Physica A 491 (2018) 887–897

Contents lists available at ScienceDirect

Physica A

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

RCA: A route city attraction model for air passengers

Feihu Huang [a], Xi Xiong [b], Jian Peng [a,*], Bing Guo [a], Bo Tong [a]

[a] College of Computer Science, Sichuan University, Chengdu 610065, PR China
[b] College of Information Security Engineering, Chengdu University of Information Technology, Chengdu 610225, PR China

HIGHLIGHTS

- A route city attraction model for air passengers is proposed.
- Air passenger travel has a balance phenomenon.
- The distance is no longer the influencing factor of air passenger movement.
- The probability that a route is selected is derived.

ARTICLE INFO

Article history:
Received 26 March 2017
Received in revised form 12 June 2017
Available online 1 September 2017

Keywords:
Human mobility
Air passenger
Throughput prediction
Aggregate model

ABSTRACT

Human movement pattern is a research hotspot of social computing and has practical values in various fields, such as traffic planning. Previous studies mainly focus on the travel activities of human beings on the ground rather than those in the air. In this paper, we use the reservation records of air passengers to explore air passengers' movement characteristics. After analyzing the effect of the route-trip length on the throughput, we find that most passengers eventually return to their original departure city and that the mobility of air passengers is not related to the route length. Based on these characteristics, we present a route city attraction (RCA) model, in which GDP or population is considered for the calculation of the attraction. The sub models of our RCA model show the better prediction performance of throughput than the radiation model and the gravity model.

1. Introduction

With mobile devices and GPS systems, it is easy to record the data generated in travel activities for exploring human space movement [1]. Studies on human space movement, in theory, help us to explain the complex social phenomena driven by human movement behaviors and understand the laws of human movement. Study results of human space movement can be applied in urban planning, traffic flow forecasting, epidemic diffusion, recommender systems [2–6].

In the exploration of human travel, it is necessary to establish a model to characterize human movement characteristics. Human mobility models have been extensively developed since 2006. These models can be classified into two categories. The first type of models are built on the characteristics of human mobility, such as continuous time random walk (CTRW) model [7], self-similar least action walk (SLAW) model [8], exploration and preferential return (EPR) individual mobility model [9], and task-based model [10]. The second type of models are based on geospatial information, such as gravity model [11], radiation model [12], population-weighted opportunities (PWO) model [13], rank-distance model [14] and the model for mobile phone users [15]. However, the data mainly used in these studies include currency flow data, mobile phone location data, and taxi trajectory data, which are associated with ground travel activities.

* Corresponding author.
E-mail address: jianpeng@scu.edu.cn (J. Peng).

http://dx.doi.org/10.1016/j.physa.2017.08.081
0378-4371/

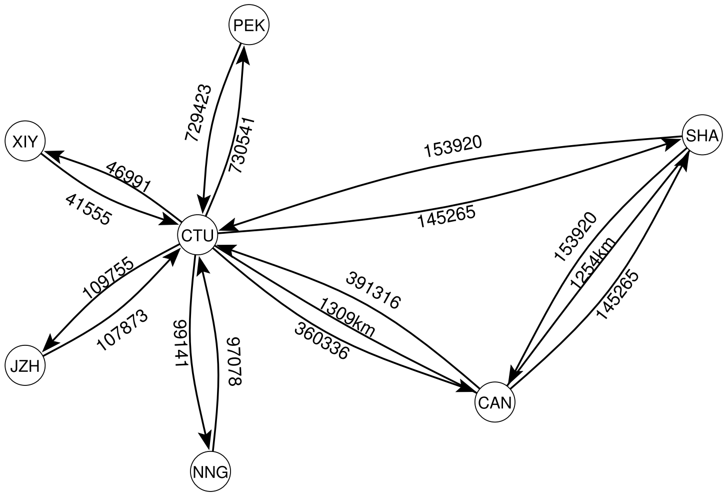

Fig. 1. Illustration of the simplified route.

Due to the development of air traffic and the growth of national economy, more passengers would like to travel by plane. Mining and analysis of the travel data of civil aviation has attracted wide attention. The studies on aviation data mainly focus on throughput forecasting [16], the influencing factors of throughput [17], aviation networks [18], etc. Some researchers studied air passenger mobility characteristics. Jiang et al. explored human mobility patterns based on massive tracking data of US flights and found that airports traffic volume follows a power-law distribution and that the travel lengths exhibit an exponential distribution [19]. Huang et al. used the reservation record data of civilian aviation passengers to analyze the characteristics of air passengers group mobility behaviors and discovered that passengers prefer to book the tickets during the holidays in advance [20]. Travel by plane is a long-distance movement, which is a space interaction process. Unlike the daily movements within city, air passengers' movement is not only influenced by individual factors but also the economic relationship among regions [1,21]. As shown in Fig. 1. It is a simplified route network and only describes the throughput relationship between the Chengdu Airport and other six airports. The number on the directed arc is the amount of route throughput. Although the route length between Chengdu and Guangzhou is longer than the distance between Shanghai and Chengdu, the throughput of route between Chengdu and Guangzhou is larger than that between Chengdu and Shanghai. As we can see, there is a stronger regional economic relationship between Chengdu and Guangzhou. Likewise, the throughput of route between Chengdu and Jiuzhai is twice as large as that between Chengdu and Xi'an. Because Jiuzhai is a well-known tourist spot, people would like to go there. So air passengers' movement Characteristics may be different from the characteristics of human movement on the ground. Study on air passengers' movement can help us to better understand the mechanisms of human mobility, and arrange the flights properly.

In this paper, we use reservation record data of air passengers to explain air passengers' movement characteristics and build a model to characterize the movement patterns of air passengers. Experimental results show that our model has the good prediction performance. The rest of this paper is organized as follows. Air passengers' movement characteristics are introduced in Section 2. The model is introduced in Section 3. The discussion is provided in Section 4. The conclusion is drawn in Section 5.

2. Movement characteristics

Human movement has the scale-free characteristic [22]. Air passenger travel also has the same characteristic. We analyzed the characteristics of air passengers' group movement behaviors and found that air passenger travel has scale-free characteristic [20]. The travel length distribution is more consistent with the stretched exponential distribution and the travel interval time follows a truncated power-law distribution. Besides, air passenger travel may have its own unique characteristics, which are different from those of human movement on the ground. Therefore, in this paper, we use reservation record data of air passengers to explore their movement characteristics. Table 1 shows the relevant socioeconomic information of the cities corresponding to these airports. For the purpose of clear descriptions, R_{A-B} denotes the route from City A to City B, R_{B-A} denotes the route from City B to City A. We call R_{A-B} and R_{B-A} "round-trip routes". T_{A-B} denotes the throughput of R_{A-B}. T_{B-A} denotes the throughput of R_{B-A}. For simplicity, we use the International Air Transport Association (IATA) code instead of the airport name in the figure (Figs. 1 and 8). Some airports' IATA codes are listed in Table 1. The air passengers' travel characteristics are summarized below.

Table 1
Population, GDP and route lengths of Chengdu and other six cities.

	Chengdu (CTU)	Beijing (PEK)	Xi'an (XIY)	Jiuzhai (JZH)	Nanning (NNG)	Shanghai (SHA)	GuangZhou (CAN)
Population (million)	14.42	21.52	8.62	0.92	7.29	24.15	12.71
GDP (billion)	1005.6	2133.1	547.5	24.8	314.8	2356	1670.7
Route length from CTU		1752.6 km	685.5 km	254.6 km	1002.3 km	1950 km	1309 km
Route length from CAN	1309 km					1254 km	

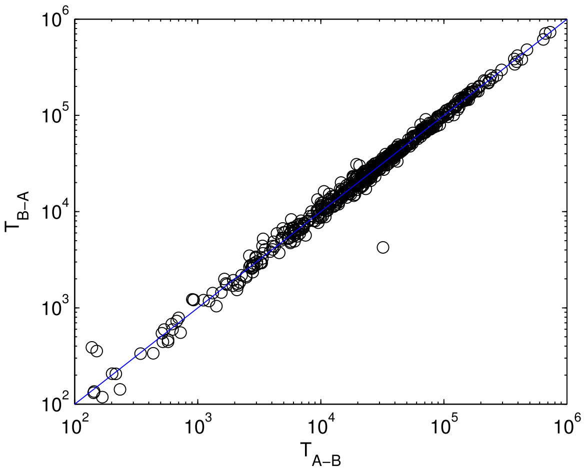

Fig. 2. Throughputs of round-trip routes.

2.1. Balance phenomenon

In the analysis of the throughputs of round-trip routes, a balance phenomenon of air passenger travel is found. As indicated in Table 1, economic indicators of these cities are different. However, the throughputs of round-trip routes are almost equal, such as $T_{Chengdu-Jiuzhai}$ and $T_{Jiuzhai-Chengdu}$ (Fig. 1). Jiuzhai Airport is located in the county of Jiuzhai, which is a well-known tourist spot. Chengdu airport is located in Chengdu, Sichuan Province. The population and the GDP of Chengdu are much larger than Jiuzhai's. However, $T_{Chengdu-Jiuzhai}$ and $T_{Jiuzhai-Chengdu}$ are 109,755 and 107,873 people, respectively. In order to verify whether this was an accidental phenomenon, we analyze the throughputs of all the round-trip routes in the dataset. Fig. 2 shows the statistical results. These points are close to the line y=x, indicating that the throughputs of the round-trip routes are balanced. Then we analyze the passenger travel records. The phenomenon that most passengers eventually return to the departure city is found, which explains the reason why the throughputs of round-trip routes maintain a balance. This is not the same as the radiation model's result that the travel flows between two regions are different if the two regions have different economic indicators.

2.2. Effects of route length on throughput

Human travel activities on the ground can be effected by the distance between the origin and destination [11,13,21]. Here we analyze the effects of route length on throughput to verify whether the distance is also a factor of air passenger travel. As shown in Fig. 1, the route lengths of $R_{Chengdu-Guangzhou}$ and $R_{Guangzhou-Shanghai}$ are 1309 km and 1254 km, respectively. Shanghai's population is more than Chengdu's (Table 1). If we consider the route length and city attraction (like the gravity model), $T_{Chengdu-Guangzhou}$ should be smaller than $T_{Guangzhou-Shanghai}$, but the fact is $T_{Chengdu-Guangzhou}$ is almost 3 times of $T_{Guangzhou-Shanghai}$. Here we use the population as a measure of the attraction to discuss the relationship between route length and throughput. Similarly, with the GDP as a measure of quality, we can obtain the same result. Fig. 3 shows a scatter plot of route length and route throughput. Obviously, there is no inverse relationship between route length and route throughput, as described in the gravity model.

The travel characteristics of air passengers are different from human traffic activities on the ground. The distance is no longer the influencing factor of air passenger travel. Furthermore, if passengers depart from City A, they will finally return to City A. Therefore, in our model, we consider this factor and avoid the distance factor.

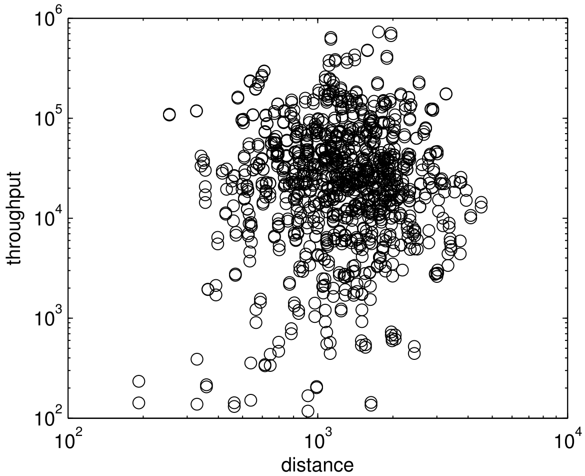

Fig. 3. Route length and throughput.

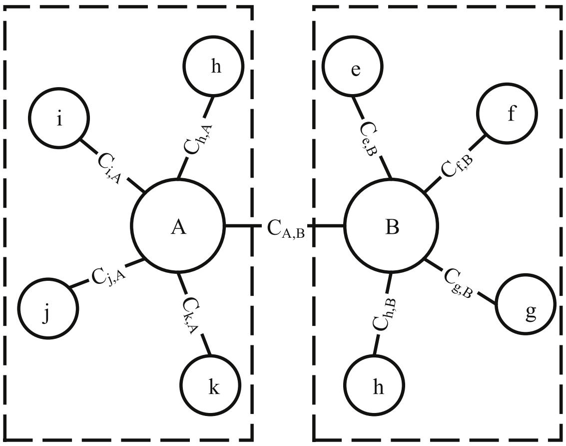

Fig. 4. Route competition.

3. Model

3.1. Route city attraction model

Human travel is the process of individual's destination selection. She/He goes to a place because of work (such as meeting customers and business meetings) or life (such as visiting friends and shopping). Here we mainly consider long-distance movements by plane. Such movements are the long-range and large-scale activities among cities. In this paper, we consider the route competition among cities to estimate the probability of which route passengers would like to choose. Fig. 4 illustrates the main idea. In route competition, three issues should be solved.

i. Selection probability. The selection probability is the probability that a route is selected. The throughputs of round-trip routes are almost equal, indicating that air passengers' movement is balanced. In our model, we believe that $P_{A-B} = P_{B-A}$, which means the probability passengers choose route from City A to City B is equal to the route from City B to City A. Route AB attracts the passengers of City A and City B, so it competes with those routes involving City A and City B. According to Fig. 4, we have

$$P_{A-B} = P_{B-A} = \frac{C_{A,B}}{\sum_{i\in N_A, i\neq B} C_{i,A} + \sum_{j\in N_B, j\neq A} C_{j,B}} \tag{1}$$

ii. Competition. The competition ultimately reflects in the route. We use $C_{A,B}$ to denote the competitiveness of $R_{A,B}$. The more developed the economy of the two cities is, the more frequent they communicate with each other. In this case, the route between them is more likely to be chosen by passengers. In other words, the competitiveness of a route is related to the two cities. F_A and F_B are the attraction of City A and City B, respectively. Then, we have

$$C_{A,B} = F_A \cdot F_B \tag{2}$$

iii. City attraction. The more developed economy of a city indicates the greater attraction and the higher commercial value. Besides, a city's attraction is proportional to its population , like the assumption in the radiation model [12] and the PWO model [13]. So, we can use GDP and population to measure a city's attraction.

In the radiation mode, the number of people is the total commuters start from location i (see Eq. (7)). For the two cities corresponding to a route, the city with the larger attraction will make this route to attract more people and the city with the smaller attraction determines the base number of passengers. Here we use function $f(\cdot,\cdot)$ to estimate the number of passengers between city A and B. The function $f(\cdot,\cdot)$ is:

$$f(T_A, T_B) = a + bT_{max} + cT_{min} \tag{3}$$

T_{max} denotes the throughput of the city with the larger attraction, T_{min} denotes the throughput of the city with the smaller attraction. We can use least square method to learn the values of a, b, and c. The error function is:

$$Q = \sum \left(f_i - \hat{a} - \hat{b}T^i_{max} - \hat{c}T^i_{min}\right)^2 \tag{4}$$

where $\hat{a}$, $\hat{b}$ and $\hat{c}$ are the estimated values of a, b and c, respectively. f_i is the number of passengers between city A and B. According to the least square method, the parameters $\hat{a}$, $\hat{b}$ and $\hat{c}$ should minimize the value of error function Q. We get the parameter calculation formula:

$$\begin{cases} \hat{a} = \bar{f} - \hat{b}\hat{T}_{max} - \hat{c}\hat{T}_{min} \\ \hat{b} = \dfrac{\sum(\dot{T}^i_{max}\dot{f}_i)\sum(\dot{T}^i_{min})^2 - \sum(\dot{T}^i_{min}\dot{f}_i)\sum(\dot{T}^i_{max}\dot{T}^i_{min})}{\sum(\dot{T}^2_{max})^2\sum(\dot{T}^i_{min})^2 - \sum(\dot{T}^i_{max}\dot{T}^i_{min})^2} \\ \hat{c} = \dfrac{\sum(\dot{T}^i_{min}\dot{f}_i)\sum(\dot{T}^i_{max})^2 - \sum(\dot{T}^i_{max}\dot{f}_i)\sum(\dot{T}^i_{max}\dot{T}^i_{min})}{\sum(\dot{T}^2_{max})^2\sum(\dot{T}^i_{min})^2 - \sum(\dot{T}^i_{max}\dot{T}^i_{min})^2} \end{cases} \tag{5}$$

where $\dot{T}^i_{min} = T^i_{min} - \bar{T}^i_{min}$, $\dot{T}^i_{max} = T^i_{max} - \bar{T}^i_{max}$, $\dot{f}_i = f_i - \bar{f}$. $\bar{T}_{min}$, $\bar{T}_{max}$ and $\bar{f}$ are the mean of T_{min}, T_{max} and f, respectively.

Now we have the selection probability and the estimated number of passengers. We can use statistical approach to generate the throughput of T_{A-B}. It is a binomial distribution. Passengers will choose the route or not with the selection probability. When the departure throughput of Airport A is T_A and the departure throughput of airport B is T_B, the average throughput of T_{A-B} is:

$$\langle T_{A-B}\rangle = f(T_A, T_B) \cdot P_{A-B} \tag{6}$$

The proposed model reflects the process of individual's destination selection: if City A is more attractive than other cities, it has a greater advantage over other cities. When the attractions of the two cities of a route are greater, the competitiveness of the route is stronger and the route throughput becomes larger. The route throughput is related to the attractions of the two cities of a route. We therefore name our model the route city attraction (RCA) model. As shown is Figs. 5 and 6, we can use the function of population and GDP to calculate the throughput approximatively. They have positive correlations with throughput. For simply, in the RCA model, we use population and GDP to calculate the attraction, respectively. When GDP is used to calculate the attraction, the RCA model is named the GRCA model. In the same way, when population is used to calculate the attraction, the RCA model is named the PRCA model. We then will demonstrate the performance of the GRCA model and the PRCA model with real travel data. In the simulation experiment, we use 10-fold cross validation to learn the values of parameters $\hat{a}$, $\hat{b}$ and $\hat{c}$. Then we calculate the mean values of their learning results as their estimated values. In the PRCA model, $\hat{a} = -1861561.1$, $\hat{b} = 1.09603$, $\hat{c} = 3.89088$. In the GRCA model, $\hat{a} = -754195.85$, $\hat{b} = 1.03846$, $\hat{c} = 2.53783$.

3.2. Model performance

3.2.1. Throughput prediction

Throughput prediction involves route throughput prediction and airport throughput prediction. The route throughput is the total throughput of the flights from departure city to destination city. The airport throughput is the number of passengers arriving at the airport. In Fig. 7, the horizontal axis indicates the empirical data and the vertical axis indicates the prediction result obtained with the model. The better prediction effect of the model indicates that the scatter points are closer to the line $y = x$. Each figure is a line box plot of data points and the blue point is the mean of each bin. We use the color of the line box map to distinguish whether the line $y = x$ passes the range of each bin. A box is marked in light green if the line $y = x$ lies between 10% and 90% in that bin and in red otherwise. As shown in Fig. 7, the line boxes of both the radiation model

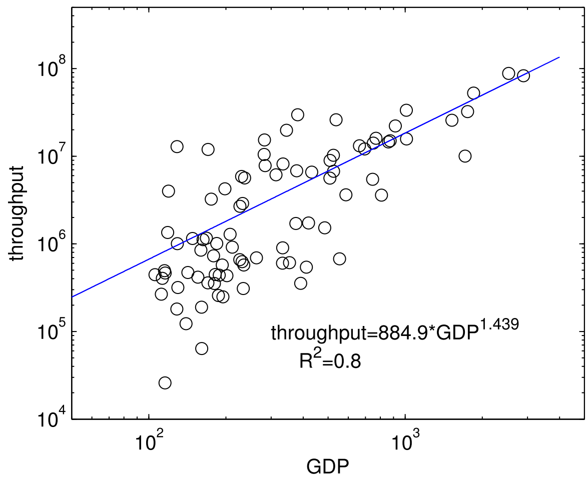

Fig. 5. GDP and route throughput.

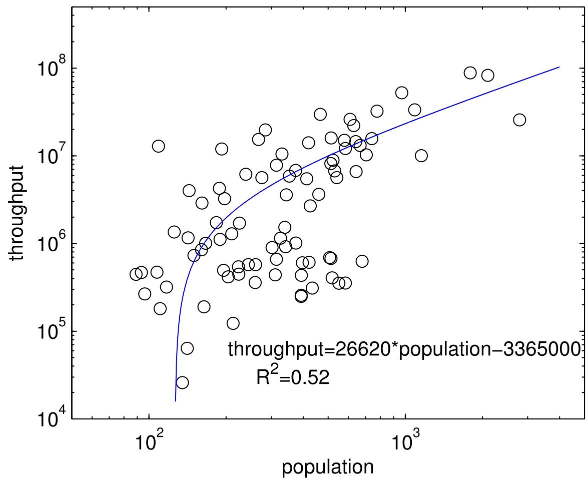

Fig. 6. Population and route throughput.

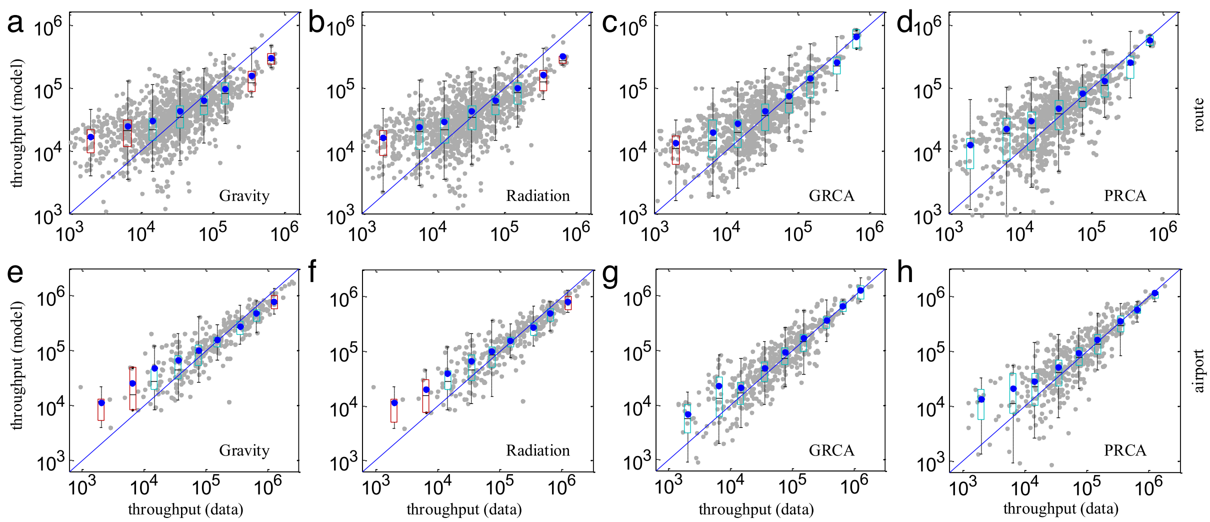

Fig. 7. Route competition.

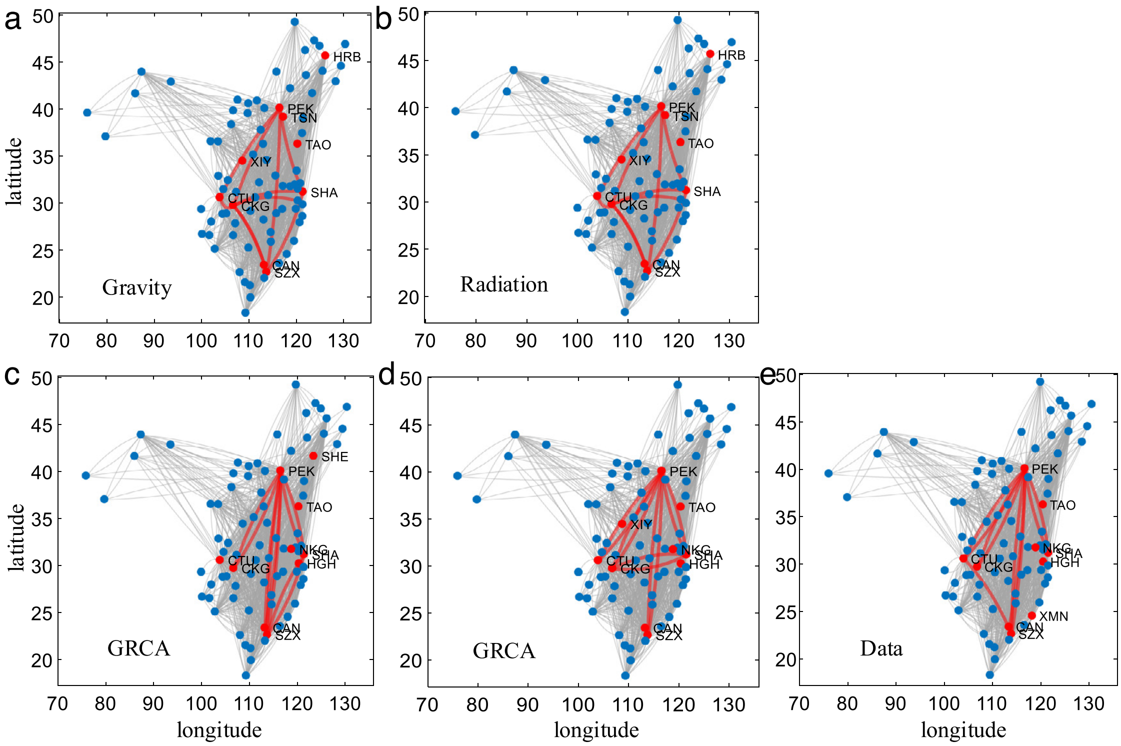

Fig. 8. Popular airports and routes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

and gravity model are red when airport (route) throughput is small and large. In our sub models, line boxes are almost green except that the GRCA model has a red line box when the route throughput is small. The results imply that our sub models have the better performance in throughput prediction.

3.2.2. Top-k prediction

In the recommender systems, the top-k recommendation is an important indicator of the model. For human travel activities, we also pay attention to the hot spots. Here we show the prediction results of the three models' popular cities and popular routes. Fig. 8 shows a route network, in which the 10 red points are popular airports and the red 10 lines are popular routes. Compared with the top-10 set of empirical data, the GRCA model can predict accurately 9 popular airports and 6 popular routes, the PRCA model can predict accurately 9 airports and 8 routes, the radiation model have 7 airports and 5 routes, the gravity model have 7 airports and 5 routes. The results of the two sub models are slightly better than that of the radiation model and the gravity model. The obvious difference is that the popular routes of the radiation model and the gravity model are unidirectional (i.e., $R_{Beijing-Chengdu}$ is a popular route, but $R_{Chengdu-Beijing}$ is not a popular route). The reason is discussed in the next section.

3.2.3. Travel distance

Travel distance is an important feature of passenger travel. It evaluates a model from a holistic perspective. Fig. 9 shows the travel distance distributions of different models. As we can see, these models have the similar result compared with the empirical data (light green circle). Travel distance's peak appears between 900 km and 1000 km, indicating that this range is the typical travel distance of domestic air passengers. We also plot route length distribution of route network in Fig. 9 (light blue fork). Compared with the models and empirical data, it also has the same trend. This implies that passengers travel distance is affected by the route network. Air passenger travel is a long distance movement and need to travel by plane. So route network limits the travel distance of air passengers.

4. Discussion

We develop a route city attraction model to characterize air passengers' travel activities. In the model, we consider the competitiveness among routes, so if we have corresponding city's GDP or population information, we can estimate the trip distribution of air passengers. Some researchers proposed that human travel has two generic mechanisms, exploration and

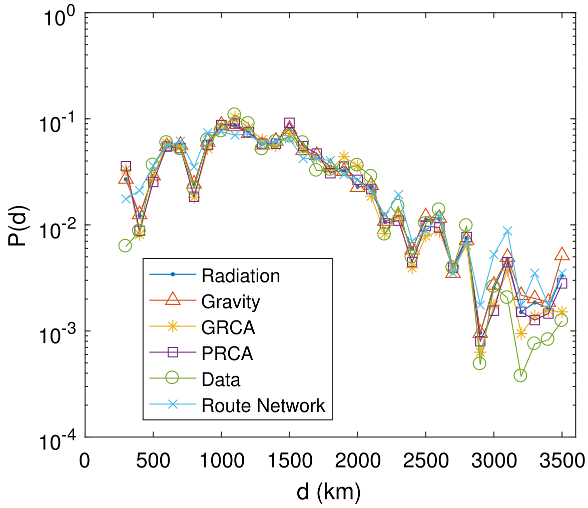

Fig. 9. Travel distance.

preferential return [9,23]. On the ground travel activities, the scope of human activities is relatively narrow and the travel cost is also cheap [21]. Therefore, when we reach a place, we may go to other places, rather than immediately return to the starting place. The cost of air travel is much expensive than that of ground travel. Thus, most passengers will go back to the starting place after arriving at the destination. Few passengers will go to another place, and return to the starting place from other places. So in our model, R_{A-B} and R_{B-A} have the same competitiveness, which characterize balance phenomenon of passengers' travel.

In the route network, there are two different characteristics compared with daily commuting network. Firstly, the connections among airports are not limited by geographical factors. For the application scenarios of radiation model, the area surrounds densely populated area also has relatively dense population. However, airports with little throughput are always connected to airports with large throughput. This makes the performance of the radiation model poor, despite it has the advantage of being parameter-free and good performance in ground travel activities. The forecast result of the airport (route) with large throughput is relatively small and the forecast result of the airport (route) with small throughput is large (Fig. 7(b) and (f)). Secondly, distance is not the key factor influences passengers' travel. Some passengers would like to choose famous scenic spots to travel, even if the distance is very far. Specially, the farther away, people are more likely to travel by plane. Besides, in gravity model, the route throughput is related to the departure airport's throughput. For route R_{A-B}, if Airport A has low throughput, the passengers arrive at Airport B are small. This does not conform the balance phenomenon. As shown in Fig. 7(a) and (e), the throughput prediction of gravity model is not very good when the airport (route) throughput is small and large.

It is noteworthy that our model describes air passengers' travel at air network scale. The prediction performance of our model is shown in Section 3, but the prediction quality is still unknown. Here, we use the Sørensen similarity index (SSI) [24] to quantify the degree of similarity with empirical datasets. As shown in Fig. 10, the SSI values of the four models (0.635, 0.633, 0.657, and 0.651) are very close in the forecast results of route throughput. For the prediction of airport throughput, our sub models outperform the radiation model and the gravity model. The SSI value is more than 0.75, which shows that our sub models' predictions have more than 75% similarity with the empirical datasets. It is noteworthy that the GRCA model's SSI is bigger than PRCA model's. This may imply that the economic factors have an influence on human long-distance mobility because many air passengers travel for work [20]. So in GRCA model, we use GDP to calculate city's attraction and the model can better describe air passengers' travel activities. Besides, we use Normalized Discounted Cumulative Gain (NDCG) [25] to measure the ranking of neighbor airports' throughputs. NDCG can test the prediction effect of the model on each airport. Fig. 11 illustrates the results. The red points are the mean values of all airports. The GRCA model and the PRCA model are better than the radiation model and the gravity model in predicting the neighbor airport ranking. This can explain why our sub models are more accurate for predicting each route. Actually, our model belongs to the aggregate model, analogous to the radiation model. The collective characteristics of air passengers are used to build the model. However, every passenger has her/his own travel characteristic. Therefore, the prediction performance may not be as good as the microscopic mobility models, which will be explored in our next studies.

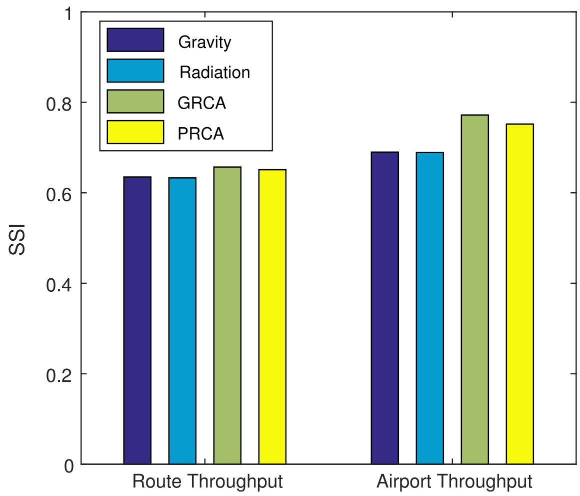

Fig. 10. GDP and route throughput.

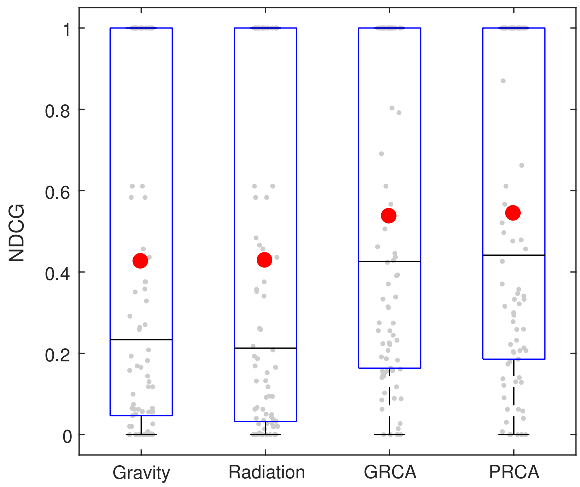

Fig. 11. Population and route throughput.

5. Materials and methods

5.1. Datasets

In this paper, we analyze the air passengers' reservation record data from an airline. The data covers the full year of 2014. It contains the passenger's flight number, seat number, flight departure, and arrival time. After eliminating the incorrect data and temporary route, we have 2 293 589 travel records of 70 366 passengers, a total of 86 airports, 1022 routes. In order to calculate the route length, we use the urban latitude and longitude instead of the airport latitude and longitude, the city's latitude and longitude information can be obtained from the website (http://api.map.baidu.com/lbsapi/getpoint/index.html). When calculating urban attraction, urban population and GDP data are obtained from the website of the National Bureau of Statistics.

5.2. Baseline model

In the experiment, we have two baseline models, the radiation model and the gravity model. They are two popular models describing the population movement between regions. The formula of the throughput between two cities in the radiation model is:

$$T_{i-j} = T_i P_{ij} = T_i \frac{m_i n_j}{(m_i + s_{ij})(m_i + n_j + s_{ij})} \tag{7}$$

Here we calculate the parameters according to the previous method [26]. T_i is the throughput of the airport i; m_i is the population of the city where the airport i is located; n_j is the population of the city where the airport j is located; s_{ij} is the sum of the population of the airports which have routes with airport i.

The formula of the throughput between two cities in the gravity model is:

$$T_{ij} = T_i P_{ij} = \frac{m_i n_j}{d(i,j)^b} \tag{8}$$

Using the same method of [27], distance function is a power function, we have $b = -0.13$. m_i is the population of the city where the airport i is located, and n_j is the population of the city where the airport j is located. $d(i,j)$ is the distance between the city where the airport i is located and the city where the airport j is located.

5.3. Evaluation metrics

In Section 4, in order to quantify the prediction performance, we respectively use SSI and NDCG to measure the similarity with empirical datasets and the neighbor airport throughput ranking performance of the model.

Sørensen similarity index. It is a statistical index to measure whether two samples are similar or not [24]. In the field of economics, this index is widely used. This indicator is also used to evaluate the performance of the human mobility model, defined as [13,26]:

$$SSI = \frac{1}{N^2} \sum_i^N \sum_j^N \frac{2min(T'_{i-j}, T_{i-j})}{T'_{i-j} + T_{i-j}} \tag{9}$$

where T'_{i-j} is the number of throughput from i to j predicted by model and T_{i-j} is the real throughput. If each T'_{i-j} is equal to T_{i-j}, SSI is 1, otherwise SSI is close to 0.

Normalized Discounted Cumulative Gain. NDCG is used to measure the sorting quality in the recommendation system [28–30]. It is defined as [25]:

$$NDCG = \frac{1}{IDCG} \times \sum_{i=1}^{N} \frac{2^{\tau_i} - 1}{log(i+1)} \tag{10}$$

where IDCG is the maximum possible DCG for a forecasting series of neighbor airports; if the predicted ranking is correct, τ_i is 1, otherwise τ_i is 0.

6. Conclusion

Air passengers' travels belong to the long-range and large-scale activities among cities and have its own characteristics, which are different from the characteristics of human travel activities on the ground. In this study, we construct a model based on the movement characteristics of air passengers, the route city attraction model. The key idea that makes our model distinct from previous ones is the balance and completion mechanism. The balance mechanism reflects the balance phenomenon of round-trip routes' throughputs. The completion mechanism reflects the competitiveness of route, which is related to the two cities. The RCA model has two sub models: the GRCA model and the PRCA model. Our sub models can reach the SSI value more than 0.75. That is to say, the models can share more than 75% common part with the real data. This suggests that the RCA model captures the fundamental mechanisms of air passengers' movement.

Acknowledgments

This work was supported by the National Natural Science Foundation of China [grant numbers U1333113, 61332001 and 61472050]. This work is also supported by the Science and Technology Planning Project of Sichuan Province [grant numbers 2014JY0257, 2015GZ0103 and 2014-HM01-00326-SF].

References

[1] F. Lu, K. Liu, J. Chen, Research on human mobility in big data era, J. Geo-Inf. Sci. 16 (5) (2014) 665.
[2] T. Zhou, X.P. Han, X.Y. Yan, Z.M. Yang, Z.D. Zhao, B.H. Wang, Statistical mechanics on temporal and spatial activities of human, J. Univ. Electron. Sci. Tech. China 6 (4) (2013) 481–540.
[3] C. Song, Z. Qu, N. Blumm, A.L. Barabási, Limits of predictability in human mobility, Science 327 (5968) (2010) 1018–1021.
[4] F. Xiong, Y. Liu, Z.J. Zhang, J. Zhu, Y. Zhang, An information diffusion model based on retweeting mechanism for online social media, Phys. Lett. A 376 (30–31) (2012) 2103–2108.
[5] M. Lenormand, A. Bassolas, J.J. Ramasco, Systematic comparison of trip distribution laws and models, J. Transp. Geogr. 51 (2016) 158–169.
[6] F. Xiong, Y. Liu, H.F. Zhang, Multi-source information diffusion in online social networks, J. Stat. Mech. Theory Exp. 2015 (2015) P07008.
[7] D. Brockmann, L. Hufnagel, T. Geisel, The scaling laws of human travel, Nature 439 (7075) (2006) 462–465.
[8] K. Lee, S. Hong, S.J. Kim, I. Rhee, S. Chong, Slaw: A new mobility model for human walks, in: IEEE INFOCOM 2009, 2009, pp. 855–863. http://dx.doi.org/10.1109/INFCOM.2009.5061995.
[9] C. Song, T. Koren, P. Wang, A.L. Barabási, Modelling the scaling properties of human mobility, Nat. Phys. 6 (10) (2010) 818–823.

[10] X.P. Han, B.H. Wang, Impacts of distance and memory in the emergence of scaling mobility pattern of human, Phys. Procedia 3 (5) (2010) 1907–1911.
[11] G.K. Zipf, The p1p2/d hypothesis: On the intercity movement of persons, Amer. Sociol. Rev. 11 (6) (1946) 677–686.
[12] F. Simini, M.C. González, A. Maritan, A.L. Barabási, A universal model for mobility and migration patterns, Nature 484 (7392) (2012) 96–100.
[13] X.Y. Yan, C. Zhao, Y. Fan, Z. Di, W.X. Wang, Universal predictability of mobility patterns in cities, J. R. Soc. Interface 11 (100) (2013) 20140834.
[14] A. Noulas, S. Scellato, R. Lambiotte, M. Pontil, C. Mascolo, A tale of many cities: Universal patterns in human urban mobility, Plos One 7 (5) (2012) e37027.
[15] B.C. Csáji, A. Browet, V.A. Traag, J.C. Delvenne, E. Huens, P.V. Dooren, Z. Smoreda, V.D. Blondel, Exploring the mobility of mobile phone users, Physica A 392 (6) (2012) 1459–1473.
[16] W.H.K. Tsui, H.O. Balli, A. Gilbey, H. Gow, Forecasting of Hong Kong airport's passenger throughput, Tourism Manage. 42 (42) (2014) 62–76.
[17] Z. Z.D., X. J.H., An analysis of major factors on airport passenger volumes, Urban Transp. China 5 (6) (2007) 54–57.
[18] T. Jia, K. Qin, J. Shan, An exploratory analysis on the evolution of the US airport network, Physica A 413 (413) (2014) 266–279.
[19] B. Jiang, T. Jia, Exploring human mobility patterns based on location information of US flights, 2011. http://arxiv.org/abs/1104.4578v2.
[20] F.H. Huang, J. Peng, M.Y. You, Analyses of characetristics of air passenger group mobility behaviors, Acta Phys. Sin. 65 (22) (2016) 228901.
[21] X.Y. Yan, X.P. Han, B.H. Wang, T. Zhou, Diversity of individual mobility patterns and emergence of aggregated scaling laws, Sci. Rep. 3 (9) (2012) 454.
[22] A. Barabási, The origin of bursts and heavy tails in human dynamics, Nature 435 (7039) (2005) 207–211.
[23] L. Pappalardo, F. Simini, S. Rinzivillo, D. Pedreschi, F. Giannotti, A.L. Barabási, Returners and explorers dichotomy in human mobility, Nature Commun. 6 (8166) (2015).
[24] T. Sørensen, A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on danish commons, Biol. Skr. 5 (1948) 1–34.
[25] K. Rvelin, Kek, J. Inen, Cumulated gain-based evaluation of IR techniques, Acm Trans. Inf. Syst. 20 (4) (2002) 422–446.
[26] M. Lenormand, S. Huet, F. Gargiulo, G. Deffuant, A universal model of commuting networks, Plos One 7 (7) (2012) 457–464.
[27] Y. Yang, C. Herrera, N. Eagle, M.C. González, Limits of predictability in commuting flows in the absence of data for calibration, Sci. Rep. 4 (2014) 5662.
[28] Y. Zheng, L. Zhang, Z. Ma, X. Xie, W.Y. Ma, Recommending friends and locations based on individual location history, Acm Trans. Web 5 (1) (2011) 5.
[29] Y. Chang, A. Dong, P. Kolari, R. Zhang, Y. Inagaki, F. Diaz, H. Zha, Y. Liu, Improving recency ranking using twitter data, Acm Trans. Intell. Syst. Technol. 4 (1) (2013) 4.
[30] Y. Liu, C. Liu, B. Liu, M. Qu, H. Xiong, Unified point-of-interest recommendation with temporal interval assessment, in: The ACM SIGKDD International Conference, 2016, pp. 1015–1024.