AIRPORT ENGINEERING

TERM PAPER

Summary & Analysis- *Flight delay prediction from spatial and temporal perspective*

Kalyani S Nair

21064014

Department of Civil Engineering

INDIAN INSTITUTE OF TECHNOLOGY

(BANARAS HINDU UNIVERSITY)

VARANASI – 221 005

1. INTRODUCTION

The swift escalation of air traffic demand and constrained airspace have increased the pressure on the air transportation system, resulting in a heightened incidence of flight delays. Significant economical and logistical burdens are imposed on airlines and airports due to flight delays. Flight delays are costly for airlines, airports, and passengers alike. According to the Federal Aviation Administration (FAA), flight delays cost U.S. airlines and passengers over $30 billion annually, including fuel costs from idling planes, additional staffing, and maintenance expenses associated with extended turnaround times. For passengers, delays cause missed connections, rescheduling costs, and lost business. Predictive models for flight delays are crucial in pre-emptively addressing potential delays further improving the service quality. In their 2022 research paper- “Flight delay prediction from spatial and temporal perspective”, Qiang Li and Ranzhe Jing present a unique integration of complex network theory and temporal modelling to capture key elements influencing delays (Li & Jing, 2022). This study introduces a novel spatio-temporal (ST-Random Forest) model designed to predict flight delays with higher accuracy. The spatial aspect leverages complex network theory to assess the connectivity between airports and account for potential cascading effects of delays in interconnected flights. Temporal elements are modelled using Long Short-Term Memory (LSTM), which are adept at analysing sequence (time series) data, such as the progression of weather patterns or crowding conditions at peak hours. The objective of this report is to provide a critical examination of Li and Jing’s methodology, assessing both its strengths and limitations in the context of aviation and airline management.

2. BACKGROUND

2.1 IMPORTANCE OF ACCURATE DELAY PREDICTION

One of China's biggest airports, Beijing Capital International Airport (PEK), had a 72.74% on-time departure performance in February 2018 (Variflight, 2018). The effects of these delays on the country’s economy are huge. In February 2018, the anticipated delay cost for departure flights alone at PEK reached up to 55 million euros, based on a benchmark recovery cost of 81 euros per minute (Cook, 2015). This data emphasizes how urgently better delay prediction systems are needed, particularly in airports with heavy traffic. Beyond direct costs, delays also complicate the logistical management of airports. For instance, when flights are delayed, the planned sequence of gate assignments is disrupted, often resulting in a cascade effect that can delay several subsequent flights.

For airlines, managing delays involves more than just rescheduling flights. Crew schedules, aircraft assignments, and passenger accommodations must all be changed in real time. Delays may cause crew members to hit their duty hour limits before flights are finished, requiring last-minute replacements, making crew scheduling particularly difficult. This can result in additional cascading delays across the airline's schedule in addition to increasing operational complexity. Flight delays directly affect passenger satisfaction, which in turn influences airline loyalty and reputation. In highly competitive markets, frequent delays can erode an airline’s market share, driving customers toward carriers with better performance records. Therefore, minimizing delays not only enhances operational efficiency but also reinforces an airline’s competitive standing in the market.

2.2 EXISTING CHALLENGES

Flight delays in high-traffic airports are largely driven by overcrowded airspace, limited runway availability, and rapidly changing environmental conditions. To effectively model these issues, predictive systems must process large data volumes and extract nonlinear patterns within them. Conventional models fall short when it comes to this. Conventional delay prediction models often rely on macro factors, such as seasonal trends, airport types, or historical averages, to assess delay likelihood. They often lack the accuracy to handle complicated, real-time circumstances like abrupt weather changes, shifting traffic volumes, or unanticipated operational concerns, even though they can offer broad insights. For example, traditional models may be unable to dynamically adapt to day-to-day or hour-by-hour variations, but may be able to predict increased delays during periods of high travel demand. This restricts their use in circumstances where prompt, flexible reactions are essential, including during unforeseen traffic jams or unfavourable weather conditions. Models that can manage real-time data, recognize intricate patterns, and generate adaptive predictions are becoming even more crucial as airports continue to grow and aircraft volumes rise.

3. SUMMARY OF PAPER

3.1 RESEARCH QUESTION

The main objective of the research paper by Li and Jing is to investigate how integrating spatial and temporal data can improve flight delay predictions. Specifically, it aims to determine whether a spatial-temporal model, combining complex network theory and LSTM-based temporal analysis, can capture delay patterns more effectively than traditional models. The goal is to enhance prediction accuracy for real-time applications in airport operations.

3.2 METHODOLOGY

Data- Departure and arrival flights crossing the Chinese airspace between Jun and Aug 2016 from VariFlight, containing 762,415 samples connecting 260 airports. The real dataset containing a wide range of attributes including OD airports, airlines, tail number, scheduled and actual departure and arrival times and weather conditions was split into three sub-datasets: 80% for model training, 10% for validation and 10% for model testing.

Feature extraction- It refers to the process of identifying and creating specific, relevant data points (or features) from the raw aviation data to improve the accuracy of the flight delay prediction model. Each feature provides important information about factors that can influence flight delays, such as airport congestion, weather patterns, and operational practices.

- **SPATIAL FEATURES**: This is a graph-based representation of the aviation network where:
  1. Nodes represent airports.
  2. Edges represent flight connections from one airport to another.
  3. Arrows in edges represent flight direction.

  Key metrics used:

  1. **Adjacency matrix**: An adjacency matrix is an n×nn \times nn×n matrix where nnn represents the number of airports. Each entry in the matrix

represents the number of flights from one airport to another. For each timeframe (e.g., 3 PM - 4 PM), two adjacency matrices are created

Scheduled Flights: Matrix representing scheduled flight connections.

Actual Flights: Matrix representing actual flight connections.

2. **Crowdedness degree of airport**: The crowdedness degree of an airport is the total number of flights arriving at and departing from that airport. The crowdedness degree is the sum of entries in a row (or column) of the adjacency matrix. Each airport will have two crowdedness values per timeframe:

   Scheduled Crowdedness: Derived from the scheduled flights adjacency matrix.

   Actual Crowdedness: Derived from the actual flights adjacency matrix.

3. **Crowdedness degree of aviation network**: The crowdedness degree of the entire aviation network represents the total number of flights across all airports in a given timeframe. The crowdedness degree of the network is the sum of all entries in the adjacency matrix, calculated separately for scheduled and actual flights. Two values are available for each timeframe:

   Scheduled Network Crowdedness: Total number of scheduled flights.

   Actual Network Crowdedness: Total number of flights actually operated.

- **ANALYSIS OF TEMPORAL FEATURES**: To capture time-dependent patterns, Long Short-Term Memory (LSTM) models are used. This is important because LSTM models can understand how previous congestion and weather conditions might affect current flight delays.

  The inputs used are:

1. Airport crowdedness (Ia)
2. Aviation network crowdedness (Ib)
3. Past weather conditions (Iwb)
4. Forecasted weather conditions (Iwf)

Each of these temporal features is processed through LSTM layers, allowing the model to capture the progression and persistence of delays over time. Two LSTM architectures are used:

1. Separate models for each input- each of the temporal features separately.

2. Combined multi-input architecture that provides a single output of the fused layer.

The multi-input fused architecture showed better performance. This could be attributed to the fact that the model can learn the dependencies and relationships between the different temporal features early on, rather than modeling each feature in isolation.

- **EXTRINSIC FEATURES:** Additional extrinsic features provide context for understanding delays that arise from operational practices and flight characteristics.

1. Delay of previous flights.
2. Airline issues (turnaround time and flight specific information).
3. Distance between origin and destination airports.
4. Scheduled fly time.

RANDOM FOREST CLASSIFIER (PREDICTION FRAMEWORK): Random Forest is the leading classifier that operates based on decision trees built from samples selected at random as well as subsets of features. It allows ensemble techniques that boost not only accuracy but also robustness by balancing the heterogeneous types of features and attenuating the effects of noise on data.

The inputs used are:

- Spatial features(from complex network analysis of the aviation network)
- Temporal features (from the LSTM models)
- Extrinsic features

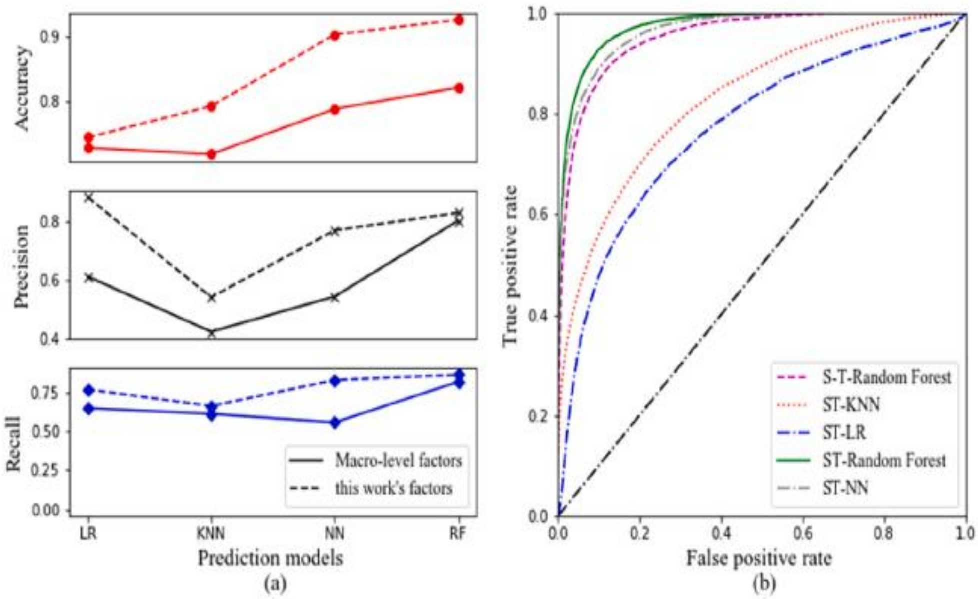
Evaluation- The ST-Random Forest model's accuracy is evaluated against benchmarks. Results show a significant improvement over baseline models. Further details about the performance are provided in the next section.

3.3 KEY FINDINGS AND PERFORMANCE

The above figure shows the metrics of prediction of the proposed model attains an accuracy of 92.39%. Approximately 86% of on-time samples are correctly identified, while delayed samples have a classification accuracy 95%. The suggested model has a strong prediction performance, as indicated by an area under the ROC curve (AUC score) of 0.89.

Insights on model performance:

- The ROC curve indicates a high sensitivity (TPR), as it is close to 1. This suggests that the ST-Random Forest model effectively distinguishes between delayed and on-time flights.
- The ROC curve also shows a low specificity (FPR), indicating that the model has a low likelihood of incorrectly predicting a delay when there is none.
- The confusion matrix shows that 14% of on-time flights were misclassified as delayed. This suggests that the model is slightly more conservative, prioritizing correct delay predictions over on-time predictions.



The above figure depicts the comparison of proposed classifier with 8 baselines:

KNN-macro- k-nearest neighbours, considering macro factors.

RF-macro- Random Forest, considering macro factors.

LR-macro- Linear regression, considering macro factors.

NN-macro- Neural network, considering macro factors.

ST-LR- Linear Regression considering both spatial and temporal features.

ST-KNN- k-Nearest Neighbours considering both spatial and temporal features.

ST-NN- Neural Network with sigmoid classifier considering both spatial and temporal features.

ST-RF- Random Forest considering both spatial and temporal features.

Key insights from the figure:

- ST-Random Forest model achieves the highest accuracy compared to other models, especially when using factors specific to this study.
- Recall varies across models but tends to be lower compared to accuracy and precision. However, the ST-Random Forest model still shows a relatively good recall, indicating a balanced performance in identifying both delayed and on-time flights.

- Overall, ST-Random Forest model combined with the unique factors of this study, delivers a more accurate, precise, and robust prediction of flight delays compared to conventional models like KNN, NN, and LR.

4. ALGORITHM AND MODEL ARCHITECTURE

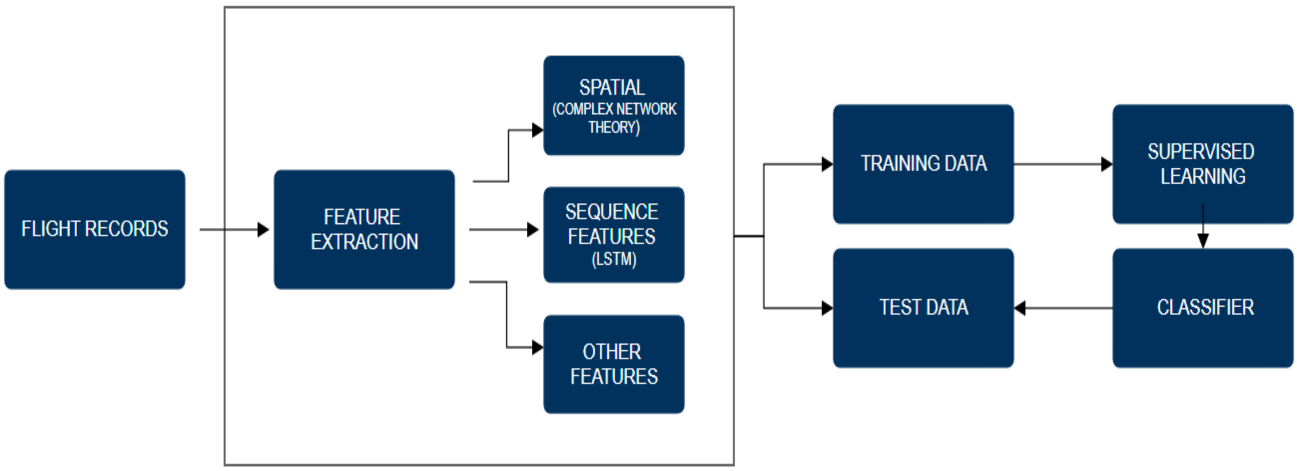


Fig. 4.1 Algorithm of the proposed model.

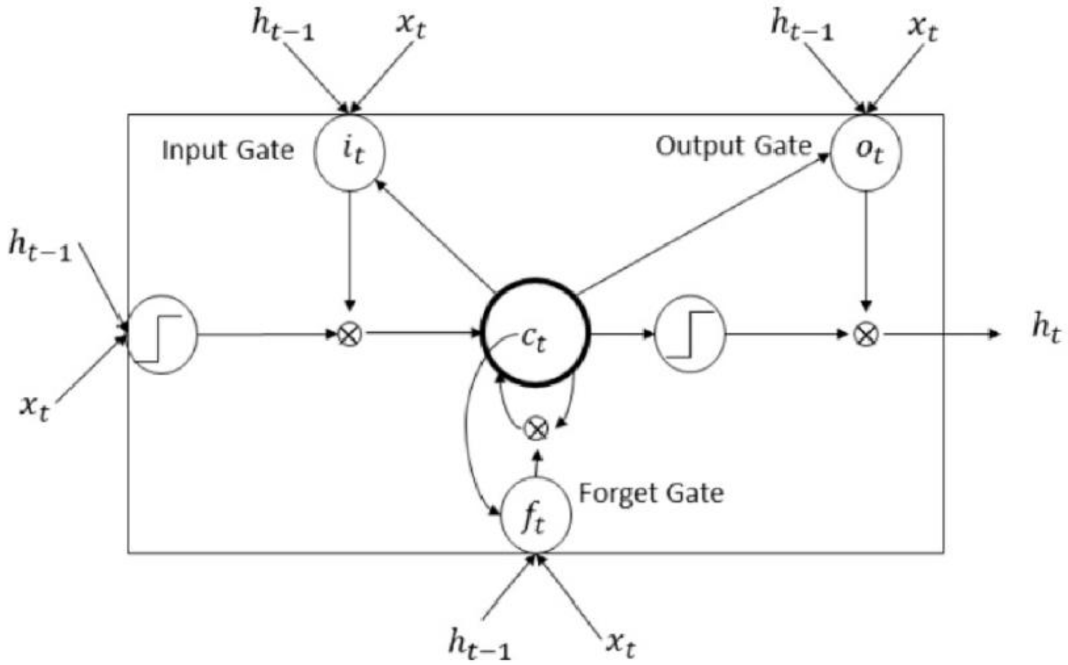


Fig. 4.2 LSTM cell (*Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference On*, 2013).

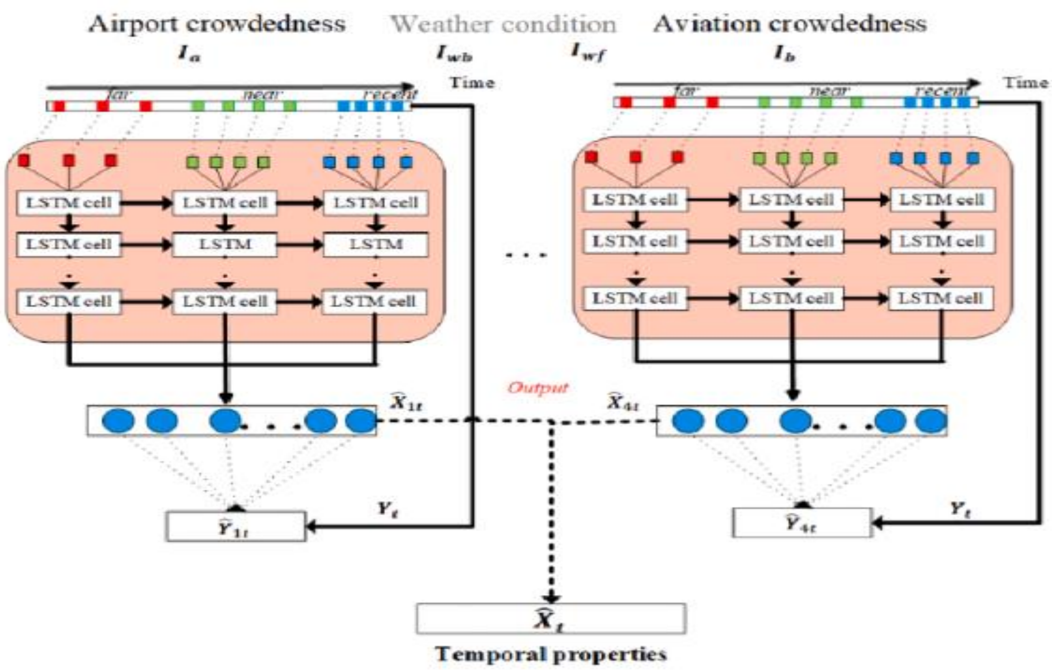


Fig 4.3 Separate Models for Each feature type. (Li & Jing, 2022)

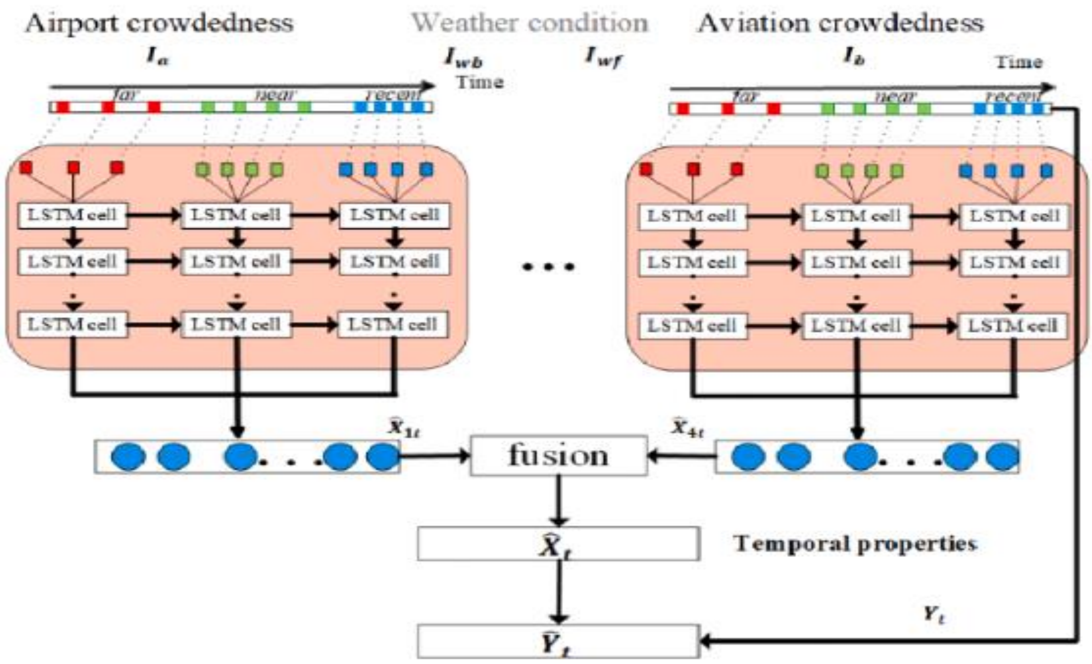


Fig. 4.4 Combined multi-input model. (Li & Jing, 2022)

5. CRITICAL ANALYSIS

5.1 STRENGTHS AND NOVELTY

What distinguishes Random Forest model from traditional delay prediction models is its hybrid structure, which gives a much stronger framework in the real-time spatial-temporal relationships it could capture. Additionally, this would give an improved accuracy rate compared with traditional models. With complex network theory, it was possible to consider airports as nodes and routes as links between nodes, illustrating how the propagation of delay across a network can affect another part of the system. More inclusion of LSTM networks makes it even more accurate since the model could now work on sequential data, for example, on time-varying parameters like the weather and levels of congestion hence capturing dependence over time. This will be of much value especially when the recent data say the weather conditions at a given point in time dictate future output values, thereby enhancing the predictive accuracy of the model. It also employs a Random Forest classifier, which is very effective in handling complex, non-linear data and provides high reliability in predicting categorical outcomes, such as delay or on-time. This makes it suitable for real-time applications where large datasets need to be processed efficiently. This design makes the model scalable for use in airports' and airlines' working systems to quickly adjust and better utilize time according to how long it will take for a flight to take.

5.2 LIMITATIONS

The test was conducted only on Chinese domestic flights which limits the generalizability of this model. Secondly, only limited spatial correlations were analysed which might lead to lower accuracy for such complex, interconnected networks. Also, the high computing requirements of both LSTM as well as Random Forest Algorithm are combined with pre-processing huge datasets. Although the model captures some of the spatial features using complex network theory, it is not able to capture correlations between nearby airports or other neighbouring entities. This might limit its ability to capture the accuracy of densely populated areas where the airports are closely linked and delays at one easily impact others. Although the model works generally well, there still exists a 5% misclassification error in delay versus on-time predictions at 14%. This is indicative of an area of potential refinement to predict more on-time flights. Overfitting is another risk. Overfitting is a tendency of some machine learning models, such as LSTM, especially when dealing with complex and high-dimensional data, such as that of this project, in which not enough representative data exists for all conditions. This risk can reduce the robustness of the model, especially in scenarios that have unusual patterns or rare events not well-represented in the training data.

6. TECHNICAL ANALYSIS OF MODEL COMPONENTS

6.1 COMPLEX NETWORK THEORY

Complex network theory is framework that aids in analysing and interpreting relationships between different components (nodes) connected by edges, forming a network. The aviation network is broad and complex, the spatial characteristics of the network cannot be fully understood based on just airport-level data. This network is made up of airports (nodes) and flight routes (edges). This theory is used to capture spatial characteristics by analysing the structure and behaviour of this interconnected system at different levels- nodes, edges and network. This network structure helps to understand how busy an airport is and how flights interact across the entire system, forming a weighted, directed graph (flights have a specific direction and number).

Individual airports are treated as nodes with certain metrics:

- Degree- Crowdedness/activity level of the airport. It is the number of flights arriving and departing. This study uses 2 measures of degree- scheduled and actual crowdedness.
- Density- Connectedness. Density is the ratio of the number of edges in the network to the maximum possible number of edges in the network.
- Edge betweenness centrality- Routes that appear frequently on shortest paths between airports are considered more central and important. Centrality of an edge illustrates the sum of the fraction of all-pairs of shortest paths that pass through it.

6.2 LSTM NETWORKS

Many factors influencing the delay of flights are time-dependent and have cascading or cumulative effects. Factors such as weather conditions, air traffic and air congestion often have non-linear and interdependent relationships and do not affect in isolation, they usually have a ripple effect on flight delays. Flight delays are often interconnected, a delay in one flight might lead to further delays throughout the day. This makes it important to capture historical delay patterns to understand how they propagate.

LSTM (Long Short-Term Memory) is a type of Recurrent Neural Network (RNN) that is particularly useful in capturing the temporal features as it excels in capturing patterns in sequence-based data (time series data). LSTMs are designed to remember information over long time spans. In contrast to traditional Recurrent Neural Networks (RNNs), which struggle with the vanishing gradient problem (less effective at long-term memory), LSTMs use special mechanisms to retain and selectively forget information. This allows the model to learn relationships across both recent and distant time points. LSTMs have a unique architecture with gates (input, forget, and output gates) that control how much information is stored or discarded. LSTMs, can capture the complex, non-linear trends of airport crowdedness and weather conditions in a way that simpler models cannot, making them ideal for understanding how these factors evolve and interact over time. This makes them an ideal choice for integrating temporal features into flight delay prediction models, helping improve the robustness and accuracy of predictions.

6.3 RANDOM FOREST

Random Forest (RF) is an ensemble learning algorithm that combines multiple decision trees to create a more stable and accurate model than a single decision tree alone. It has the ability to handle diverse features. In the context of the current study, the model combines three types of features- temporal, spatial and extrinsic. RF allows integration of these diverse features without needing any extensive preprocessing. Flight delay data can be noisy due to unpredictable factors like sudden weather changes or last-minute operational issues. Since RF uses an ensemble of trees, it is more robust to noise and reduces the risk of overfitting. RF provide built-in feature importance scores, allowing researchers and airport operators to understand which features are most influential in predicting delays. This interpretability is valuable in the aviation context, as it enables stakeholders to focus on the key factors driving delays, such as specific weather conditions or high-density routes, and make data-driven decisions accordingly. The above-mentioned benefits make it a strong choice for the task of flight delay prediction, where real-time accuracy and robustness are essential.

7. CONCLUSION

The study tackles the issue of flight delays, that lead to significant economical losses in the aviation industry. The study was performed on domestic flight data from China (June to August 2016). A novel and highly accurate prediction model ST-Random Forest is proposed. It combines spatial and temporal data by incorporating complex network theory and LSTM (Long Short-Term Memory) methods. Complex network theory is used to extract spatial features of the aviation network at different levels (edge, node, and network), allowing the model to understand how airports and flight routes are interconnected. The LSTM component captures temporal dependencies, such as recent patterns in crowdedness and weather, which are crucial in influencing delays. This combined approach enhances the model’s predictive accuracy, outperforming other conventional methods. The ST-RF model in intended for integration into airport and airline information system, it helps operators keep an eye on possible delays and make real-time adjustments to flight scheduling, runway and gate allocation. This helps airlines improve operations, lower passenger dissatisfaction and lessen financial effects.

8. FUTURE RESEARCH

- Test the model with international data allowing the model to capture a variety of delay patterns influenced by varying international weather conditions, air traffic control protocols, and operational practices.
- Enhance spatial features by incorporating spatial correlations across neighbouring airports. Adding these correlations would enable the model to predict how delays at one airport might propagate to nearby locations, particularly in high-traffic areas where interconnected delays are common.

9. REFERENCES

Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on. (2013).

Cook - University of Westminster, A. (2015). *European airline delay cost reference values - updated and extended values (Version 4.1)*.

Li, Q., & Jing, R. (2022). Flight delay prediction from spatial and temporal perspective. *Expert Systems with Applications*, *205*. https://doi.org/10.1016/j.eswa.2022.117662