Airport check-in counters assignment using predicted passenger load: A Case Study

MA Nang Laik
Associate Professor, School of Business
Singapore University of Social Sciences
461 Clementi Road, Singapore 599491
nlma@suss.edu.sg

Murphy Choy
Heriot-Watt University
Edinburgh Business School
Edinburgh, Scotland, United Kingdom
goladin@gmail.com

Abstract

Before Covid19, airport around the world are faced with limited capacity issues and predicting the passenger demand at the airport is vital for resource planning at various customer touch points ranging from check-in counters and immigration. In this case study, we use one-year historical data to forecast the departure demand for each airline and develop a simulation model to determine the number of check-in counters required to optimize the resource utilization at the airport. The overall objective of this paper is to uncover the real-world business problem in the classroom setting and the authors will share how to go through this case study in detail.

Keywords: forecasting demand, check-in counter, decision tree, regression, departure passengers

1. Introduction

The growth in airline traffic has been striking, despite economic recession cycles and increases in the oil price. The global air travel market has been growing at a steady rate for the past decade. Since 2001, the total number of air travel passenger has increased by 25%. This increase translates to approximately 2.25% growth annually in passenger load. With increasing number of passengers traveling by air, the revenues of the airlines will also increase. There has been an increase of 83.33% in revenue from 2001 to 2010. On the other hand, expenses have also increased at almost the same rate of 80% which effectively wipes out the overall profit margin for the airlines. Due to this poor profit margin and high competition between various airport terminals, it is extremely difficult for airport terminals to increase the existing fees that they are charging airlines for the use of terminal and services. Many airports in the world have benefited from the increase in their passenger volume by increasing their profitability through the use of shopping malls, duty free shopping and improve their operational efficiency with limited resources.

Modern airports need substantial infrastructure investment for their long runways, taxiways, airport operational equipment, passenger terminal areas and expensive ground handling equipment. An airport serves either as a transit point or terminal point for the passengers during a trip. Airport operations can be roughly divided into airside and landside. Most of the Asian airports are very complex and requires proper coordination and efforts to facilitate the daily operations. Some of the core processes in the airport are: handling of passengers and baggage, servicing, maintenance and engineering of the aircraft, ground handling activities, leasing of rental spaces for retail shops,
aviation support facility (air traffic control) and finally custom and immigration for the passengers. Airport operators do not operate alone. They normally form partnership with various partners such as ground handlers to handle the passengers and baggage, catering companies to be in-charge the meals on board, and engineering companies to take care of aircraft maintenance.

The airport terminal operator is a big player in the industry and operates one of the world’s most highly rated airports in Asia. It has turned to data analytics journey in an attempt to predict the passenger load for departure flights to estimate the flow of passengers during daily-operations. This desire to understand their business stems from the increase in passenger traffic, limited terminal resources at check-in, drop in passenger satisfaction levels and long waiting time at the check-in counters. They would like to harness the power of analytics to gain useful insights based on the past trend and travelling patterns of the passengers and predict the passenger loads based on historical information which could potentially assist them in improving their operational efficiency and making their business more sustainable. On a day to day basis, either the airline or the ground handlers will request the airport operators to open more counters than required to service the departing passengers and to improve customer experience at the counters. Thus, knowing the predicted passenger load from the past data becomes the main motivation for our study.

This case has been derived from many year of consultancy work done by the authors with the airport operators. We have framed the real-business problem to a 3-hours case study to be used in one of the data analytics class with the objective of bringing real-world scenario into the classroom. This case study’s main objectives are to help students in the following:

- Discuss the business problem and identify inputs and outputs
- Understand data and prepare the data for analysis
- To inculcate an appreciation of forecasting techniques and how the forecasts have an impact on the subsequent business decision
- To appreciate the relationship between demand forecasting and resource planning problem
- Develop a simulation model to mimics the real-world scenario for scenario planning
- Develop the daily resource requirement based on forecasting demand and simulation
- Evaluate and validate the results and share valuable insights with the users

2. Literature Review

In our literature review, we begin by looking at past works which solved the problem of assigning check-in counters. Chun and Mak (1999) developed a comprehensive intelligent resource simulation system to predict on a daily basis how many check-in counters should be allocated to each departure flight while providing passengers with sufficient quality of service. Yan et al. (2004) developed an integer programming model to assist airport authorities to assign common check-in counters on a monthly basis with the objective of minimizing passengers walking distances. The complexity of the problem required the development of a heuristic method.

Joustra and van Dijk (2001) developed a simulation toolbox to study the behavior of check-in counters and the toolbox is validated by experts of Amsterdam Airport Schiphol on the basis of a study. van Dijk and van der Sluis (2006) first used the simulation to determine minimal numbers of desks in order to meet a service level for each separate flight. Next, integer programming formulations are provided to minimize the total number of desks and the total number of desk hours under the realistic constraint that desks for the same flight should be adjacent. They also introduced some operating realistic constraints to solve the problem for real cases. However, the formulation of the resulting allocation problem turns out to be a new NP-hard problem, whose solutions have to be computed by using heuristic approaches. Finally, preliminary results for real world shows a triple win in waiting time performance, in number of desks and in number of desk hours (staffing).

Related to check-in counter assignment, there is another group of papers which looked at performance of the airlines. The simulation model developed by Haeme et al. (1988) helped the airlines evaluate their on-time arrival performance. Correia and Wirasinghe (2007) developed a methodology to study the service standards at the
passenger airport based on the user perceptions using the surveys. The check-in counter is identified as one of the key components of the measurement which include processing time, waiting time and space available between each passenger. The study uses data obtained from a passenger survey conducted at São Paulo/Guarulhos International Airport, Brazil.

There are also papers which look at baggage flow problems, where one of them is Atkins et al. (2003), who combined queue theory and optimization to solve the baggage flow problem at Vancouver International Airport. The result showed that through efficient scheduling and job deployment, 90 percent of Vancouver International Airport passengers could expect to wait no longer than 10 minutes at pre-board screening security points.

Airport terminal is a complex system which involves multiple stakeholders, agents, different passengers flow, government agencies and various operational policies. In order to help the airport manager to make strategic – high level decision, Manataki and Zografos (2010) developed a generic and flexible decision-support tool to facilitate the high-level decision-making related to fundamental changes in the structure and operation of the airport terminal system. This tool has been used to access the performance of the Athens International Airport passenger terminal under different demand and resource deployment scenario.

Simulation is commonly used for planning and design of airport. Jim and Chang (1998) presented a simulation tool using SLAM II simulation language for the final design of airport passenger terminal. The tool allows the management of airport passenger terminal to plan the different design and improvements for the existing or proposed terminal before the actual construction of an airport terminal.

Ma et al. (2014) analyzed the passenger load from the past historical pattern and developed a predictive model using decision tree (DT) to forecast the passenger load based on certain criteria. The model is being tested against the actual data given for a particular month and the root mean square error of 3%-12% is observed for all the airlines at the airport.

Ma (2017) focus on the forecasting of monthly departure passenger movements for one of the busiest airport in Asia. The author forecasted the monthly airport departure passenger flows for the next 12 months for macro level planning. Next, she used SAS Forecast Studio for detailed-level planning based on airline and per airline-city combinations using hierarchical forecasting. The result shown that in most cases, the mean absolute percentage error is less than 3%, which indicates the usefulness of our model for better decision making.

From the literature review, we understand that check-in counters at the airport have been studied before. But the contribution of our paper is in the use of a simple spreadsheet model to simulate the passengers’ arrival and to collect various performance statistics such as average queuing time, average system time and average queuing length at the check-in counters. Apart from the ease of use, spreadsheet model is an effective tool to display results visually. It is also very intuitive as users can change the input parameters easily to see the impact of changing number of counters on the service level.

3. Case Study Analysis

We teach this case in “Business Analytics Application”. At the beginning of the course, most students have very little knowledge about data analytics and we have to teach them various data analytics skills over a few lessons. Over time, they become quite familiar with solving this kind of problem. This case study is given to them towards the end of the course where they are expected to combine various topic to solve a real-world business problem of this kind.

This case study has been targeted for undergraduate and post-graduate students pursuing a degree in any data analytics or decision science subjects. Requires prior experience with Excel, R or any data mining software such as SAS or SPSS. Students should have some basic understanding of the logistics and transportation industry.

The case study requires the application of forecasting models, simulation and resource allocation methods to solve the daily operational problem faced by one of the largest airports in the world. It is recommended that the instructor
using the case study to break the case across multiple lessons and solving a specific aspect of the problem in each class.

Over the years, our course, with this approach using a set of similar simple to hard cases for students to learn, has gained an excellent reputation as one of the “best and most useful courses offered” in our university. We have used the case to train hundreds of graduate students who are enrolled into data analytics course in our university.

3.1 Business understanding

Airlines are the major customers to the airport operators and the main objectives are to ensure the on-time departure of the aircraft and improve the passengers experience and convenience at the airport. The arrival and departure processes of aircraft at the airport are two major operations which trigger various subsequent activities at the airport. Figure 1 shows the overview of the departure process of a passenger at the airport.

![Figure 1. Departure process](image)

### Passenger Load

Currently, most of the airlines who use the terminal, do not submit the passenger load to the terminal as the airlines are still selling ticket until the last minute and the competition is stiff. Passenger load information is only available a few hours before the flight’s departure and arrival. However, from the airport planning perspective, the number of passengers arriving and departing through the airport is one of the crucial factors to do the necessary planning for resources and man-power allocation. There is no proper tool for them to use to aid in their daily decision making. The terminal planning managers are using a fixed rate of Y% passenger load to do the daily planning and assign the number of check-in counters required based on “gut-feel” and experience. Over estimating of the passenger load will require more resources such as check-in counters which are not fully utilize during the operations and resulted than higher operational cost and wastage. Conversely, under-estimating of the load will result in shortage of counters opened and passengers need to wait for a long time at the check-in counters and failure to meet the service level agreement.

Since the airport has all the historical data about the actual passenger load, we would like to build a predictive model to determine the passenger load for each flight based on a number of attributes given in the dataset. For airline passenger load analysis, we are interested in the following performance measures or indicators.

- Passenger load (pax. load) = total passenger on board / max seating capacity

Max seating capacity varies according to the airline configuration and aircraft type which we can derive using the past data to get the maximum seating capacity based on airline and aircraft type.

The objective of this case study is to develop various predictive models to forecast the departure passenger loads based on the historical demand and evaluate the accuracy rate of the models. The management is interested to identify the importance variables which affects the passenger load.
Resource planning

A very common objective from a corporate point of view is the minimization of the operational cost by utilizing the resources efficiently and effectively to guarantee the required SLA. For the departing flight, the passengers will arrive at the airport approximately 2.5 hours before the departure time where most of the counters will be opened for check-in. The check-in counter will be opened for 2 hours and it will be closed 30 minutes before the scheduled departure time to ensure that the passengers have enough time to board the plane and load the baggage on board.

In the case of check-in counter management, this objective can be translated into the determination of the minimum number of check-in desks to be opened in a given time interval to ensure the service coverage. In this case, a dynamic rather than a static management policy can help to minimize costs, as opening a greater number of desks in peak hours and closing them when service demand is decreasing allows an optimization of check-in resources.

It is a difficult and challenging operational problem faced by airport operator in the attempt to decide the number of check-in counters while balancing the operational costs and the service level passengers have to be provided with, in terms of queue length and waiting times.

3.2 Case Discussion

The instructor should begin the case by introducing the following concepts to the students:

- Airport terminal operations
- Current workflow for departure passenger at the airport
- Forecasting methods
- Simulation model of queues
- Resource planning in general
- Capacity planning

The analysis of the problem in this case study starts with understanding of the problem and making appropriate assumptions. After initial “brainstorming”, students should be able to list the key inputs of the model are:

- Historical data for departure passengers
- Customer arrival pattern at the check-in counters
- Service time at the check-in counters for each passenger
- Number of counters available at the terminal

You need to understand the domain and study the current business processes and identify some potential pin points where data analytics can be applied to overcome the business and operational issues in the current business environment. In addition, you must perform an initial data analysis and present your finding and propose analytics solutions as proof-of-concept to the executive management team, which will include people both from the business and IT divisions.

In order to achieve this, you should do the following tasks (but not restricted to):

a. Perform exploratory data analysis on the data that is given
b. Develop predictive models to predict the departure passengers load based on airline, destination or origin?
c. Compare and validate the models developed above and suggest the most suitable for implementation
d. Analyse the pros and cons of the chosen model and model limitations
e. Determine the check-in counters requirement for daily operations

Data given

1 year of historical data (2010 data) and the data fields are described in Table 1. Data has been masked to protect the clients’ identify, but it will not affect the quality of your analysis:
Table 1. Flight Data

<table>
<thead>
<tr>
<th>Field name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>flight_dt</td>
<td>Flight date</td>
</tr>
<tr>
<td>Airline</td>
<td>airline</td>
</tr>
<tr>
<td>Flight_No</td>
<td>Flight number</td>
</tr>
<tr>
<td>Arr_or_Dep</td>
<td>A - Arrival, D - Departure</td>
</tr>
<tr>
<td>Terminal</td>
<td>Terminal name</td>
</tr>
<tr>
<td>Ori_Des</td>
<td>Original or destination (city code)</td>
</tr>
<tr>
<td>Total_Pax</td>
<td>Total number of passenger on the flight</td>
</tr>
<tr>
<td>Max_Seat</td>
<td>Maximum flight capacity</td>
</tr>
<tr>
<td>Schedule_time</td>
<td>Schedule time of departure or arrival</td>
</tr>
<tr>
<td>Actual_time</td>
<td>Arrival time of arrival or departure</td>
</tr>
</tbody>
</table>

Students are also given 2011 – one-week flight schedule so that they can use the predictive model developed above to test the validity of the model. The overview methodology to solve the problem is illustrated in Figure 2.

![Figure 2. Overall methodology to solve the case](image)

Specific topics discussion questions are also listed below.

**Forecasting**
- What are the popular techniques for forecasting?
- Do you foresee issues with small amount of data for forecasting?
- Given the need to present the forecasts to senior management who are more comfortable with Excel spreadsheets, do you think using a simpler model that is available in Excel spreadsheet will be more readily accepted?
- Do you think advanced forecasting techniques such as neural network will work well in this case?
- What are the suitable measures to determine the best forecasting model?

**Simulation model**
- How do you collect the parameters to build simulation model?
- What are the kinds of queues that are available (Bank queue, Priority queue, Multiple servers’ multiple queues)?
- What would be the key performance indicators to collect?

**Resource planning**
- Determine the decision variables for resource planning problem
- What is the optimal number of check-in counters required to meet the service level agreement over 24-hour time period?
4. Case development & models building

4.1 Exploratory data analysis (EDA)

Based on our initial analysis, pax. load varied according to day of the weeks and weekends. Students can use various graph that they have learnt to display the result. There is also a seasonality effect on the demand such that there are higher pax. load during the festive season in December and during the school holiday in March and June.

There is also a highly volatile in pax. load due to other factors such as airline and destination of travel. For budget airlines, the pax. load is relatively higher than the full-cost carriers in general as there are more people taking the budget airline due to the attractive low cost to fly. The histogram of pax. load and the general statistics is given in Figure 3.

<table>
<thead>
<tr>
<th>NAME</th>
<th>MEAN</th>
<th>STD</th>
<th>MIN</th>
<th>MAX</th>
<th>SKEWNESS</th>
<th>KURTOSIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pax_load</td>
<td>0.63346</td>
<td>0.2113</td>
<td>0.01539</td>
<td>1</td>
<td>-0.32893</td>
<td>-0.63111</td>
</tr>
</tbody>
</table>

Figure 3. Histogram of pax_load and mean statistics of pax_load

Students are also expected to identify what are the top ten destination based on pax. load to have a better understanding of the situation.

4.2 Predictive models

Three data mining techniques are commonly used for predictive models, namely, decision tree, regression, and neural networks. These are models from machine learning, statistics, and artificial intelligence. Usually, all three are used and the results are evaluated to identify the best model, based on prediction error typically measured by mean squared error, which measures the difference between the predicted value (pax. load) and actual value (pax. load).

\[
\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - \bar{y}_i)^2 \text{ where } \hat{y}_i \text{ is the predicted value and } \bar{y}_i \text{ is the actual value}
\]

There is no one best data mining tool for predictive modelling as each of these models has its own pros and cons. For example, decision trees are easy to interpret and can also handle missing values well, regression is easy to apply and use but it is cumbersome to include non-linear and interaction effects. Neural networks are a very good universal approximate but their results are not easy to interpret.
Students are expected to use three data mining techniques and identify the important factors which can predict the pax. load. Based on the data given, both airline, destination, day of week and months are important parameters and they are significant.

Using the model comparison, students could choose the best model. Table 2 shows one of the expected outcome of the models developed. From the outcome of the model, we have observed that the decision tree has the lowest MSE as compared to regression and AutoNeural and useful to predict the passenger load for the real-world with MSE of 3%.

<table>
<thead>
<tr>
<th>Selected</th>
<th>Description</th>
<th>Valid: MSE</th>
<th>Train: MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>YES</td>
<td>Decision Tree</td>
<td>0.031624</td>
<td>0.030798</td>
</tr>
<tr>
<td></td>
<td>Regression</td>
<td>0.031917</td>
<td>0.030962</td>
</tr>
<tr>
<td></td>
<td>AutoNeural</td>
<td>0.046676</td>
<td>0.045576</td>
</tr>
</tbody>
</table>

The model has some limitations; it assumed that past demand is good estimation of the current pax. load. We haven’t taken into consideration of external factor such as airline promotions, long weekends or holidays or new destination or growth factor. We also do acknowledge that if we use a few years of data, then we can build a more robust model but due to the lack of data available, we will proceed with using the result of this model for the resource allocation problem.

### 4.3 Simulation Model

We have done field study at the airport for a few airlines to identify the input parameter required (passengers’ arrival pattern, service time etc.) to simulate the check-in process where a passenger will drop the baggage and waiting for boarding pass to be issued. We have collected the data for 7 days of the week and we are able to fit it to the statistical distribution for the purpose of simulation.

The inter-arrival time of the passengers based on their travelling profiles are generated based on the data collected. From the data collected, we have identified that 20% of the passengers who tend to check in between 2.5 hours and 2 hours before STD, the majority of the passengers 60% of the total predicted passenger load will check in during one to two hours before STD and the remaining 20% will check in during the last 30 minutes before counters are closed.

Given the number of passengers for a particular flight and the assumption that all the passengers will arrive within the 2-hour of check-in counter opening time, we can estimate the average inter-arrival time of the passengers using the formula.

\[
\text{Inter-arrival time} = \frac{\text{Counter Opening Period}}{\text{total \# of passengers}}
\]

In addition, we also assume that the check-in process service time per passenger is 1 minute and 30 seconds and it also follows the exponential distribution.

The descriptions for the arrival time, service start time, service end time, waiting time, system time and system length are given below.

\[
\text{Arrival Time} = \text{Previous Arrival Time} + \text{Inter-Arrival Time}
\]
Service Start Time = If the queue length is less than number of counter, then the person will be served upon arrival (Service Start Time = Arrival Time), otherwise, the Service Start Time is the earliest Service End Time of all the counters

Service End Time = Service Start Time + Service Time
Wait Time = Waiting time from the passenger’s Arrival time till he gets served = Service Start Time – Arrival Time
System Time = Total time spends in the system from the time the passenger arrives at the queue to the time the passenger leaves the counter = Service End Time – Arrival Time = Wait Time + Service Time
System Length = The number of people in front of the passenger upon his arrival at the queue

Students are expected to develop an Monte Carlo simulation model using the predicted passenger load. The simulation models should output several performance indicators including the 90% percentile of the system time, average waiting time, average system time (average time spent in the system from joining the queue to leaving the check-in counter) and average system length (average queue length upon arrival) for a given number of check-in counter.

The number of passengers per flight will vary according to the seasonal demand such as school holidays, Christmas season, route of flight and other economic factors. The airport operator is interested to find out the number of check-in counters required given the passenger load. We have run the simulation using the above model by varying the number of passengers for the flight and obtained the optimal number of counters required to meet the service level agreement. A graphical representation which shows the relationship between the number of counters and the pax load is also shown in Figure 4.

The equation for a linear trendline is given as $y = 0.0137x + 0.5074$ and the $R^2 = 0.9632$ which indicates a very good approximation. We can just use the equation to determine the number of counters needed given the number of passengers. Since the number of counters is an integer value, we need to get round up to the nearest integer value. Using the linear relationship above, we can compute the number of counters required given the departure passenger load for a particular flight. We use the scheduled departure time (STD) as a reference and counters are only open two and half hours before STD and they will be opened for two-hours duration for the passengers to do their check-in. We have calibrated 24 hours to 48 time-slots each 30-minutes time window.
4.4 Check-in counters assignment problem

To do the final resource planning, students need to use the 2011 flight data and predict the passenger load using the developed predictive model. Each time block of 30 minutes is required as time period and there will be 48 time periods (t) in 24 hour. Using the outcome of simulation and predictive models, we can determine the number of check-in counters required for 24 hours.

The predictive model will give the predicted pax. load for each flight and the predicted passenger number can be derived using the formula:

\[
\text{Predicted passenger} = \text{predicted pax. load} \times \text{maximum capacity} \text{ and round up the number to the nearest integer value for planning.}
\]

For each flight in the day, using the predicted passenger, we need to derive the number of check-in counters required. Example if the passenger load is 0.6914 and maximum capacity is 212, the predicted passenger is 147 and 3 check-in counters will be required for the flight. Students also need to take note the check-in counters are opened 2.5 hours before STD and counters will be closed 30 minutes before STD. The counters are required throughout the check-in period. We also assume that there is no sharing of check-in counters for different flights. However, it may not be true in reality, there are some airlines which have dedicated check-in roles and they can allow passengers to check in 24 hours before. They will require more check-in counters and has made special arrangement with the airport operator. We can also use the same terminology to determine how many check-in counters required based on the airline code.

The daily requirement of the resources can be computed using the above mentioned method and the graph such as Figure 5 should be displayed for the operations managers and highlight the peak and alert the manager if 80% of the resources are occupied for necessary actions and follow-up.

![Figure 5. Counters requirement for a typical day at the airport](image)

5. Conclusion

In this case study, students are given one year of historical data to build the predictive models to forecast the demand. Students need to explore various predictive data mining techniques and choose the best model for implementation. As a data analytics profession, students must be able to share these actionable insights with the terminal managers which enables them to do better planning of terminal resources such as check-in counters. In our research, we also developed the simulation models for the check-in counters and finally allocating the resources using optimization.

© IEOM Society International
Using the decision support system, the terminal managers can change various inputs to run different scenarios in simulation to plan for unexpected events and anticipate the changes required. Harnessing the power of analytics, the airport operators can make a better decision and assist the departure passengers so that the waiting time at the check-in counters can be minimized, to give a better impression to passengers of departure passengers one of world best airports.

Acknowledgements

This case study has been developed based on the authors the consultancy work done by the faculty members for one of the busiest airport in the region. The client would like to remain anonymous and we appreciate the company for the support and knowledge sharing.

References


Biographies

Nang Laik is an Associate Professor in School of Business, Singapore University of Social Sciences (SUSS). She has more than ten years of teaching undergraduate courses and master courses throughout her academic career. She teaches undergraduate core modules and master level courses in the areas related to business modeling using spreadsheet, data analytics and logistics and supply change operations. She holds a PhD from Imperial College, London where her research focused on operations research (OR) in the area of optimization of resource. Her research expertise lies in the simulation and modelling of large scale real-world problems and the development of computationally efficient algorithms to enable sound and intelligent decision making in the organization. Nang Laik is an expert in transportation and logistics industry, she serves as a consultant for one of the best airports - Changi Airport Group to use data and decision Analytics to generate insights, make better decision and improve the business efficiency and productivity. She has also previously worked in one of the largest container ports to develop and implement multi-million dollars’ decision support system.
Murphy is an experienced analytics specialist with extensive experience in risk analytics, marketing analytics, social media and Big Data analytics. He has extensive experience in developing new techniques and models to achieve business objectives in real world application. He is also passionate about lifelong learning and is an active contributor to several international publications.