

Predicting Taxi Times Using Airport Surface Movement Data

1st Min-Kyun Park

*Department of Aerospace Engineering
and the Program in Aerospace Systems Convergence
Inha University
Incheon, Republic of Korea
par2k1157@inha.edu*

2nd Hyeon-Su Hwang

*Department of Aerospace Engineering
Inha University
Incheon, Republic of Korea
hyeonsss@inha.edu*

3rd Jae-Young Ryu

*Department of Aerospace Engineering
and the Program in Aerospace Systems Convergence
Inha University
Incheon, Republic of Korea
JaeyoungRyu@inha.edu*

4th Hak-Tae Lee

*Department of Aerospace Engineering
and the Program in Aerospace Systems Convergence
Inha University
Incheon, Republic of Korea
haktae.lee@inha.ac.kr*

Abstract—Taxi time prediction is an essential component of efficient airport surface traffic management, especially for arrival and departure scheduling. In this study, the surface movement data at Incheon International Airport is analyzed to find taxi routes and taxi speed distributions, and then a taxi time prediction model is developed that can reflect the current surface traffic conditions. The airport surface is modeled using a node-link structure so that a route is expressed by a sequence of nodes and links. Previous studies have mainly estimated taxi times based on the average speed of each link, or developed prediction models using regression or machine learning techniques with limited traffic variables. Building on these approaches, this study uses a decision tree-based model, eXtreme Gradient Boosting, to more accurately predict the taxi time by incorporating the ground traffic conditions. Features are divided into two categories, one is the real-time observational features that are captured at the initial time snapshot of the target aircraft such as total number of aircraft on the surface and the number of aircraft along the taxi route. The other is the route-specific static features such as route lengths and number of links. In addition, surface movement statistics obtained from the dataset such as average taxi speeds are included in these static features. The importance of each feature is investigated and with the final feature set of 13 features, an R^2 value of 0.74 was achieved, where the taxi times are predicted within $\pm 20\%$ for most of the routes. The proposed method is expected to be an effective tool for improving the surface traffic management such as arrival and departure scheduling or traffic flow management.

Index Terms—Taxi Time Prediction, Node-Link Model, Airport Surface Operations, ASDE-X, Feature Engineering, XGBoost

This work was supported by the Korea Agency for Infrastructure Technology Advancement (KAIA) grant (RS-2021-KA163373) funded by the Ministry of Land, Infrastructure and Transport of the Korean government and by the National Research Foundation of Korea (NRF) grant (RS-2022-NR070854) funded by the Ministry of Science and ICT of the Korea government.

*Hyeon-Su Hwang is currently with Hanwha Aerospace, Daejeon, Republic of Korea.

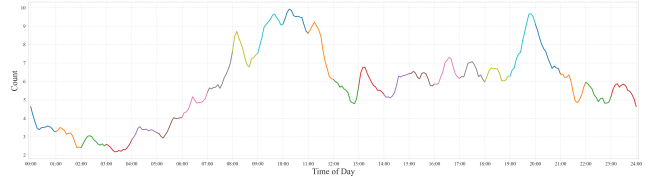


Fig. 1: Number of departing aircraft by hour of day

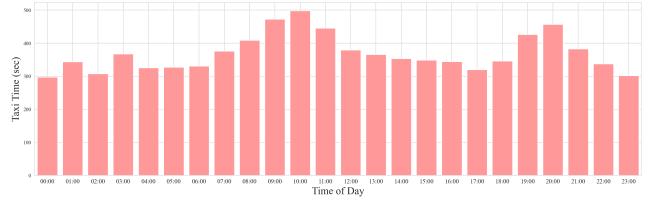


Fig. 2: Taxi time by hour of day

I. INTRODUCTION

The continuous growth in air traffic demand, coupled with the physical limitations of the airport infrastructure, has led to increased surface congestion at major airports [1]. Taxi time has become a critical factor in airport operational efficiency and flight schedule management, directly impacting various aspects of airport operations, such as take-off scheduling and runway capacity management. In particular, uncertainty in taxi times can cause takeoff delays and bottlenecks in surface traffic flow, which in turn adversely affect the overall efficiency of airport operations.

To reduce surface congestion and improve flight punctuality, it is essential to predict taxi times quickly and accurately. Accurate taxi time prediction is a critical foundation for Airport Collaborative Decision Making systems [2], helping

to optimize runway and taxiway traffic and reduce departure queues and delays.

Surface operational conditions at airports exhibit different patterns depending on various factors such as the day of the week, time of day, and the number of aircraft on the ground. These operational characteristics are expected to directly influence surface movement flows and, consequently, taxi times. As can be seen in Figs. 1 and 2, average taxi times vary depending on the time of day.

Therefore, leveraging airport operational patterns allows for more precise taxi time prediction, and it is necessary to design prediction models that reflect such patterns. This study aims to incorporate the patterns of surface traffic in terms of the statistical taxi speed characteristics into the input features.

Previous studies have modeled airport surfaces using node-link structures and proposed taxi time prediction methods based on average link speeds [4], [5] or polynomial regression models [6]. [3] produced a statistically significant results by first predicting the size of the departure queue and estimating the taxi-out time using the queue size. In addition, some research has used machine learning approaches such as random forests to improve prediction accuracy [7]. These studies have helped to improve the accuracy of taxi time prediction by incorporating specific route characteristics or aircraft types.

Building on these previous approaches, this study proposes a taxi time prediction method that comprehensively reflects both the ground traffic conditions observable when the target aircraft enters the initial node and the inherent route-specific characteristics. This allows the model to predict the taxi time without the need for complex forecasting of future surface conditions. While machine learning techniques such as XG-Boost have been used in previous studies [8], [9], this study further incorporates a node-link-based network structure into the feature generation process to explicitly capture the structural characteristics of individual taxi routes. By leveraging the model's fast training and inference capabilities, the proposed system aims to enable scalable, potentially real-time taxi time prediction in large airport environments.

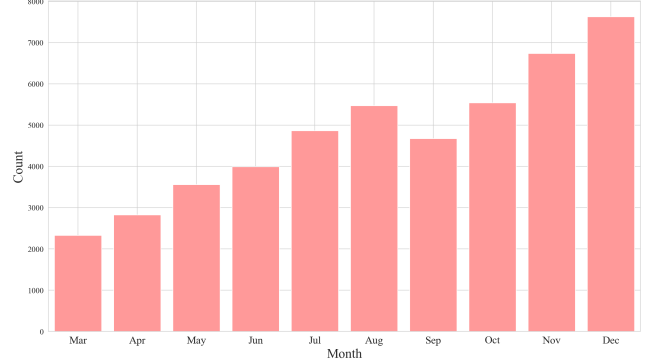
The remainder of this paper is organized as follows. Section II describes the dataset and the preprocessing procedures. Section III defines the problem and presents the key feature engineering process. Section IV details the development of the prediction model, while Section V presents the experimental and performance evaluation results. Finally, Section VI concludes the study and suggests directions for future research.

II. DATA AND PROCESSING

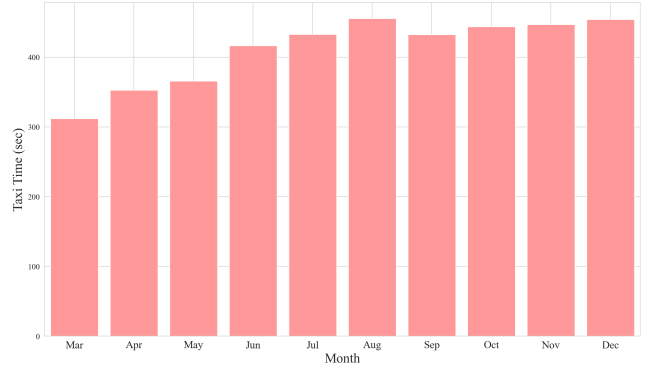
A. Source of Data

This study uses surface movement data of aircraft operating at Incheon International Airport (ICN). The primary data sources include surface trajectory data collected by the Airport Surface Detection Equipment, Model X (ASDE-X) [10], and airport infrastructure information provided by the Aeronautical Information Publication (AIP) [11]. The ASDE-X data contain high-resolution records of aircraft position, time, speed, and other variables, enabling precise tracking of ground movement

paths. The AIP data provide detailed geographic information on gates, ramps, taxiways, and runways, which has been used to structure ground movement paths into a node-link framework.



(a) Number of departures



(b) Average taxi out time

Fig. 3: Number of departures and average taxi out time by month

The ASDE-X data used in this study consists of 47,984 departure trajectories collected from March to December 2022 for which taxi routes could be clearly identified. The average number of aircraft per day over the entire period was 159.

These data were preprocessed based on the method proposed in [12]. In the filtering step, consecutive trajectory points with a spatial interval of 10 meters or less were removed, as they were considered potentially redundant or erroneous. As a result of this smoothing and filtering procedure, the average time interval between the remaining trajectory points was approximately 3.7 seconds. In this study, the preprocessed data were further interpolated to a 1-second interval, and the interpolated data were used for the analysis.

During this period, the traffic volume steadily increased as the air travel recovered from the COVID-19 restrictions as shown in Fig. 3a. The average taxi out time peaked in August and remained steady as shown in Fig. 3b.

B. Airport Node-Link Model Construction

In this study, node-link model constructed based on AIP data of ICN [13], as shown in Fig. 4 was used. The model was

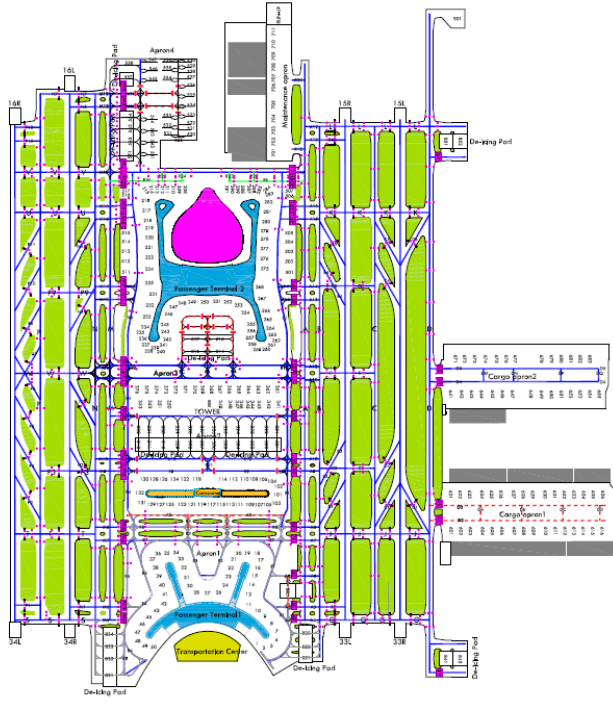


Fig. 4: Surface layout of ICN

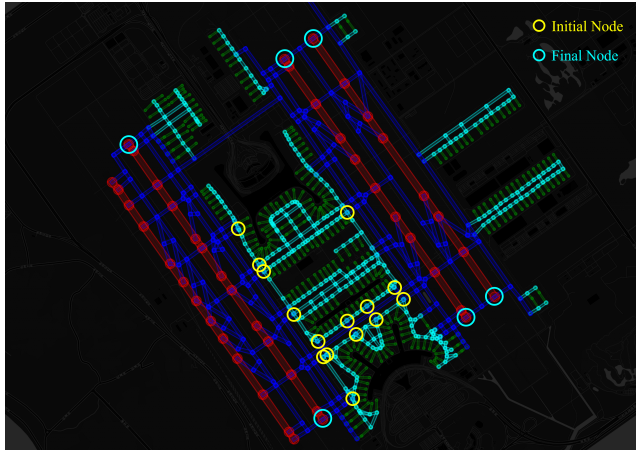
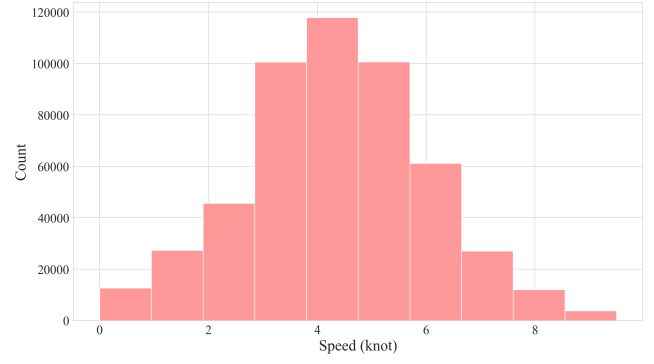


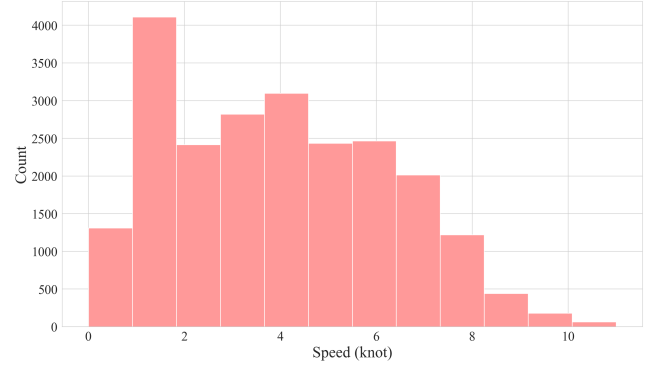
Fig. 5: Node-link model of ICN

developed by referencing the ramp area and taxiways indicated in the AIP, using the geographic coordinates and names of the gates, and by defining intersections as nodes using Google Earth Pro to manually obtain structured latitude and longitude data. As shown in Fig. 5, the model consists of 775 nodes and 1,007 links. Each node is classified as gate, ramp area, taxiway, or runway. This model allows for the conversion of aircraft surface trajectories expressed in a sequence of latitudes and longitudes into a sequence of nodes and links, which enables the extraction of quantitative features such as route length and number of links.

To map the latitude and longitude coordinates of the ASDE-X data to the node-link structure, the approach described



(a) Speed distribution at the most used link



(b) Speed distribution at the least used link

Fig. 6: Speed distributions

in [14] was used. Basically, if a coordinate point is inside a rectangle that encloses a link with a certain threshold, the point is considered to traverse the link. This approach enabled the identification of aircraft positions relative to the network structure to analyze the speed of the aircraft, the stop status, and the presence of a nearby aircraft. Metrics such as average speed within each link, speed distribution characteristics, and time to traverse the links were also quantified. These data provide a critical foundation for a quantitative understanding of ground traffic flow.

Figure 6 shows the taxi speed distribution at two different links. The speed distribution for the most used link shown in Fig. 6a closely follows a Gaussian distribution, while the one for the least used link shown in Fig. 6b shows some variations. In general, the speed distributions at most of the links were close to a Gaussian distribution.

C. Trajectory Data Pre-processing

Some of the trajectories in the data set were incomplete or anomalous. For departures, only the trajectories that originate from gates and pass through runways are filtered. These trajectory data were then mapped to the node-link model as previously mentioned. This preprocessing ensured the reliability of the data and prepared them in a form suitable for model training.

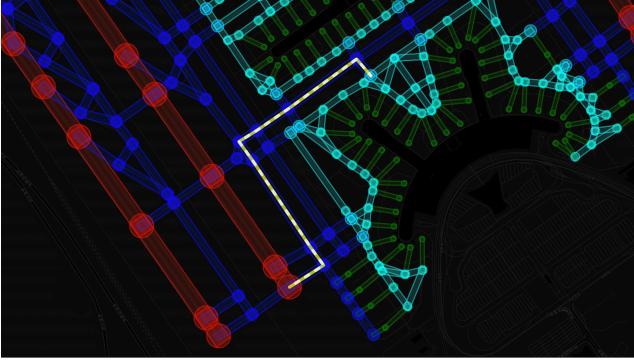


Fig. 7: Route example

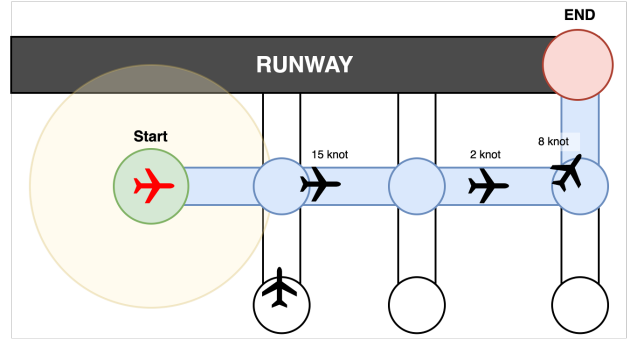


Fig. 8: Taxi route and features

D. Taxi Route Identification

Each aircraft's taxi out path was identified based on its initial node (start point) and final node (end point). The initial node was defined based on the node-link model as the node where the aircraft enters the taxiway from the ramp, that is the last ramp area node. The final node was defined as the node where it enters the runway from the taxiway, that is the runway threshold node. An example of a taxi out route is shown in Fig. 7. Even for identical initial/final node pairs, actual taxi paths could differ, and such differences were taken into account.

Figure 5 shows 15 initial nodes in yellow circles and 6 final nodes in cyan circles. Total of 33 initial/final combinations and 56 routes were identified. During the data collection period, Runway 16R/34L was not used for departures.

III. PROBLEM FORMULATION AND FEATURE ENGINEERING

A. Problem Definition

Predicting taxi time is formulated as a regression problem where the target variable is a continuous value measured in seconds. This differs from classification problems, which select from a set of discrete categories. The input variable X consists of information available at the moment when the aircraft enters the initial node including airport ground traffic conditions and route-specific attributes. The output $\hat{y}$ represents the expected taxi time for a flight, and their relationship is expressed as in (1).

$$\hat{y} = f(X) \quad (1)$$

B. Feature Design and Categorization

In this study, input features were systematically designed to improve the prediction accuracy of taxi time. The features were divided into two main categories. The first includes *real-time observational features*, which capture airport traffic conditions observable at the moment when the aircraft enters the initial node. The second category consists of *route-specific features*, which reflect structural characteristics inherent to each taxi route.

Real-time observational features include the total number of aircraft on the ground, the number of aircraft within 1 km of the initial node, the number of aircraft on the planned route, and the number of stopped aircraft. In addition, the average and standard deviation of aircraft speed on the route are taken into account, as well as congestion indicators such as the number of aircraft per route length and the number of aircraft near adjacent routes. Temporal information such as time of day and day of week are considered. Aircraft Wake Turbulence Category (WTC) is also included to reflect the variations in the aircraft performance.

Route-specific characteristics represent structural attributes that are fixed for each route, including the route id, total route length, and the number of links that make up the route. In addition, two features are constructed from the statistical data analysis. As the aircraft speed distributions for each link are obtained, the arithmetic means of each link's average speed and the standard deviation of speed are calculated to represent the flow characteristics.

Figure 8 illustrates some of the key features along a taxi route. The shaded blue segments represent the expected taxi route from the initial node (green) to the final node (red). White segments that are connected to the nodes that are part of the route are the adjacent links. Multiple aircraft may be on this route at this moment. Features such as the number of aircraft on the route, the number of aircraft stopped, and the mean and standard deviation of their speeds are included to reflect these traffic conditions. Especially, the speed differences between segments can have a significant impact on the total taxi time.

Table I and Table II summarize the real-time observational features and static route-specific features defined in this study. Each feature was input to the model in its original form without normalization. In the table I, N_{seg} represents the N_{route} disaggregated by dividing the sequence of links along the route to start, middle, and end segments based on the number of links, and by separately counting the number of aircraft in each segments. Similarly, N_{wtc} separately counts the number of aircraft in five WTC categories, which are light, medium, heavy, super, and unknown. This is expected to increase the granularity of the features.

TABLE I: Real-Time Observational Features

Symbol	Feature	Description
N_{tot}	Total aircraft count	Total number of aircraft present on the airport surface
N_{near}	Aircraft count near the initial node	Number of aircraft located within 1 km radius of the initial node
N_{route}	Aircraft count on the taxi route	Number of aircraft on the taxi route
N_{seg}	Aircraft count per route segments	Number of aircraft in start/middle/end segments (3 sub-features)
N_{wtc}	Aircraft count by WTC	Number of aircraft for five WTC classes (light/medium/heavy/super/unknown) on the taxi route (5 sub-features)
N_{adj}	Aircraft count on adjacent links	Number of aircraft located on all links adjacent to the route
$N_{stopped}$	Number of stopped aircraft	Number of aircraft that is slower than 3 knots along the taxi route
P_{den}	Aircraft count per route length	Number of aircraft on the route divided by total route length
V_{ave}	Average speed	Average taxi speed of all aircraft on the route
V_{dev}	Standard deviation of speed	Standard deviation of the aircraft speeds on the route
T_{hour}	Time of day	Hour-of-day (24 1-hour bins)
T_{day}	Day of the week	Day of week from Monday to Sunday
A_{wtc}	WTC	Target aircraft's WTC

TABLE II: Route-Specific Static Features

Symbol	Feature	Description
A_{id}	Route id	Unique identifier for each taxi route
R_{length}	Total route distance	Total length of the route in meters
N_{links}	Number of route links	Number of links comprising the route
$V_{ave,link}$	taxi speed metric	Mean of the average aircraft speeds calculated for each link on the route
$V_{dev,link}$	taxi speed variance metric	Mean of the speed standard deviations calculated for each link on the route

Furthermore, $V_{ave,link}$ and $V_{dev,link}$ in Table II were calculated by arithmetically averaging the average and the standard deviation of speed for each link, respectively, based on the data such as that shown in Fig. 6.

C. Taxi Route Selection and Dataset Filtering

The taxi routes of the aircraft were classified based on the combination of initial node (start point) and final node (end point). Even for the same initial/final node combination, multiple routes could be used, and the usage ratio for each route varied. An analysis of the average speed distributions for the major routes showed that most of them closely followed a Gaussian distribution. Figures 9 and 10 show the number of uses per route for two initial/final node combination as pie charts, and the average speed distributions as histograms. For the first initial/final node pair, two routes existed where the major route among the two has an 80% share. For the second initial/final combination, only a single route existed.

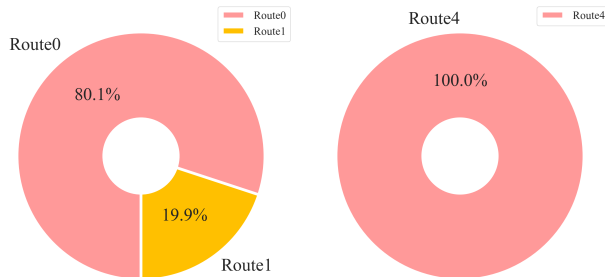


Fig. 9: Usage distribution among different routes for the same initial/final node pair.

To ensure the reliability of the data and the stability of the model training, taxi routes that were used more than 100

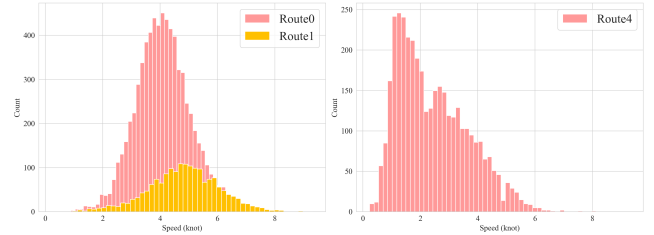


Fig. 10: Histogram of average speed distributions for selected routes.

times for the same initial/final node combination were selected for training and evaluation. Among the selected routes, only samples whose taxi times did not exceed the $\pm 3\sigma$ range of the route-specific mean were retained, resulting in a total of 45,319 trajectory data points used for model training.

Even within the same initial/final nodes combination, each route was treated as an independent learning target if the path was different.

IV. PREDICTION MODEL DEVELOPMENT

A. Overview of XGBoost

In this study, a tree ensemble method based on gradient boosting, namely eXtreme Gradient Boosting (XGBoost) [15], was used to predict taxi times. Boosting sequentially combines multiple weak learners to gradually improve prediction performance, with each stage focusing on correcting the residuals of the previous predictions. XGBoost is particularly suitable for this task due to its excellent efficiency in terms of prediction accuracy, training speed, and flexibility in handling different input variables. In addition, the incorporation of regularization terms helps prevent overfitting and enables stable learning even with large datasets. These features make XGBoost highly suitable for the taxi time prediction problem, where both

surface traffic conditions and route characteristics interact in a complex manner.

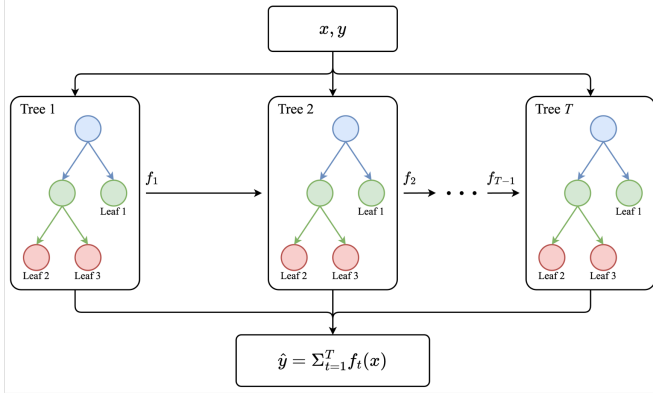


Fig. 11: XGBoost diagram

The XGBoost model operates as a nonlinear regression model that takes an input matrix, $\mathbf{X} \in \mathbb{R}^{N \times k}$, consisting of N samples and k features and predicts the taxi time, $\hat{y}_i$, for each sample. The learning process is performed iteratively, and, at each step, the prediction is updated by adding the output of a newly trained regression tree to the previous prediction. This process is mathematically expressed as:

$$\hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + \eta \cdot f_t(x_i) \quad (2)$$

Where $\hat{y}_i^{(t)}$ is the cumulative prediction for the i -th sample up to the t -th tree, η is the learning rate, and $f_t(x_i)$ is the output of the t -th regression tree trained to reduce the residual error.

XGBoost minimizes a regularized objective function that balances prediction accuracy and model complexity. To optimize this objective efficiently, the loss function $\ell(y_i, \hat{y}_i)$ is approximated using a second-order Taylor expansion:

$$\ell_i^{(t)} \approx \ell_i^{(t-1)} + g_i^{(t)} f_t(x_i) + \frac{1}{2} h_i^{(t)} f_t(x_i)^2 \quad (3)$$

Here, $g_i^{(t)}$ and $h_i^{(t)}$ denote the first- and second-order derivatives of the loss function with respect to the prediction $\hat{y}_i$ at stage $t - 1$.

The overall objective function to be minimized at each iteration includes both the accumulated loss and a regularization term $\Omega(f_t)$:

$$\mathcal{L}^{(t)} = \sum_{i=1}^N \left[g_i^{(t)} f_t(x_i) + \frac{1}{2} h_i^{(t)} f_t(x_i)^2 \right] + \Omega(f_t) \quad (4)$$

The regularization term penalizes model complexity and helps prevent overfitting. While the exact solution involves computing optimal leaf weights and evaluating all possible tree structures, these details are handled efficiently by the XGBoost algorithm internally.

B. Training Setup and Input Feature

The input features were used without normalization and were fed directly into the model. A detailed list of the features can be found in Tables I and II. A total of 45,319 taxi trajectory samples were used to train the model. The entire dataset was randomly divided into training and test subsets in a 7:3 ratio. The training data was used to optimize the model parameters, while the testing data was used to evaluate the taxi time prediction performance.

C. Model Training Process

XGBoost provides fast training speed and efficient parallel processing, making it well suited for repeated experiments with different hyperparameter combinations. The main hyperparameters used in this study are summarized in Table III.

TABLE III: Hyperparameter Settings for XGBoost

Parameter	Value	Description
Learning rate	0.05	Tree contribution weight
Maximum depth	5	Max tree depth
Number of estimators	400	Number of trees
Subsample ratio	0.8	Row sampling ratio per tree
Feature subsample ratio	0.8	Column sampling ratio per tree

Hyperparameter tuning was performed using a grid search method combined with 3-fold cross-validation. The best-performing configuration was selected to build the final model. Model training was performed using the prepared training and validation dataset. During training, the XGBoost model was iteratively optimized and performance was periodically evaluated on the validation set to avoid overfitting. Early termination was applied if no improvement was observed after a certain number of epochs. After training, the final model was evaluated on the separate test dataset using the Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Coefficient of Determination (R^2) metrics.

V. RESULTS AND DISCUSSIONS

A. Feature Importance Analysis

Before constructing the XGBoost-based taxi time prediction model, the prediction models performances are evaluated using different combinations of input features. In the initial experiments, the model was trained using only general features such as the total number of aircraft on the airport surface, the number of aircraft near the initial node, the time of day, and the WTC of the aircraft. However, under these conditions, the model showed limited predictive performance, with the R^2 score staying near 0.6 and both the MAE and RMSE values remaining relatively high.

Subsequently, when route-specific static features such as route id, total route distance, and number of route links were added to the input, a significant improvement was observed. In particular, the inclusion of these route features led to a sharp increase in the R^2 score, while MAE and RMSE were significantly reduced. This suggests that, in addition to the overall airport traffic conditions, the taxi route plays a critical role in predicting the taxi out times.

TABLE IV: Feature Combination by Case

Feature	Case1	Case2	Case3	Case4	Case5	Case6	Case7	Case8	Case9	Case10	Case11	Case12
N_{tot}	✓	✓	✓		✓	✓			✓	✓		✓
N_{near}		✓	✓		✓	✓				✓		✓
N_{route}	✓	✓	✓		✓	✓			✓	✓		✓
N_{seg}		✓	✓		✓	✓				✓		✓
N_{wtc}		✓	✓		✓	✓				✓		✓
N_{adj}		✓	✓		✓	✓				✓		✓
$N_{stopped}$		✓	✓		✓	✓				✓		✓
P_{den}		✓	✓		✓	✓				✓		✓
V_{ave}			✓			✓					✓	✓
V_{dev}			✓			✓					✓	✓
T_{hour}				✓	✓	✓	✓				✓	✓
T_{day}				✓	✓	✓	✓				✓	✓
A_{wtc}				✓	✓	✓	✓				✓	✓
A_{id}							✓	✓	✓	✓	✓	✓
R_{length}								✓	✓	✓	✓	✓
R_{links}								✓	✓	✓	✓	✓
$V_{ave,link}$								✓	✓	✓	✓	✓
$V_{dev,link}$								✓	✓	✓	✓	✓

TABLE V: Prediction Performance for Each Case

Metric	Case1	Case2	Case3	Case4	Case5	Case6	Case7	Case8	Case9	Case10	Case11	Case12
MAE	102.49	81.42	81.19	123.44	79.81	79.51	94.14	94.14	67.26	64.94	71.33	63.06
RMSE	132.51	108.47	108.40	156.55	106.36	106.20	117.49	117.49	88.87	86.29	94.19	83.90
R^2	0.3804	0.5848	0.5853	0.1352	0.6008	0.6020	0.5129	0.5129	0.7213	0.7372	0.6870	0.7516

Table V shows the MAE, RMSE, and R^2 scores for twelve different feature combinations listed in Table IV. It can be seen that the prediction accuracy improved significantly when route-specific features were included. Based on the experimental results of different combinations of features, the final model was trained using the full set of features with 13 traffic-related features and five route-specific features.

The final model achieved an MAE of 63.1 seconds, an RMSE of 83.9 seconds, and an R^2 score of 0.7516 in the test dataset. Considering that the average taxi time for the entire dataset is approximately 401.8 seconds with a standard deviation of 170.1 seconds, the MAE is approximately 15.7% of the average and the RMSE is approximately 20.9%. In addition, the RMSE is approximately 49.3% of the standard deviation, indicating that the model achieved prediction errors smaller than the inherent variability in the data.

To analyze the contribution of the input features to the model's predictions, a SHapley Additive exPlanations (SHAP) analysis was performed. SHAP is a method based on game theory that quantitatively evaluates each input feature's contribution. This allows key features in the prediction process to be visualized and interpreted. As shown in Fig. 12, the total route length (R_{length}), the number of aircraft on the route (N_{route}), and the standard deviation of aircraft speed (V_{dev}) were identified as the most influential factors that determine the taxi out time.

This indicates that both the physical characteristics of the route itself and the ground traffic conditions play a critical role in taxi out time prediction. In addition, the total aircraft on ground (N_{tot}) and the time of the day (T_{hour}) also made significant contributions to the model's decision process.

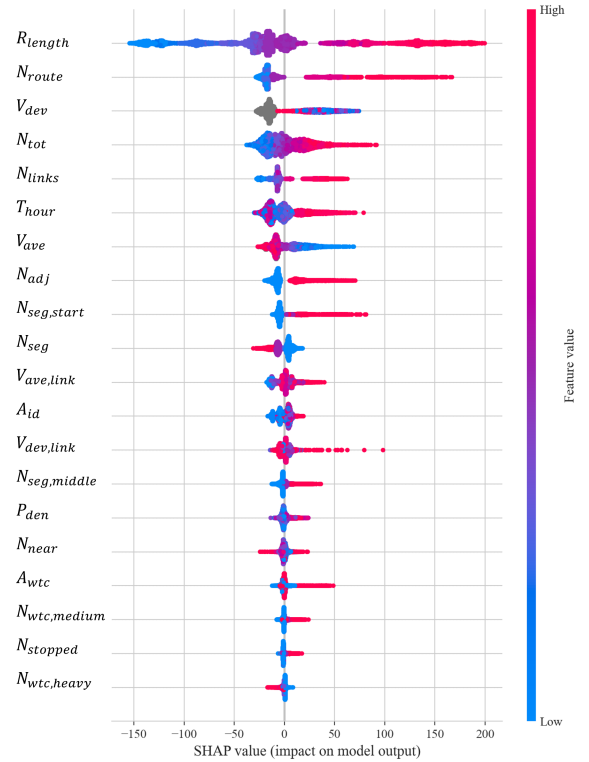


Fig. 12: Feature importance analysis using SHAP values

B. Actual vs. Predicted Taxi Time Comparison

To evaluate the predictive performance of the model, the relationship between actual and predicted taxi times was analyzed. Figure 13 shows a scatter plot of the actual versus predicted values for the entire test dataset. Most of the data

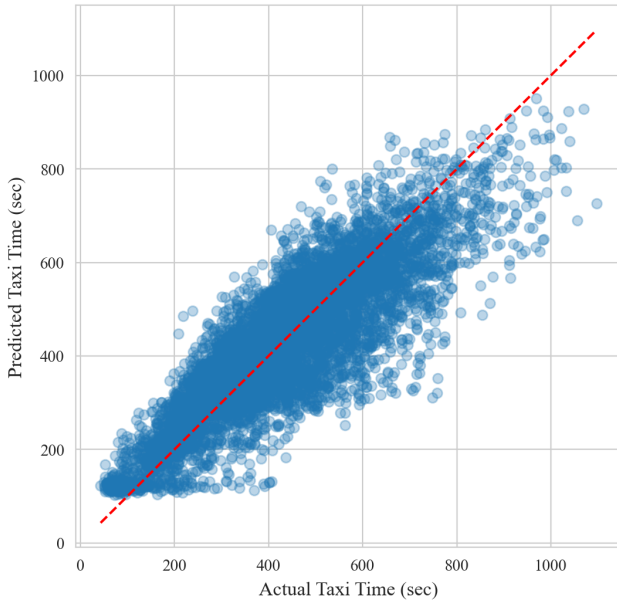
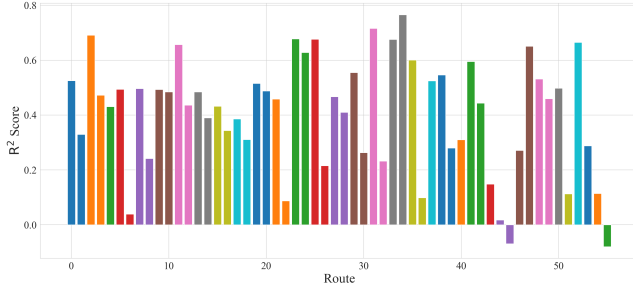


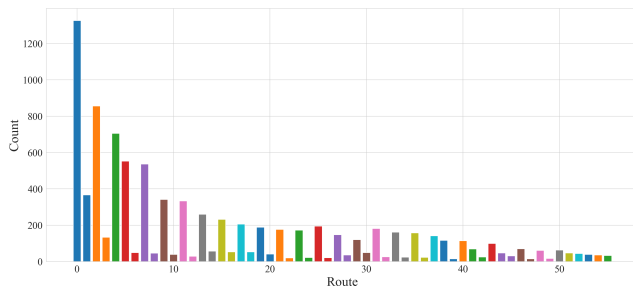
Fig. 13: Actual vs. predicted taxi time

points are distributed near the diagonal line, indicating that the proposed XGBoost model effectively learned the patterns of the actual taxi times.

C. Analysis by Taxi Route



(a) R^2 scores for all 56 routes



(b) Number of data for all 56 routes

Fig. 14: R^2 scores and number of data by route

As shown in Fig. 13, most of the predicted values closely follow the actual values. However, when analyzing the prediction performance on a route-by-route basis, variations in

model performance were observed. As shown in Fig. 14, each bar represents a route, and consecutive bars of the same color indicate alternate routes that share the same initial/final node pair. While most routes achieved favorable R^2 scores, some had values close to zero or even negative, indicating that the model's prediction for those routes was unstable.

Figure 15 shows the prediction errors for all 56 routes. It can be observed that for most routes, the maximum prediction errors did not exceed 50%, and the interquartile range (25%-75%) remained within $\pm 20\%$, indicating that the model maintained an overall stable prediction performance.

Figure 16 shows the route with the highest R^2 score. Although both Figs. 16a and 16b show no unusual patterns in the route geometry, the route in Fig. 16a is a departure on a frequently used runway, resulting in a sufficient amount of training data. In contrast, Fig. 16b corresponds to a route that crosses another runway to depart at a runway that is primarily used for arrivals, resulting in a limited amount of training data. Moreover, since this route involves a runway crossing, it is assumed that a combination of different operational factors contributed to the larger error.

In particular, it was observed that this tendency became more pronounced for routes with relatively fewer data samples. Figure 14b shows the number of data samples for each route, where colors represent identical initial/final nodes combinations. The results indicate that routes with fewer data points tend to have lower prediction performance. For example, routes with comparatively sufficient training data consistently maintained R^2 scores around 0.5, while routes with only a few hundred samples often had negative R^2 scores. Nevertheless, even for some low-data routes, it was confirmed that certain path options within the same initial/final pairs achieved relatively higher R^2 scores.

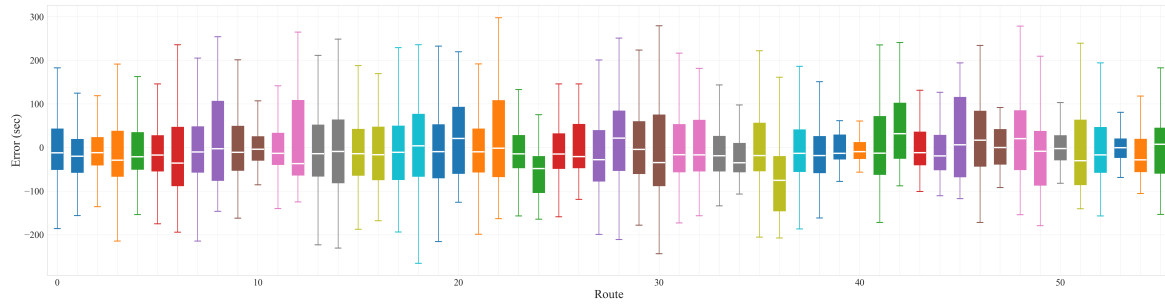
Figure 17 shows the route-wise MAE and the mean absolute percentage error (MAPE). Even though the MAPE remains below 20% in most cases, there were also cases where the relative error was relatively high despite a high R^2 score, which appears to be due to the nature of how the coefficient of determination is calculated.

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} \quad (5)$$

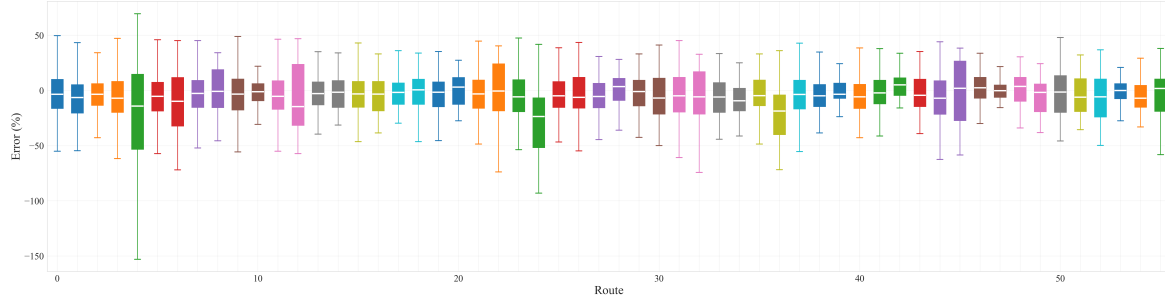
According to (5), the R^2 score evaluates the prediction relative to the mean. Therefore, even if the prediction error of the model is higher than in other routes, a high R^2 score can still be achieved as long as the prediction is closer to the actual value than the mean estimate.

D. Discussion

This study empirically demonstrates that it is possible to rapidly predict an aircraft's taxi out time based on information available at the moment when the aircraft enters the initial node combined with the historical statistics. By considering not only static features such as route id, total distance, and number of links, but also real-time traffic features such as the number of aircraft on the route, adjacent traffic congestion,

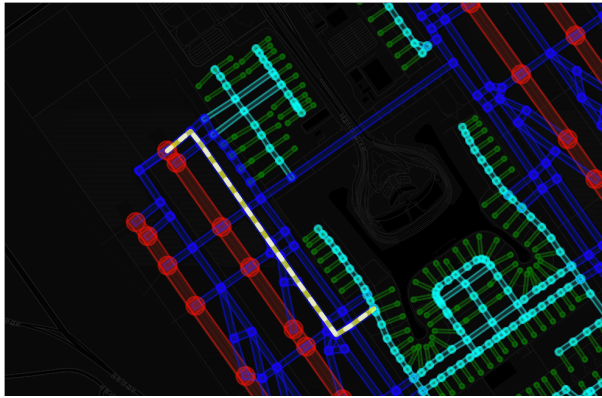


(a) Signed error box plot (sec)

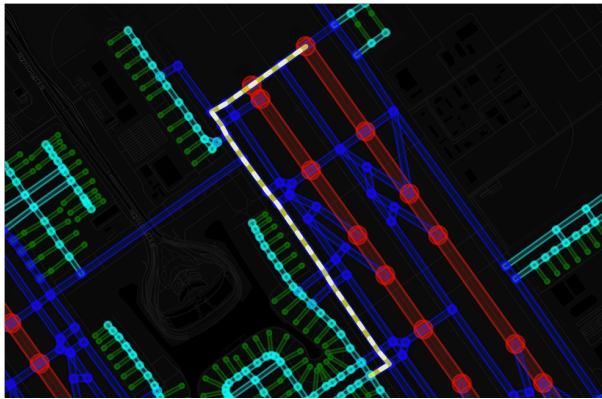


(b) Signed relative error box plot (%)

Fig. 15: Box plot of signed error and signed relative error for all 56 routes

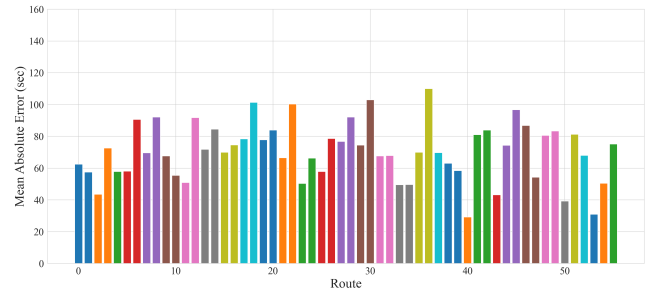


(a) Route with the largest R^2 score

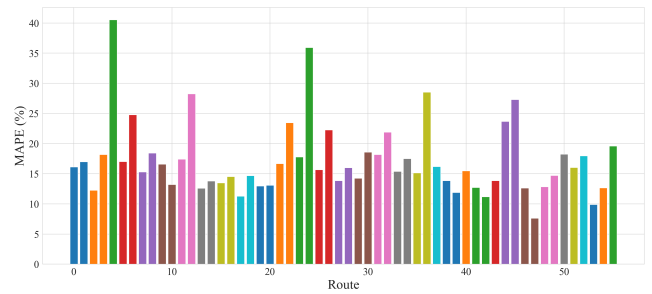


(b) Route with the smallest R^2 score

Fig. 16: Routes with the largest and the smallest R^2 score



(a) MAE



(b) MAPE

Fig. 17: Comparison of MAE and MAPE

and average taxi speed, the sensitivity and adaptability of the model were significantly improved.

The SHAP-based analysis quantitatively interpreted the influence of key features in the model's decision-making process, providing valuable insights for future model improvement and airport management strategy development.

Meanwhile, route-by-route prediction analysis revealed that routes with insufficient data or imbalanced features had lower prediction stability. This result is attributed to a generalization limitation caused by the training imbalance among the routes, suggesting that future studies should consider data augmentation, integrated learning based on route clustering, or transfer learning techniques to address this issue.

In addition, the current study was limited to taxiways where routes are relatively well defined. Extending the predictive model to ramp areas leading to taxiways is expected to allow better scheduling and traffic control from the earliest stages of ground movement.

VI. CONCLUSIONS

In this study, a model capable of predicting aircraft taxi times using surface movement data from ICN was developed. Based on the airport layout information, a node-link model was constructed, and the surface movement data were preprocessed to standardize the aircraft trajectory into a sequence of nodes. In this process, the input features were designed by comprehensively reflecting both the ground traffic conditions observable at the initial time and the static characteristics of the taxi route.

The predictive model applied the XGBoost algorithm, which is based on gradient boosting and uses a variety of traffic-related and route-specific features as input variables. The model achieved a MAE of 63.06 seconds, a RMSE of 83.90 seconds, and a R^2 of 0.7516 on the test dataset. The MAE is approximately 15.7% of the average taxi time, suggesting the potential of the model for practical use in airport operations.

In addition, by applying SHAP analysis, key features contributing to model predictions were identified, enabling quantitative identification of factors that determine the taxi times. This not only improves prediction accuracy, but also provides actionable insights for airport management strategies and ground traffic optimization.

Future research is expected to focus on improving the prediction performance for routes with limited data, extending the model to different airport environments, and further enhancing the model by incorporating a wider range of features.

ACKNOWLEDGMENT

The authors would like to thank Seok-Hwan Lee for processing the raw ASDE data to be used for the analysis.

REFERENCES

- [1] K. Chen, A. Marzuoli, and H. Balakrishnan, "Understanding congestion in airport surface operations using 3D fundamental diagrams," SSRN, Preprint, Jan. 2025. [Online]. Available: <https://ssrn.com/abstract=5119235>
- [2] EUROCONTROL, "Airport Collaborative Decision-Making (A-CDM) Implementation Manual", 2017. [Online]. Available: <https://www.eurocontrol.int/publication/airport-collaborative-decision-making-cdm-implementation-manual>
- [3] M. Jeong, Y. Eun, and D. Jeon, "Taxi-out time prediction using departure queue length estimation," in *Proc. Korean Inst. Aeronaut. Space Sci. Conf.*, Nov. 2016.
- [4] M. S. Jeong, S. Hong, Y. Eun, and D. Jeon, "Aircraft taxi time prediction based on node-link model," in *Proc. Korean Soc. Aeronaut. Space Sci. Conf. Abstract Book*, Jeju, Korea, Nov. 2016.
- [5] M. S. Jeong, Y. Eun, H. Kim, and D. Jeon, "A study on taxi route extraction based on a node-link model for aircraft movements on airport surface," *J. Korean Soc. Aviat. Aeronaut.*, vol. 25, no. 3, pp. 51–60, 2017.
- [6] S. Kim, D. Jeon, and Y. Eun, "A development of data-driven aircraft taxi time prediction algorithm," *J. Korean Soc. Aviat. Aeronaut.*, vol. 26, no. 2, pp. 39–46, 2018.
- [7] H. Cho, S. Oh, and S. Han, "Analysis of variable taxi time at Gimpo International Airport using machine learning," in *Proc. Korean Soc. Aeronaut. Space Sci. Conf. Abstract Book*, Gangwon, Korea, Nov. 2023, pp. 569–570.
- [8] Z. Zhao, S. Feng, M. Song, L. Hu, and S. Lu, "Prediction method of aircraft dynamic taxi time based on XGBoost," *Adv. Aeronaut. Sci. Eng.*, vol. 13, no. 1, pp. 76–85, 2022.
- [9] L. Yuan, J. Liu, and H. Chen, "Additional taxi-out time prediction for flights at busy airports by combining nutcracker optimization algorithm and XGBoost," *Appl. Sci.*, vol. 14, no. 21, p. 9968, 2024.
- [10] D. M. McNulty, A. Koros, and A. Poston, "Airport surface detection equipment—Model X early user involvement event," FAA Tech. Rep. ACT-530-2001-2a, Atlantic City Int'l Airport, NJ, 2001. [Online]. Available: <http://www.tc.faa.gov/its/worldpac/techrpt/asde-x-v1.pdf>
- [11] International Civil Aviation Organization (ICAO), "Specimen AIP", 2006. [Online]. Available: <https://www.icao.int/airnavigation/information-management/Documents/Specimen%20AIP.pdf>
- [12] S. Lee, B. Park, H. T. Lee, "Augmenting ADS-B Data near Airports Using ASDE Data," in *Proc. Korean Institute of Navigation and Port Research Conf.*, Seoul, Korea, Oct. 2023.
- [13] Ministry of Land, Infrastructure, and Transport, "e-AIP," 2022. [Online]. Available: <https://aim.koca.go.kr/eaipPub/Pakage/history-enGB.html>
- [14] M. Park, S. Lee, H. Hwang, S. M. Han, and H. T. Lee, "Analysis of taxi time distribution between runway and gate using surface movement data of Gimhae International Airport," in *Proc. Korean Soc. Aeronaut. Space Sci. Conf. Abstract Book*, Gangwon, Korea, Nov. 2023.
- [15] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, San Francisco, CA, USA, Aug. 2016, pp. 785–794.