Dynamic Re-accommodation: Automated Tools for Adapting to Airline Schedule Shifts

Team 99

MPhasis Foundation Mid-Prep Quantum Computing Challenge Report



Team 99
Table of Contents-

| Sr. No. | Header | Page No. |
|---------|---|----------|
| 1 | Abstract | 3 |
| 2 | Problem Statement and Project Objectives | 3 |
| 3 | Methodology and Procedure | 5 |
| 4 | Results | 9 |
| 5 | Passenger Emailing Facility | 11 |
| 6 | Business Profiling | 12 |
| 7 | Potential Applications of Quantum Computing in this problem with a study on its Effectiveness | 13 |
| 8 | Conclusions | 15 |

Abstract

This report presents an innovative automated solution to address the challenges of routine schedule changes in the airline industry. Airlines frequently adjust their flight schedules due to seasonal demands, new routes, daylight savings, flight numbers, and operating frequency modifications. These changes can impact passengers, requiring them to be re-accommodated on alternate flights. Our proposed system focuses on efficiently identifying affected flights and passengers, optimizing alternate flight solutions, ranking them based on various factors, and performing Business Profiling. We have optimized the data sets given to us by the Mphasis Foundation with the parameters and weights assigned to us in the rulebook and have reviewed the potential applications of Quantum Computing in this specific problem.

Problem Statement

Airlines undergo frequent adjustments to their flight schedules for various reasons, such as changes in seasonal demand, the need to identify new routes, adjustments for daylight savings time, modifications to flight numbers, operational frequency, and other considerations. These alterations in itinerary can have significant repercussions for numerous passengers, necessitating the creation of new re-accommodation schedules.

To address these challenges, airlines require a robust solution to assess the implications of the scheduled itinerary modifications for their passengers. This entails an automated system that intelligently determines the most suitable alternative flights for affected passengers, considering availability, convenience, and efficiency. The process involves finding replacement flights and adjusting to accommodate the changes *seamlessly*.

In addition to re-accommodating passengers, airlines must evaluate the broader impact on their flights and the associated logistical elements. This includes assessing the availability of aircraft, adjusting operational frequencies, and managing the overall timing of flights. The goal is to optimize the scheduling changes to minimize passenger disruptions and the airline's operational efficiency.

The complexity of this task underscores the need for a sophisticated system that can analyze various interconnected factors simultaneously. This system should be capable of identifying potential issues, proposing solutions, and making informed decisions about the most appropriate course of action. By automating this process, airlines can streamline their operations, enhance customer satisfaction, and navigate the dynamic and ever-changing landscape of the aviation industry without difficulty.

Project Objectives

- Develop a system that organizes flights according to arrival time and the impact on various supplementary services purchased.
- Assess and prioritize passengers for re-accommodation by considering factors such as being an unaccompanied child, an employee on duty, customer loyalty ratings, and the paid class of service. This involves creating an appropriate assessment framework that dynamically adapts to changing scenarios as given by the user.
- ❖ To showcase the system's capability to modify business rules governing passenger and flight rankings dynamically. This ensures flexibility in responding to evolving business needs, providing a solution that can easily adjust to changes in criteria and preferences.
- Implement a ranking system that objectively evaluates and prioritizes different solutions. This involves considering various parameters and criteria and ensuring the ranked solutions align with the airline's goals and priorities.
- Create the files for the flight and passenger re-accommodation solutions, then rank them. Also, to perform the Business Profiling of the entire procedure. Adding an alternative emailing option for affected passengers to choose the flight is a bonus.
- Set clear and actionable guidelines regarding the expiration of specific data or decisions. This ensures that information remains relevant and up-to-date, contributing to the overall efficiency and effectiveness of the system.

Procedure

- Platform Choice: We opted for Google Colab due to its collaborative nature and seamless integration with Google Drive, enabling simultaneous code and folder updates.
 Files like <u>Inventory_dataset</u>, <u>SCH-ZZ_data</u>, and <u>Passenger_Booking_data</u> from the Mphasis data models (Google Drive) were moved to a new folder accessible in Colab, overcoming the view-only file limitation.
- **Data Processing:** The above-mentioned Excel files were transformed into DataFrames (inventory_df, flight_df, and passenger_df) for ease of modification. Specific functions were defined to calculate scores, such as calculate passenger score for passenger df.
- *Passenger Score Calculation*: The function calculates scores based on SSR, cabin, class, downline, and pax. Class-to-cabin mapping is defined in the ruleset, and class scores are assigned alphabetically. SSR data is assumed to be 0 or 1(when this report is written). The total score is the sum of cabin, class, SSR, downline, and pax scores.

For passenger df, we use function(calculate passenger score):

- We can find out the PNR-based score given in the rule set. So, we define a function that calculates the score based on ssr, cabin, class, downline, and pax.
- From the class given, we can find the respective cabin for which the mapping is provided in the ruleset.
- Some classes are not assigned to any cabin, so the remaining classes are assumed to be given to cabin 'Y.'
- The scores for cabins J, K, and Y are given in the ruleset, but for class, the range in which the score lies is given, i.e., 500 to 1000. So, we have assigned a score of 500 to class A, which keeps increasing alphabetically until a score of 1000 is assigned to class Z.
- For SSR, we took it as either 0 or 1.

The total score is calculated by the sum of cabin+class+ssr+downline+pax scores.

Passengers for an impacted flight will be assigned to an alternate flight based on the total

score obtained for an individual PNR. (A higher score will be assigned first)

Flight Data Processing: For flight_df, an impacted flight is selected using ScheduleID and DepartureDate. Constraints from the ruleset are applied, removing flights with ETD > 72 hours. Total scores for each flight are computed using the calculate_flight_score function, considering factors like arrival_delay, SPF, equipment matching, city pairs, and stopovers.

Methodology for flight_df

- 1. Suppose we select an impacted flight to sort out the available alternate flights. In that case, we can give the ScheduleID and DepartureDate of the flight as input to extract the impacted flight data (unique for each flight).
- 2. Now, we have a constraint in the Ruleset. If ETD is >72 hours for the proposed flight, we reject this flight. So, if the Departure time of the proposed flight is > 3 days away from the impacted flight, they are removed.
- 3. For the above-proposed flights, the total score for the flight is computed. We use the function (calculate_flight_score) for each flight to compute the total score.
- 4. Arrival_delay is calculated from the difference in the impacted and proposed flight departure dates. According to the ruleset, the scores are assigned based on the arrival delay.
- 5. Similarly, SPF is also graded according to the ruleset, i.e., the difference between the arrival time and departure time of a flight.
- 6. We assumed the Equipment No of a flight as ScheduleID of the same flight. So, the score is calculated if the equipment no of the impacted flight matches the proposed flight according to the ruleset.

Scores based on *citypairs* and *Isstopover* will also be computed. We used function assign(grade) to assign grades based on their total score:

This is used to eliminate low-grade flights from the solution.

• *Inventory Data Processing*: For single impacted flight passengers allocated to a single alternate flight, inventory data for the proposed flight is obtained, and available seats are distributed equally among classes. Passengers are assigned based on their total score, considering class seat availability. Unallocated passengers go to the exception list.

• Methodology for inventory df:

For single impacted flight passengers allocated to a single alternate flight case:

- In this instance, we select one of the suggested flights and acquire the inventory information for the suggested ScheduleId and DepartureDate (ScID & Ddate are unique for an inventory).
- Oversold (FC_Oversold) indicates the number of first-class tickets that are still available.
 Additionally, FC_CD provides the total number of seats in a first-class together with the corresponding classes ('F', 'P'). However, the number of seats available in each class is not stated. Therefore, we assumed the seats would be the same for all classes.
- To ensure that no class in the first class has more seats than available, function(distribute_equally) distributes the seats across all classes equally, likewise for the other classes.
- Based on their overall score, passengers will be assigned. If seats are still available in the
 requested class, passengers will be assigned to that class. If not, they may be placed in
 other classes, as specified in the Ruleset, if there are open seats.
- Passengers who are not assigned will be added to an exception list.

For single impacted flight passengers allocated to multiple alternate flight cases:

- This method is effective only when multiple flights are being considered. The underlying logic remains the same; however, in cases where passengers remain in the exception list after being assigned to a single aircraft, these passengers are distributed among the other proposed flights.
- In the case of several impacted flights, passengers are assigned to a single alternate flight.

This is done by inputting the Schedule ID and Departure Date of the affected flights. In this case, passenger data is aggregated based on the number of affected flights, while everything else stays the same.

- In the case of multiple impacted flights, passengers allocated to multiple alternate flights case can be done using the same logic as stated above.
- *Handling Multiple Scenarios*: The process is adapted for cases where passengers are allocated to multiple alternate or impacted flights.
- **Solution Ranking:** Solutions are ranked based on low scores for a high exception list, lower mean arrival delay, and other weights specified in the ruleset.

The solution is ranked based on the following factors:

- Low score for high exception list.
- More score for less mean arrival delay.
- The remaining weights are given as mentioned in the Ruleset.

Results-

We only conducted the 1-1 case. We have to provide input to choose a flight that has been affected.

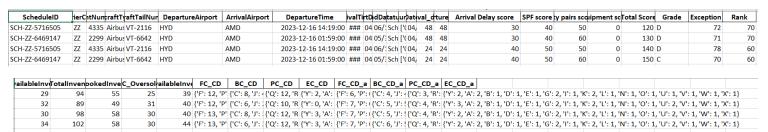
- The selected_DEP_KEY provides a distinct identifier that pertains to the affected flight data. This is the sole input required to determine an affected flight.
- The PNR numbers are taken from the PNRB-ZZ-20231208_062017.csv file based on the DEP_KEY that matches the selected_DEP_KEY. Thus, we obtain the passenger data at the PNR level for those affected.
- The passenger-level data, including names and phone numbers, is extracted from the Passenger Booking data file using the PNR-level data.
- The PNR and passenger-level data columns have been combined to obtain the necessary data in a unified data frame. Originally, we possessed coach data for passengers, categorised as F, J, Q, and so forth. Currently, we only possess class-specific data, such as First class and Business class. Each passenger will be assigned coaches randomly, depending on their class. If a passenger has a first-class seat, they will be assigned an F or P coach. The same applies to passengers in other classes.
- By utilising this consolidated data frame, scores are computed. For example, scores such as Class, PAX, and Downline scores.
- Using selected_DEP_KEY, we can get the Available_flight data, which has ETD < 72 hours and (DepartureAirport and ArrivalAirport) same as the impacted flight.
- These flights are scored accordingly (Arrival Time delay score, SPF score, etc).

By utilizing the Available_flight data, we can extract the inventory level data from the INV-ZZ-20231208_041852.csv file by using the Available_flight data.

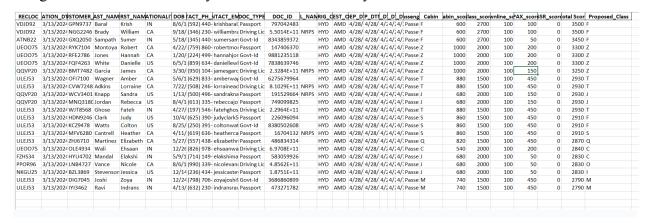
- As the distribution of oversold seats among coaches is unknown, we allocate them evenly, as done with the prior data set. (It is located in the preceding process)
- Additionally, we own both passenger data and inventory data. Passengers are assigned seats based on the availability of seats in their coach type. If seats are unavailable in their coach type, they are checked for seat availability in their suggested class. They are placed on the exception list if seats are unavailable in either their coach type or proposed class. The allotted passengers are derived by excluding those listed in the exemption list from the initial affected passengers. (The same methodology was applied to the prior data set

as well).

- Currently, the solutions are prioritized according to the total score obtained by combining
 the exception list score and the average arrival time delay score. The score for the
 exception list is calculated as 100 minus the number of passengers on the exception list.
- The mean arrival time delay score is calculated by subtracting the arrival time delay from 100. The passengers, exception list, available inventory, and solution ranking of the available flight data are stored in the **Solutions.xlsx** file.
- Example-



Available inventory as shown above, FC_CD shows the total seats in First class and its coaches. FC_CD_a shows the available seats computed from Oversold data. Passengers are allocated using these available seats only. It is also done similarly for other classes.



The allocated passenger's data contains the score of passengers, their initial class before and allocated class (proposed class) after, etc. The exception list is prepared similarly.

Passenger Emailing Facility

We have created an additional facility to email the passengers about the re-scheduled flight using the Gmail API. The automated email system employed by the team utilises the Google Cloud Platform and integrates the Gmail API. This sophisticated system incorporates HTML templates for various elements such as plane tickets, email formats, and other appropriate send-out features. The solution file is downloaded to a server (facilitated by a laptop). Subsequently, the system dispatches updates to all passengers via email, keeping everyone informed of updated information.

Business Profiling

Business Concept: The application is designed to address the challenges associated with planned schedule changes in the airline industry. The primary objective is to identify optimal and best alternate flight solutions for all impacted passengers, ensuring compliance with a set of provided rule sets. It goes beyond mere rebooking by using advanced optimization algorithms and ranking mechanisms to enhance the overall passenger experience. Overall, our application is positioned to revolutionize the way airlines handle planned schedule changes, providing a comprehensive and efficient solution that prioritizes passenger satisfaction while adhering to airline-specific rule sets and policies.

Target Audience: An ideal target audience for this application consists of various commercial airliners. This will satisfy the need of airliners to accommodate passengers in the case of rescheduling, which is becoming necessary in the modern, volatile world where everyday uncertainties can cause trouble for an airline seeking positive customer feedback.

Market analysis and overview: The industry is scattered with various other airline reservation software like TravelPerk, myCWT, Deltamatic (used by American airliner Delta Airlines), Amadeus (operated by American budget airline Southwest), SAP Concur, and so on. While these reservation systems went along without problems for the most part since its establishment, they happened to backfire heavily during the pandemic. Several passengers were hit with rescheduling conflicts, which went wrong and left disgruntled customers behind. This leaves a permanent scar on the face of the airline's reputation- a lost customer would never come back to the same airline, considering the industry's competitiveness. This is where our application can go into place- an accurate application that leaves the airline satisfied with satisfied customers and assures the return of the customer if they were to fly again.

Marketing and Sales Strategy: The idea would be to focus on local airlines first, build a strong reputation then focus on the larger ones. This is because these airlines require more customer satisfaction to remove possible fliers from the giant airliners. Our product should be able to satisfy and convince them to buy it- the catch being, that there is a limited amount of airlines- so an initial loss (selling the product at below-margin rates) to rake in more airline customers.

<u>Potential Applications of Quantum Computing in this problem with a study</u> on its Effectiveness

Theory-

In classical computing, a bit represents information as either $|0\rangle$ or $|1\rangle$ in Dirac notation, corresponding to the column vectors (1 0)' and (0 1)', respectively. In quantum computing, a qubit is the fundamental unit of information expressed as $|\psi\rangle = a |0\rangle + b |1\rangle$ in superposition. Here, $|\psi\rangle$ is the quantum state, a and b are complex numbers, and the normalization condition is maintained $a^2 + b^2 = 1$. When measuring $|\psi\rangle$, the probability of observing 0 is a^2 , and the probability of observing 1 is b^2 .

- 1. Score Calculation using Quantum Computing- We first attempted to calculate the total scores (passenger score and total flight score assuming these weights and constraints) using Quantum Computing. The technique used was Quantum Fourier Transformation, which has proven to be the simplest way to perform arithmetic on Quantum Computers using the least number of qubits. We performed a sample addition and a multiplication operation, and the code can be viewed here as Arithmetic.ipynb.
- 2. Simulating the multiplication of large numbers on quantum computers demands a substantial number of qubits, and the simulation runtime surpasses one minute. Even with the IBM Eagle providing a maximum of 127 qubits and scalable technologies struggling to exceed 5000 qubits, executing intricate arithmetic operations like calculating the total scores in the flight re-routing problem on quantum computers appears impractical. In such cases, classical calculations are preferred due to current limitations in quantum

computing capabilities.

3. Quantum Machine Learning- Quantum Approximation Optimization Algorithm and Quantum Annealing

To offer flexibility in business rules, instead of the classical optimisation techniques, we could use Quantum Machine Learning (QML), which uses the quantum properties of particles to implement ML and optimisation algorithms and shows higher success with more constraints.

The first step is formulating the algorithm, considering all the constraints and weights. If we don't have pre-defined data (for testing flexibility in business rules), we could transform this formulation into a QUBO (Quadratic Unconstrained Binary Optimization) or an Ising Hamiltonian, depending on the type of problem. This technique is used for combinatorial optimization problems.

Conversion to QUBO

The following code below converts the problem into QUBO form using the QUBOvert package in Python. The QUBO form can then be optimised based on the company's requirements using techniques like Quantum Annealing, with the use of simulated annealers like PyQubo (D-Wave Quantum Annealer)

[∞] QUBOvert.ipynb

The QUBO is then converted to the Hamiltonian energy form. After obtaining the Hamiltonian, the next step is optimisation to find the lowest possible energy, representing the optimised function analogous to the ground state in quantum machine learning. Traditional Quantum Approximate Optimisation Algorithm (QAOA) might not be ideal for this problem due to its classical-hybrid nature and challenges in ranking alternative solutions. Instead, Quantum Annealing (QA) offers a promising approach. Similarly, we could also convert the problem to Ising Hamiltonian Formulation using *dimod* or *PyQubo* packages.

Quantum Annealing involves using a specialized quantum annealer, such as the D-Wave

Annealer, to find the energies of the Hamiltonian. Multiple rounds of annealing can be executed to discover the lowest energy state, considering all parameters as the optimized cost function. QA has an advantage in increasing the probability of reaching a global minimum, providing more accurate results. D-Wave provides the PyQubo package, facilitating the direct conversion of the problem into the Ising Hamiltonian and optimization. Additional samples can be generated to identify the second-lowest energy state, representing the second-best alternative, enabling the ranking of solutions.

Furthermore, QA can be leveraged to determine the ranking by addressing the requirement to rank alternative flight routes and present choices to customers. The solution with the second-lowest energy becomes the second-best, the third-lowest for the third-best, and so on. This ranked information can be used to offer customers alternatives in a prioritized manner. This technique can only be used if we wish to test the flexibility of any algorithm for different weights and cannot be implemented for pre-defined data, as given in this problem statement.

Conclusions

Thus, this elaborates our solution for the flight-rescheduling problem, and the solution files, code, and other necessary files have been zipped. We have used classical optimization to work on our solution based on certain assumptions. The applications of Quantum Computing have been mentioned, and business profiling and the emailing facility have been included in the report file.