MS-E2177 Seminar on Case Studies in Operations Research 2017

Predicting and managing group cancellations in air passenger traffic

Client: Finnair

Project plan

February 21, 2018

Tuomas Koskinen (project manager)
Elias Axelsson
Olli Herrala

Background

Finnair is a Finnish airline company with a long history (founded in 1923). The geographical location of Finland gives Finnair a competitive advantage in flights between Europe and Eastern Asia. About 40% of Finnair's revenue comes from Asia flights[1] and that's where group cancellations, our project topic, are a big topic.

Traveling from Asia is slightly different from what we are used to in Finland. Groups of 50 people are not unusual, and group cancellation rates are high (even 50%). This is very challenging for the current revenue management systems, which more or less assume all customers to be independent.

Revenue management is based on three techniques: overbooking, fare class mix and traffic flow control.[2] In this project, we will be focusing on overbooking. The concept of overbooking comes from the significant no-show rates in air passenger traffic. Airline companies often allow cancellations without large sanctions in higher fare classes, which results in any such empty seats only lowering revenue. This lost revenue is referred to as "spoilage". This is countered by selling more tickets than the physical capacity of the plane, thus the term "overbooking". If the number of passengers eventually showing up for a flight is greater than capacity, the airline must reimburse the customers who can not fly on the plane despite having a ticket. This is referred to as "denied boardings", and is strictly regulated by the EU (regulation 261/2004).

Both denied boardings and spoilage cause expenses for the airline company, and models for minimizing the expected total costs have been produced. The models usually give an optimal number of tickets to be sold based on the mean and variance for the no-show rate. However, as stated earlier, these models assume the customers to be completely independent of each other. When a group of 50 people reserves a trip from Eastern Asia to Finland, it is different from single customers or even small families. The group either shows up or does not, and the no-

show rates can be very high. The problem this causes is easily demonstrated if we think of a flight with a capacity of 100 passengers, with 75 passengers who are very likely to show up, and a group of 50 with a statistical no-show rate of 50%. Describing the group with a model that assumes independent customers with a normal distributed actual no-show rate, we come to a conclusion that this is close to (depends on the ratio of denied boarding and spoilage costs) an optimal solution with the lowest expected costs for the company. However, in reality we have either 25 denied boardings or spoilage, and if there's a difference in these costs, they quickly add up when multiplied by 25 passengers.

Objectives

The objective of the project is to develop an optimal overbooking strategy which takes into account group reservations by utilizing a statistical forecasting model and stochastic simulation. Historical group reservation data is used for fitting a model for predicting the probability of cancellation. The data consists of group reservations from Finnair's flights during years 2016 and 2017. For each reservation there's various pieces of information available. These include point of departure and arrival, reservation and departure dates, travel agent information, cancellation status and others. The goal is to develop a model which predicts the cancellation probability of a given group reservation based on these features. A model with good predictive accuracy makes it possible to better anticipate cancellations and take this into account when determining an optimal level of overbooking. Intuitively, when a high cancellation risk reservation is identified using the developed model a higher level of overbooking can be used to maximize expected revenue in the long run.
Additionally, a stochastic simulation will be programmed to model the reservation and cancellation process. Different overbooking strategies can be evaluated using the simulation. Lastly, a simplified overbooking strategy will be proposed based on the results of the predictive modeling and simulation. The goal is to maximise revenue while minimizing denied boarding risk.

Tasks

1. **Planning the project**

Planning consists of defining the objectives of the project and specifying and scheduling the tasks necessary for achieving these objectives. In addition resources and risks are evaluated.

2. **Literature review**

In order to familiarize ourselves with the subject a literature review on airline revenue management is necessary. Literature concerning forecasting models is also studied.

3. **Exploring the data**

The data is complicated and messy. It will take some time to get an understanding on the relevance of all the provided features.

4. **Data preprocessing**

In order to to fit a model on the data it is first necessary to apply needed preprocessing procedures. These include feature selection, mapping and normalization.

5. **Model fitting**

After the data has been cleansed and preprocessed different statistical models will be fitted to find one which best predicts cancellation risk.

6. **Model testing and comparison**

The models will be tested and compared with each other to find the most suitable one.

7. **Writing the interim report**

8. **Programming the simulation**

A stochastic simulation will be programmed to model the reservation and cancellation process. Together with the forecasting model the simulation results are used to propose an overbooking strategy.

9. **Formulate an overbooking model**

A simplified overbooking strategy is proposed based the the forecasting and simulation results. The goal is to find an overbooking strategy which maximises flight revenue while minimizing the risk of denied boardings.

10. **Writing the final report**

Schedule

The tasks specified in the project plan are scheduled on a weekly resolution and can be seen in the following table.

Task \ Week	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Planning the project															
Literature review															
Exploring the data															
Data preprocessing															
Model fitting															
Interim report															
Model testing and comparison															
Programming the simulation															
Formulate an overbooking model															
Final report															

Resources

Our project team consists of three members: Elias Axelsson, Olli Herrala and Tuomas Koskinen. We are all master students at the Aalto University majoring Systems and Operations Research. We have also experience in programming, machine learning and stochastics.

The contact person inside Finnair is Antti Tolvanen, with whom we aim to discuss frequently. He has already briefed us for the assignment and he will also prepare the data for us. The course staff, including Professor Ahti Salo and assistant Ellie Dillon, give us feedback and also support when needed.

Our most critical resource is of course the reservation data received from Finnair (the content was still unclear when writing this). Technical resources include different computing softwares which we use to build our implementation. Additionally, we have academic access to scientific literature and Antti have also sent us few articles on the subject.

Risks

We identified few risks which could affect the final outcome (table below). The occurrence probability and impact are evaluated as low, moderate or high. We have also proposed ways to mitigate the risks. As we have already identified the risks, we can proactively try to avoid them.

Risk	Probability	Effects	Impact	Mitigation measures
Member inactivity or dropout	Moderate	Increase in the workload of other members	Low to high	Transparency in scheduling
Inability to stay on schedule	Moderate	Workload grows large towards the end and implementation may remain incomplete	Low to high	Frequent meetings between team members
Weak communication with customer	Low	Implementation proves to be unsatisfying	Moderate	Frequent meetings with customer
Data proves to be too messy	Low	Inability to build a model	High	Active communication with customer
Inability to build a reasonable model	Moderate	Little or no value creation for customer	High	Studying the subject carefully and scoping with customer and course staff
Model doesn't satisfy customer needs	Moderate	Low value creation for customer	Moderate	Model comparison and frequent meetings with customer

References

[1] Finnair annual report 2016 https://investors.finnair.com/~/media/Files/F/Finnair-IR/documents/fi/reports-and-presentation/2017/finnair-vuosikertomus-2016-fi-v2.pdf

[2] Belobaba, Peter, Amedeo Odoni, and Cynthia Barnhart, eds. The global airline industry. John Wiley & Sons, 2015.