating asymmetric distribution.

3.2 Time Series Analysis

The components of the time series are the time of the phenomenon and the index value of the development level of the phenomenon. In this problem, time is taken as the independent variable, and the number of passes and tickets sold is taken as the dependent variable over time. Using the Python version of Prophet algorithm to optimize the time series prediction, get prediction data.

The effects of individual factors, trends, seasonality, and holiday effects were analyzed separately. It is found that a single deterministic factor (long-term trend fluctuation or seasonal effect) can affect the sequence. Furthermore, the interaction between various deterministic factors and their comprehensive influence on the sequence are deduced. Based on the above Prophet algorithm, run the program to predict daily passes and ticket sales for the second half of 2019, and the graph based on the true and predicted values can be obtained. Take pass tickets as an example:

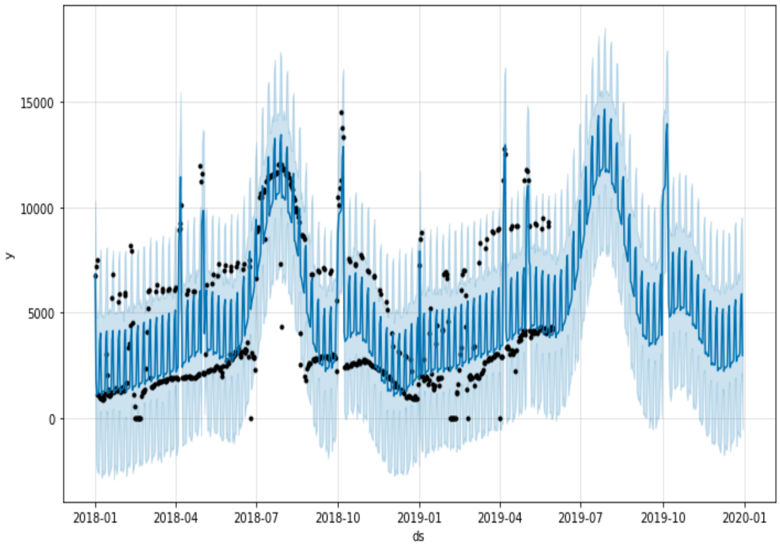

Figure 2 Forecast number of pass tickets sold

In figure 2, black represents the original discrete points of the time series; the dark blue line represents the value obtained by using the time series fitting, while the light blue line represents a confidence interval of the time series. There are peaks in the summer holidays and small peaks in the holidays. During the two days of the weekend, the weekend effect is noticeable. The latter part is the predicted value, which is based on the actual forecast curve of the second half of 2019. A seasonally induced spike will also occur in 2019. The fitting result is very close to the truth value.

3.3 Queuing Theoretic Models

Monday, August 5, 2019, during the summer vacation, and it does not belong to a particular holiday, which makes the model meaningful and improves its universality. Predicted the number of visitors for each project, and calculated the average waiting time and the average number of visitors for each project with the queuing theory model. As shown in table 1:

Table 1 The chart of total number and average waiting

Project	Total	Duration	Waiting people	Project	Total	Duration	Waiting people
A1	5362	48	411	B9	3000	49	238
A2	6222	74	669	B10	6791	52	571
A3	9530	22	308	B11	9562	147	2241
A4	9418	6	93	B12	341	0	0
A5	8106	119	1535	B13	299	0	0
A6	10037	33	488	B14	9319	97	1409
A7	6854	37	377	B15	8950	47	621
A8	8800	146	1870	B16	5548	36	324
A9	7861	112	1282	B17	7181	41	475
A10	5704	63	593	B18	3600	60	342
A11	9120	101	1502	B19	10859	25	405
B1	7790	114	1293	B20	9788	130	1852
B2	8238	18	246	B21	979	0	0
B3	9528	138	2088	B22	8186	18	218
B4	10200	169	2752	B23	9528	138	2088
B5	10463	26	447	B24	8846	28	368
B6	8800	146	1870	B25	8034	17	224
B7	8745	6	80	B26	8551	19	246
B8	10742	94	1604	B27	10200	169	2752

As can be seen from table 1, people's low willingness to enjoy may be caused by the following reasons: fewer people, the project is limited or expensive, high risk, the project itself is not interesting. However, due to the fact that there is no limit on the number of people in the project or the number of people can be accommodated at one time, such projects generally do not need to wait in line.

3.4 Dijkstra Arithmetic

From the perspective of tourists, explore the routes to have fun in more items at the same time. For the optimal path, the starting point is set as an A4 facility near the amusement park gate with an average waiting time of 6 minutes. The total walking time, waiting time and playing time from one item to another are taken as weights. Divide the amusement projects into two groups, one is marked amusement projects, and the other is unmarked amusement projects. Then, in the unmarked items, find the amusement items closest to the starting point A4 in the order from near to far, and put each amusement item into the marked amusement items. After Matlab software programming calculation, if time is enough, the play routes of 38 amusement projects are as follows: A4 - B21 - B12 - B23 - B25 - B2 - B22 B26 - B7 - B19 to B16 - A3 - B5, B15 - B9 - B24 - A6 - A7 - B17 - A1 - B18 - B10 - - A2 - B14 A10 - B8 - B1 - A11 - A9 - A5 - the B20 - B3 - B23 - bl1 - A8 - B4 to B6, B27

The amusement park is open for 11 hours a day. During the 11 hours, the first 17 items can be visited on average. This is given the optimal planning route that most people would choose, However, considering the difference of people, each amusement park can provide the average waiting time, allowing tourists to choose and give the best route by reference.

4. Conclusions and Prospects

4.1 Conclusion

(1) It uses the time series prophet algorithm to get the predicted number of people, which will

peak in summer vacation and peak in holidays. During the two days of the weekend, the weekend effect is obvious. During New Year, National Day and other holidays, there will be a sudden change, the flow of holidays accounted for the vast majority of the annual flow, in line with the phenomenon of economic life. Compared with the traditional time series prediction method, the prophet algorithm performed multi-part optimization and separately analyzed and expressed the effects of various factors.

(2) The average waiting time was calculated based on the queuing theory model to explore the possible reasons for the emergence of items with the short average waiting time. In the queuing theory, the arrival time interval of any two customers obeys k-order Erlang distribution, which is close to the actual situation and can accurately predict the average waiting number and time.

(3) The optimal plan for the most fun items is given. On average, 17 items can be enjoyed in a day. Accordingly, each amusement park can give the average waiting time of each project selectively for tourists to choose independently, and take the best route as the reference.

4.2 Prospects

(1) When the maximum value of the model proposed in this paper is given, the result is biased due to the influence of the weekend effect. During National Day, the number of tourists decreased. We can try to use ARIMA model to comprehensively compare the two or select other more effective factors to optimize the model.

(2) When figuring out the shortest path, the weight of the edge is relatively obvious to determine, and the utility and time cost of each tourist project can be combined to obtain the score. The difficulty is in the difficulty of quantifying the diminishing marginal utility of tourists and the limited amount of data. After solving these two problems, we can build a wavelet neural network and train a network according to the data set. After input the corresponding parameters, the utility value can be obtained through the network. Finally, the time cost is combined to get the score to get the weight, which can make the model more accurate.

References

[1] Wagner, Anita K., et al. "Segmented regression analysis of interrupted time series studies in medication use research." Journal of clinical pharmacy and therapeutics 27.4 (2002): 299-309.

[2] Zhang Yanjun, Yang Xiaodong, Liu Yi, Zheng Dayuan, Bi Shujun. Research on the Frame of Intelligent Inspection Platform Based on Spatio-temporal Data. Computer & Digital Engineering [J], 2019, 47(03): 616-619+637.

[3] Yang, Kun, Meie Pan, Yi Luo, Kexin Chen, Yisong Zhao, and Xiaolu Zhou. "A time-series analysis of urbanization-induced impervious surface area extent in the Dianchi Lake watershed from 1988–2017." International journal of remote sensing 40, no. 2 (2019): 573-592.

[4] Lang, Kun, Mingyuan Zhang, Yongbo Yuan, and Xijian Yue. "Short-term load forecasting based on multivariate time series prediction and weighted neural network with random weights and kernels." Cluster Computing (2018): 1-9.

[5] Yi Liu, Jiawen Peng, and Zhihao Yu. 2018. Big Data Platform Architecture under The Background of Financial Technology: In The Insurance Industry As An Example. In Proceedings of the 2018 International Conference on Big Data Engineering and Technology (BDET 2018). ACM, New York, NY, USA, 31-35.

[6] Z. Zhao, J. Wang and Y. Liu, "User Electricity Behavior Analysis Based on K-Means Plus Clustering Algorithm," 2017 International Conference on Computer Technology, Electronics and Communication (ICCTEC), Dalian, China, 2017, pp. 484-487. doi: 10.1109/ICCTEC.2017.00111

[7] Deng, Chengbin, and Zhe Zhu. "Continuous subpixel monitoring of urban impervious surface using Landsat time series." Remote Sensing of Environment (2018): 110929.

[8] Bechtel, Benjamin, and Panagiotis Sismanidis. "Time series analysis of moderate resolution

land surface temperatures." In Remote Sensing Time Series Image Processing, pp. 111-142. CRC Press, 2018.

[9] Y. Wu, Y. Liu, S. H. Ahmed, J. Peng and A. A. A. El-Latif, "Dominant Dataset Selection Algorithms for Electricity Consumption Time-Series Data Analysis Based on Affine Transformation," in IEEE Internet of Things Journal. doi: 10.1109/JIOT.2019.2946753

[10] Arvor, Damien, Felipe RG Daher, Dominique Briand, Simon Dufour, Anne-Julia Rollet, Margareth Simões, and Rodrigo PD Ferraz. "Monitoring thirty years of small water reservoirs proliferation in the southern Brazilian Amazon with Landsat time series." ISPRS journal of photogrammetry and remote sensing 145 (2018): 225-237.

[11] Zhang, Tao, and Xin Huang. "Monitoring of urban impervious surfaces using time series of high-resolution remote sensing images in rapidly urbanized areas: A case study of Shenzhen." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 11, no. 8 (2018): 2692-2708.