Taxi-speed Prediction by Spatio-Temporal Graph-based Trajectory Representation and Its Applications

Thanh-Nam Tran*, Duc-Thinh Pham†, Sameer Alam‡, Vu Duong§
Air Traffic Management Research Institute, School of Mechanical and Aerospace Engineering
Nanyang Technological University, Singapore
Email: {*thanhnam.tran | †dtpham | ‡sameeralam | §vu.duong}@ntu.edu.sg

***Abstract*—Airport surface movement systems require aircraft taxing speed as a key input to perform ground movement optimization and path planning processes. With the increasing availability of surface movement data from systems such as A-SMGCS, a data-driven framework using a spatio-temporal graph-based trajectory representation is proposed in this paper to predict aircraft taxing speed. The proposed framework includes a data preparation module for converting track points data to graph-based representation and a developing predictive model module for learning taxi-speed model. The Random Forest algorithm is selected as our predictive model. The model predicts the aircraft taxi-speed with an error of ± 1.08 m/s for taxi-out procedure and ± 0.97 m/s for taxi-in procedure, when compared with the actual taxi-speed from A-SMGCS data, respectively. Further, three applications of our approach are discussed which are taxi-speed profile, unimpeded taxi time and potential conflict detection. The results of our methods outperform all baseline methods. In detail, for generating taxi-speed profile, our method obtains the error ± 1.38 m/s while for computing unimpeded taxi time, our method outperforms the baseline model with the mean absolute percentage error is 11.03% for the taxi-in and 16.8% for taxi-out procedure, respectively.**

***Keywords*—airport surface movement, map-matching, airport graph, data analytics, A-SMGCS Data, unimpeded taxi-time prediction, speed profile, conflict detection, trajectory prediction**

I. Introduction

Airports are fast becoming a bottleneck in an air transportation system [1]. One of the most critical problems faced by airports around the world is surface movement congestion [2]. Thus, there are many decision support systems and optimization tools that are developed, such as Surface Management System (SMS) [3], Surface Operations Simulator and Scheduler (SOSS) [4], etc. to address this issue. Along with that, the research topics related to airport operations, such as airport congestion [5], airport performance assessment [6] have attracted much attention recently. All such systems and research studies require accurate estimates of the taxi-speed of the aircraft. For example, the simulators require the aircraft speed to simulate the taxiing scenarios and the surface optimization and planning algorithms use the taxi-speed as input requirement.

This research / project is supported by the Civil Aviation Authority of Singapore under the Aviation Transformation Programme.

However, in most systems like SMS and optimization algorithms [7] [8], the taxi-speed of the aircraft is modeled as a constant value. Because the surface operation is a complex logistics system, an assumption that the aircraft always moves with constant taxi-speed is almost impossible to happen in reality. Thus, building a more realistic model for the aircraft taxi-speed estimation/prediction is essential to improve the efficiency and accuracy of the system and for research studies on the surface movement problems.

Some studies considered obtaining a better estimation for aircraft speed instead of using a constant value. For example, [9] models taxi-speed as a function of procedures and turning rate. Then, they use a fixed percentage of initial speed as an uncertainty amount to reduce the speed in some specific cases such as the aircraft follow other aircraft. However, they observed that uncertainty speed reduces the validity of prediction so it only uses in limited conditions. [10] model the aircraft motion using the Kinematics model and requires to manually define different maximum speed for different airport areas (push back, active movement areas, etc.) and aircraft type (e.g., B737, A380, etc. [4]. Nowadays, the available data from the surface movement surveillance system, such as Advanced Surface Movement Guidance and Control System (A-SMGCS) and Airport Surface Detection Equipment, Model X (ASDE-X), the data-driven approach to model the aircraft speed is a promising direction. [11] proposed the method of analyzing ASDE-X data to detect two kinds of segments: straight segment and turn segment. Based on historical data of unimpeded flight, they compute the average speed on each kind of segment and use this information to estimate the speed on a new trajectory given the taxi route. However, in prediction, it is required to assign the kind of segment for each taxi segment in the new taxi route manually. Thus, it is challenging to train and test that model with a large dataset.

In this study, we proposed a data-driven framework by using a spatio-temporal graph-based trajectory representation for developing the unimpeded taxi-speed predictive model. The proposed framework can extract historical data to build the dataset for training and testing automatically by using spatio-temporal graph-based trajectory representation. The unimpeded taxi-speed model, implemented using Random Forest algorithm, can predict aircraft speed given any taxi route assignment, which implies that it can also be used for generating taxi-speed profile, computing the unimpeded taxi-time and detecting potential conflict which later are useful as the input for any surface movement optimization and planning system. Moreover, the unimpeded taxi-time, computed by speed model, is general for any position of aircraft at any time while almost taxi-time prediction model fixed to specific setting [12] [13] [14] such as gate to the runway, the wheel-on place to the runway, etc.

II. Proposed framework

Fig. 1 illustrates our proposed framework which includes two main modules: preprocessing data and developing a predictive model for aircraft's speed. In preprocessing step, we focus on standardizing the airport surface surveillance data such as A-SMGCS data to support automatically extracting analytical features from trajectories. Firstly, the impeded trajectories are detected and removed from our dataset since they introduce noises into aircraft speeds. To reduce the noise of tracking data before applying feature engineering, we then propose to match taxiing trajectories into the airport-taxiway network by a Map-Matching method [15]. In this way, we can get the graph representation of the aircraft trajectory as the sequence of taxi-segments. This representation provides not only the spatial-temporal position of the aircraft but also the physical characteristics of traveled taxi-segments.

In developing predictive model step, we use the graph-based taxiing trajectories for feature engineering. Then Random Forest Regression algorithm is used to build the predictive model, named speed model, for unimpeded taxi-speed. Given the information about the current taxi-segment and the target next taxi-segment of any aircraft, the speed model is able to return the predicted speed of that aircraft in the next segment. This model can also be used to predict the unimpeded taxiing speed profile for any aircraft given the assigned taxi-route, which is discussed in more detail in Section VI-A.

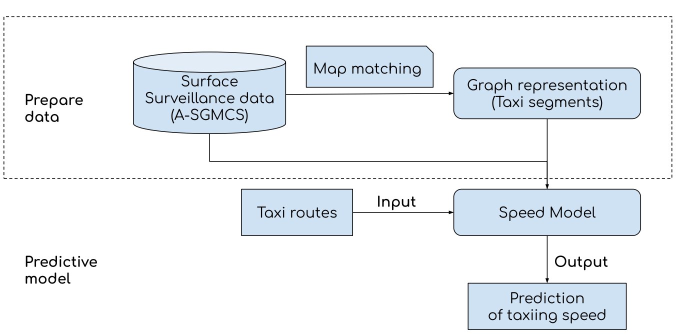

Fig. 1: The process flow of the proposed framework including two main modules: data preprocessing and predictive model

III. Data preparation

A. Dataset

The dataset is used for this study is two months A-SMGCS data in October and November 2017 at Singapore Changi Airport (the Fig. 2 shows the layout of Changi airport). The total flights recorded are 42427, including 21677 departure and 20750 arrival flights.

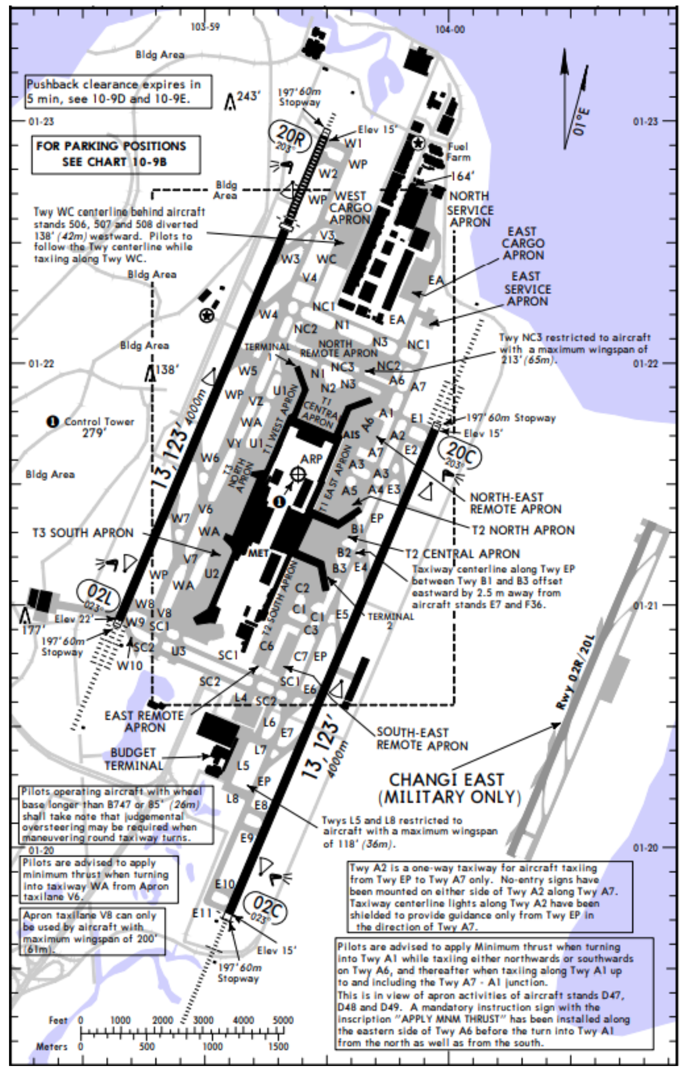

Fig. 2: Changi airport layout [16]

B. Selecting unimpeded trajectories

As per Federal Aviation Administration (FAA), the aircraft unimpeded taxiing movement is considered as the movement under optimal operating conditions (when congestion, weather, or other delay factors are not significant) [17]. In practice, when there are operational issues such as avoiding other surface traffic, the taxiing aircraft often have to stop [11]. So the taxiing movement can be classified into two modes: unimpeded taxi and stopped [5]. From this point of view, in this study, we consider the unimpeded flight as the flight without any stopping point in its movement. We also apply the outlier detection to remove the flights that contain the abnormal speed (outside 1.5 times the interquartile range above the upper quartile and below the lower quartile [18] of speed's value population of each taxi segment).

The aircraft stopping point is detected based on the instantaneous velocity of the aircraft (this information is available in A-SMGCS data). We consider the aircraft is in stopping state if their movement speed bellow 0.5 m/s. To avoid the noise in data, the stopping point only is recorded if the aircraft be in stopping state longer than 10 seconds. Fig. 3 shows the spatial distribution of stopping points in the Changi airport map. The high-density areas of stopping points often are the taxi segments close to gates and the crossing points due to the high possibility of conflict.

Fig. 3: Heatmap distribution of aircraft' stopping points

We build the unimpeded flight dataset by filtering out the flights that contain any stopping point or abnormal taxi speed. The number of unimpeded flights for keeping is 13274 flights for arrival and 1640 flights for departure. The departure flights often are affected by traffic than the arrival flights.

C. Spatio-temporal graph-based trajectory representation

The A-SMGCS data provides the tracking information (position, velocity, etc.) of the aircraft every second. The high tracking frequency of the surveillance system generates a large amount of data, including noise. Instead of applying the smoothing method to reduce the noise of tracking data before input to speed model, we first propose to match the tracking information of the aircraft into the airport-taxiway network by a Map-Matching method.

The applied map-matching [15] includes two steps, which are called spatial map-matching and temporal map-matching. The Fig. 4 shows the spatial result of map-matching. The tracking positions of the aircraft are matched to the physical taxi segments on the airport graph network. Then we obtain the graph representation of aircraft movement as a sequence of taxi segments. The noise in spatial tracking data is also eliminated with the new representation. In temporal map-matching, the graph representation of trajectory is enriched by information about the time and the moving speed of the aircraft. Each taxi segment has a start time and an end time corresponding to the time aircraft enter and leave this segment. As a result, the aircraft movement speed on each taxi segment is estimated and recorded. Fig. 5 shows the speed profile of the aircraft by tracking points and by applying temporal matching.

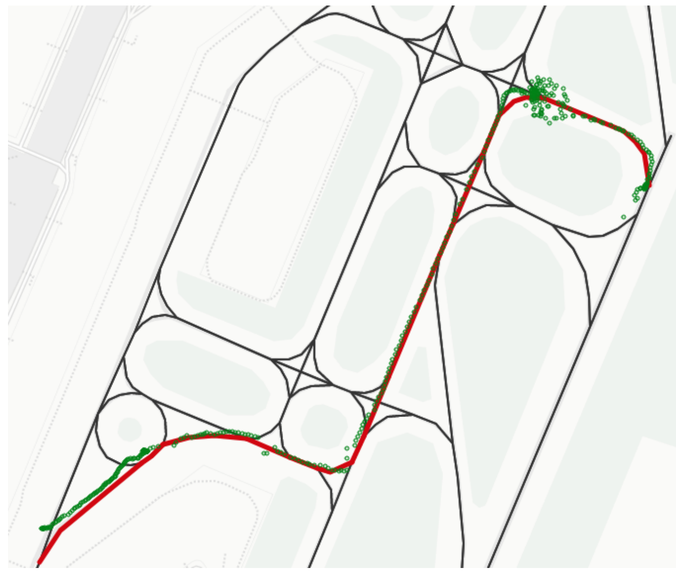

Fig. 4: Illustration of the spatial map-matching where the algorithm converts the green trajectory points of a departure flight into the red path on the graph.

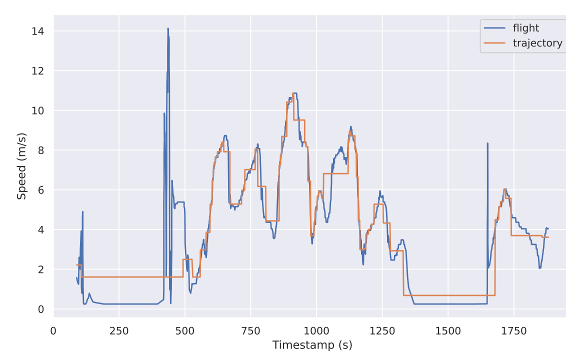

Fig. 5: Illustration of the temporal map-matching where the algorithm estimates the average speed for each taxi-segment (orange line) from the real flight trajectory points (blue line). Two segments with low speeds are corresponding to stops at the gate and holding position.

TABLE I: The list of features

11 features
angle_with_next_segment, prev_median_speed, next segment length, prev_speed, hour, prev segment length, angle_with_prev_segment, segment length, aircraft_type_M, aircraft_type_H, median_speed

IV. Data-Driven Taxi-Speed Model

A. Feature engineering

In the first step of building the data-driven model, we should define the list of features that have a strong relation to the aircraft taxi-speed. More important, such features must be extracted automatically from historical data and the airport graph to make the model easy to scale up and adapt to any airport. The first kind of feature is related to the characteristics of taxi segments. Instead of defining the fixed classes of taxi segments such as straight taxiways, turn taxiways [11], etc. we reflect the characteristics of taxiway by quantitative features as the distance of the taxi segment, the angle of the segment to the previous and the next taxi segment in the trajectory. These features are extracted using the graph representation of the trajectory after the map-matching step. The second kind of feature is the statistical information extract from historical data that includes the median speed of the current taxi segment, previous taxi segment, and next taxi segment. It reflects the field information about the taxi-operations of the airport in the historical data. One important feature is the information about the speed of the previous taxi segment. It provides the current status of the aircraft on the trajectory. Although we have an assumption that the aircraft speed we want to predict is unimpeded, the speed can depend on several factors related to an aircraft, such as the pilot/airline procedures, the type of aircraft-engine (APU) used in taxi operations, fuel loading, etc. The previous taxi segment speed can help to reference such information, especially in the real-time prediction setting. We also consider the type of aircraft and hour of the flight as the feature. The final list of features is reported in TABLE I. Each data point in the processed dataset corresponds to each taxi segment and is independent of the other data points.

B. Predictive model

In this study, we apply the Random Forest Regression algorithm to build the speed prediction model due to its stability and robustness [19]. The Random Forest model is implemented by Sklearn package [20]. The model predicts the speed (m/s) for any taxi-segment given 11 input features described in TABLE I.

V. Experiments and Results

A. Metrics

Two metrics are considered in our study: Root mean square error (RMSE) and Mean absolute percentage error (MAPE):

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y'_t - y_t)^2} \tag{1}$$

$$\text{MAPE} = \frac{100\%}{n}\sum_{t=1}^{n}\left|\frac{y'_t - y_t}{y_t}\right| \tag{2}$$

where n is the size of evaluated data, and y'_t is the predicted speed for input x_t with the known speed y_t.

B. Speed model: Real-time predictive model for taxi-speed

After filter noise, the dataset contains 35360 taxi-segments for departure flights and 246944 taxi-segments for arrival flights. The dataset is split randomly, with a ratio of 70:30 for training and testing data. As a result, for departures, we have 24752 segments for training and 10608 segments for testing, while for arrivals, we have 172860 segments for training and 74084 to testing.

The training and testing data are used for developing the Random Forest Regression model. The grid-search technique is applied for parameter tuning (number of estimators and max depth of trees). The best parameter set is selected for our model is n_estimators = 500 and max_depth = 10.

To evaluate our speed model, we use the RMSE metric, which can reflect our model's performance directly in terms of m/s unit. The RMSE of the speed model on the test set is 1.08 for departures, and 0.97 for arrivals that means the speed prediction has corresponding errors is ± 1.08 m/s and ± 0.97 m/s compared with the actual speed. It is worth noting that the standard deviation of actual speed on the test set is 2.96 m/s for departures and 2.91 m/s for arrivals. Our model can reduce the uncertainty of up to 60%. Fig. 6 shows the distribution between the prediction speed and actual speed in case of arrival flights. We observe that those two distributions are similar in both shape and magnitude.

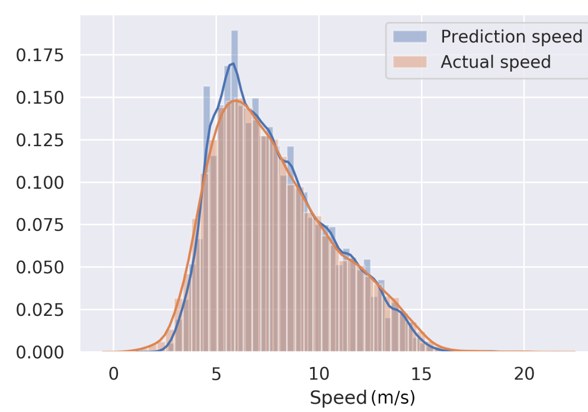

Fig. 6: Distributions of actual and predicted speed for arrival flights

Fig. 7 shows the list of features and their Gini-importance scores provided by the Random Forest model. Gini-importance is one of the metrics used to evaluate the importance of features in the context of ensembles of randomized trees [21]. The feature $previous\_speed$ is the most important feature due to the current state information it provides for the speed model. The second important feature is $median\_speed$ of the given taxi segment in historical data. The characteristics of the taxi segments such as angle with the next segment, the previous segment length, the angle with the previous segment contribute a small effect to the model.

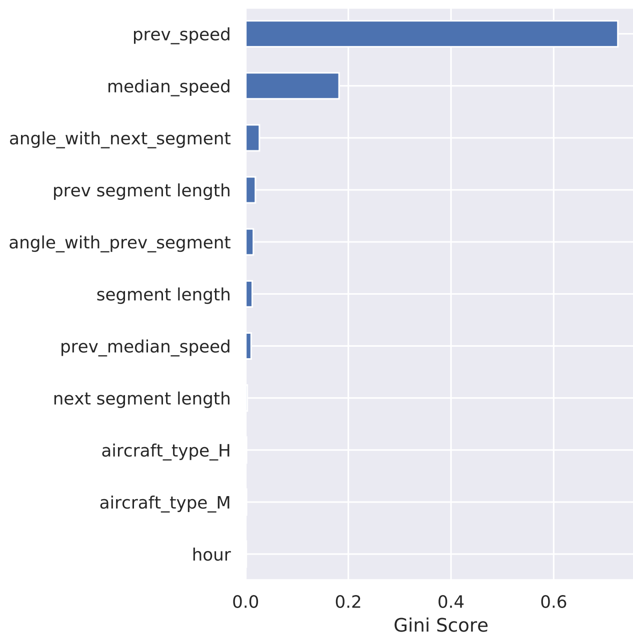

Fig. 7: The ranked list of features and their feature importance scores in speed model.

VI. Discussion on applications of speed model

A. Predicting taxi-speed profile

In many applications, it is useful in predicting the taxi speed for the far future or the whole taxi-speed profile. Giving the list of taxi-segments and initial speed, and a speed predictive model, the process can generate the whole speed profile for both taxi-in and taxi-out. In detail, we predict the taxi-speed of aircraft in the $segment_i$ then use it as the input to predict the taxi-speed of the $segment_{i+1}$ and so on. The generating process is illustrated in Fig. 8.

Fig. 9 shows the taxi-speed profile of an arrival flight from runway 02l/20R to gate G4. By giving the taxi-route and the initial speed of aircraft before entering runway exit W6, our method can predict the whole speed profile for the taxiing. The curve of predicted speed has a similar shape with the curve of actual speed. Our method can predict the decreasing trend when the aircraft is close to the turning points, such as turning points between taxiway WP and taxiway SC1 or between taxiway SC1 and taxiway L6. In the long taxiway such as WP or SC1, our method also successfully predicts the increasing trend of the aircraft speed.

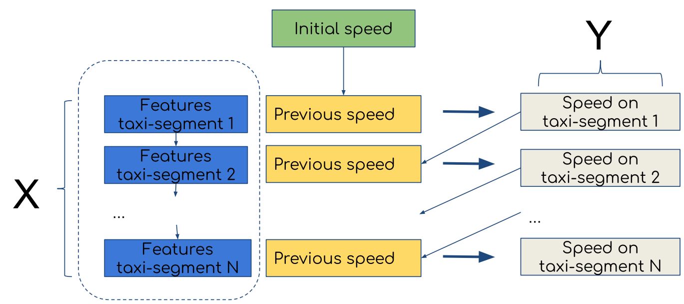

Fig. 8: The process of generating speed profile. The predicted speed on $step_i$ will be the input for predictive model of $step_{i+1}$.

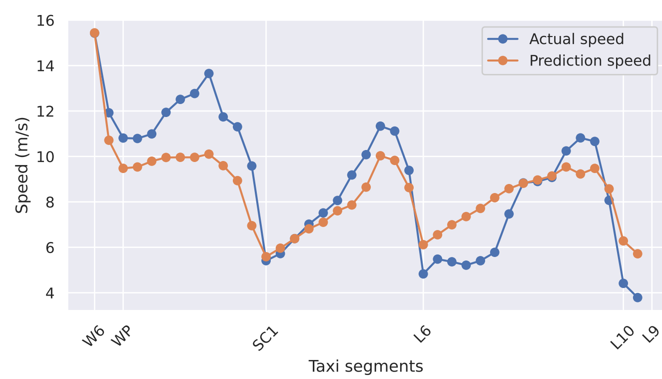

Fig. 9: Taxiing speed profile of arrival flight from runway 02L/20R to gate G4

This approach for predicting taxi-speed profile can face the problem with accumulation error in which the predicted error in the previous step can affect the predicted result of the following step. However, from our observation of Fig. 9, the speed of aircraft around turning points are slow and quite consistent, which are good positions to absorb the accumulation errors and improve the performance of our model for this approach.

The baseline for comparison with our method is the statistical method where the median speed for each taxi-segment from historical data is selected to generate the speed profile. This baseline is enabled and reliable since in our framework, all historical trajectories have been mapped to taxi-segments and all average speeds have been computed. The violin plots in Fig. 10 illustrates the predicted results of our method for speed profile generation in both MAPE and RMSE. The two methods show similarity in the shape of the distribution. But our method outperforms the baseline method in both metrics.

On average, our model achieves RMSE = 1.38, MAPE = 0.175 for generating the whole speed profile. As we have discussed, due to accumulation error, the RMSE of our method is increased compared to the result when predicting only the speed for one segment (RMSE = 1.08). However, the results of our method are only $\approx 77\%$ comparing to the baseline method for both MAPE and RMSE. In addition, we also observe outliers with significant errors in the results. More analysis will be performed in the future to understand this issue, which may then improve the performance of our model.

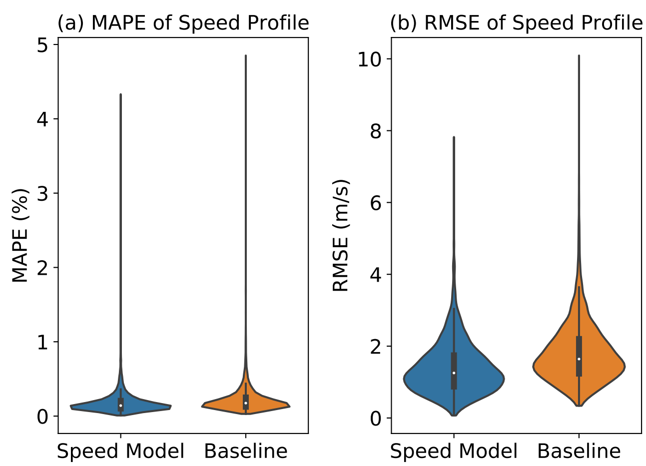

Fig. 10: The distributions of errors for speed profile prediction for our predictive model and prediction using statistical median speed for each segment. The figures (a) and (b) show the results in terms of MAPE and RMSE, respectively.

B. Unimpeded taxi-time prediction

The direct application of the unimpeded taxi speed model is computing the unimpeded taxi time of any flight plan. Using the method for generating speed profile as discussed in Section VI-A, we are able to compute the unimpeded taxi-in time and taxi-out time.

In this experiment, we compare our proposed method with the two baseline methods, which are the median method as $baseline\_1$ (mentioned in Section VI-A) and the 20th percentile method as $baseline\_2$. The 20th percentile method is widely used in the literature to estimate the unimpeded taxi time [22]. This method groups the flights with the same properties and get the 20th percentile value of the actual taxi-time as the unimpeded taxi time. In this study, we group the flights by its gate and runway. The MAPE metric is used for evaluating the performance of methods for predicting unimpeded taxi-time. The TABLE II shows the MAPE of three methods for arrival and departure flights. In both cases, the proposed method outperforms two baseline methods with the MAPE for arrival flights is 11.03%, and departure flight is 16.8%.

TABLE II: Mean absolute percentage error (MAPE)

	Arrival	Departure
20th Percentile Model	31.46 %	56.01 %
Median Model	14.05 %	20.49 %
Proposed Model	11.03 %	16.8 %

C. Potential conflict detection on the airport's airside

Aircraft speed is the factor in detecting the possible conflict among aircraft on the airside during their taxiing phase. When the controller assigns a specific taxi route for an aircraft, a 4D trajectory prediction must be performed to detect any potential conflict on the airside between this aircraft and other active aircraft. By early detect the possible conflict, we can decide the better time to push back and reduce the number of times aircraft need to stop. It will help to save cost, taxi-time, and aircraft's fuel burn.

Since our method can generate a speed profile with high accuracy for any taxi-route, it is useful for predicting potential conflict. In this section, we discuss one case study on applying our method for conflict detection.

Fig. 11 shows the graph representation of two departure flights. One flight move from gate A4 to the runway 02C/20C at 17:06:49 (blue path), another flight move from gate D41 to runway 02C/20C at 17:04:16 (yellow path). By the spatial of taxi route plans, two aircraft can meet at the crossing between taxiway SC2 and taxiway EP. However, the conflict only happens if two aircraft go to the crossing area at the same time.

Fig. 12 shows the actual tracking data of those two flights in which the red circle represents the position where the aircraft of blue trajectory stops. In this situation, the controller ordered the aircraft on a blue trajectory stop to resolve the conflict with the other one.

Using the proposed speed model, we can estimate the future position of two aircraft and the distance between them. The conflict detection with speed model can raise the alarm when the distance of two aircraft below a specific threshold distance. It should be noted that if we use Euclidean distance to compute the distance between two aircraft, it will cause a confusing alarm when two aircraft is close, but they move on the two parallel taxiway such as SC1 and SC2. Instead, we use the distance of the connection path between two aircraft on the airport graph. Moreover, the path on the graph also reflects the direction of the two aircraft. In that way, the system can detect the case that many aircraft are close together but move in the same direction is the case caused by the controller's traffic flow instead of a case of conflict.

The experiment results are demonstrated in Fig. 13 and Fig. 14. Fig. 13 shows the distance between two aircraft through time. The system detects two aircraft that will be moving very

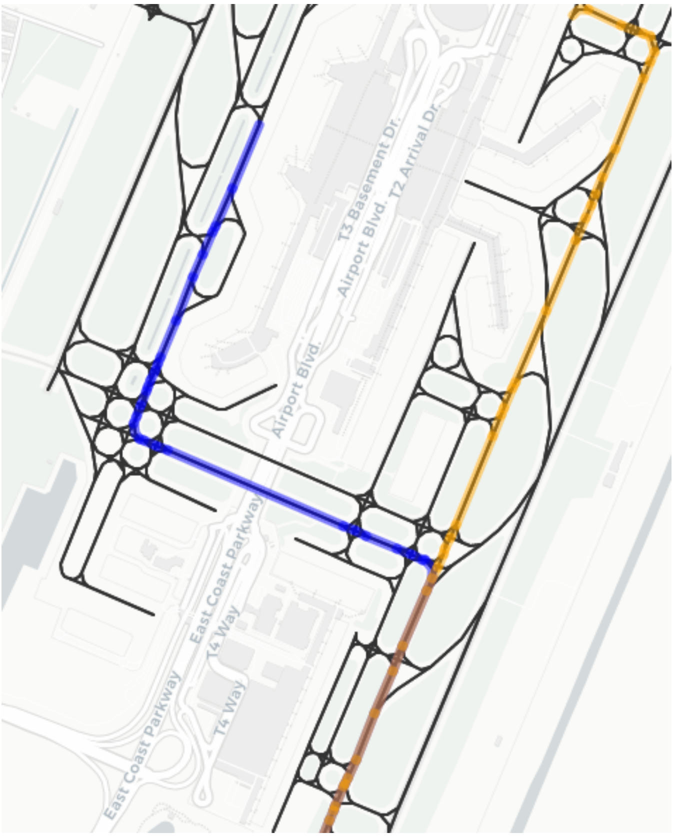

Fig. 11: A case of surface movement conflict of two departure flights. One flight taxis from gate A4 to the runway 02C/20C at 17:06:49 (blue path), another flight taxis from gate D41 to runway 02C/20C at 17:04:16 (yellow path).

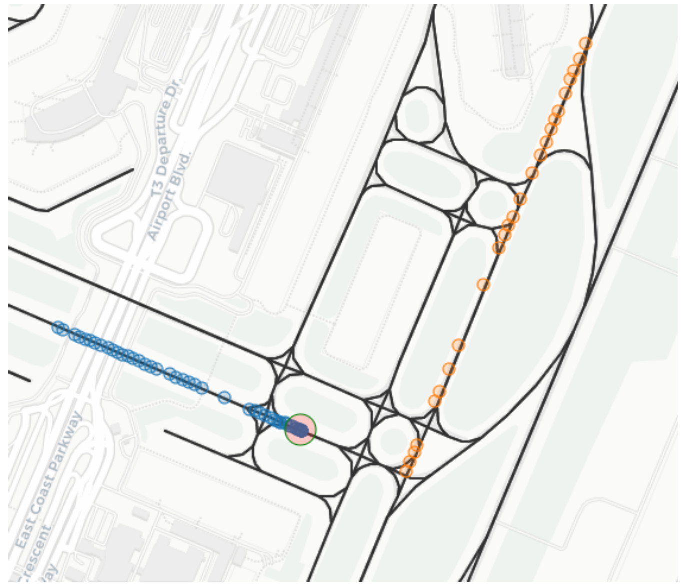

Fig. 12: A case of surface movement conflict in actual data

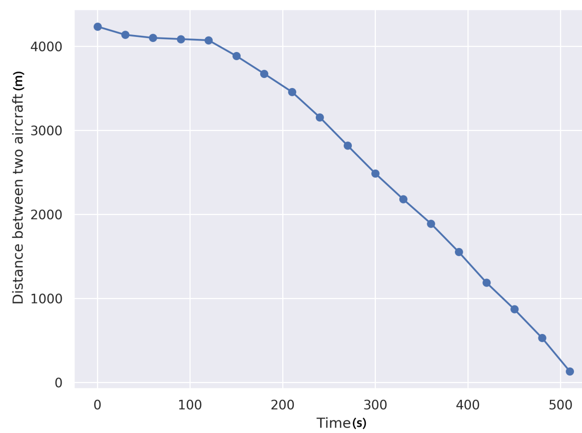

Fig. 13: The predicted distance between two aircraft through time using our method

close to each other near the crossing between taxiway SC1 and SC2 as shown in Fig. 14.

VII. Conclusion

Most of surface movement systems and studies assume the aircraft movement speed as the constant value that is an unreliable assumption. In this study, we proposed a data-driven approach by using a spatio-temporal graph-based trajectory representation for developing the taxi-speed predictive model, which may benefit for improving the performance of surface movement planning and optimization systems.

The proposed speed model predicts the speed with error ± 1.08 m/s for departures and ± 0.97 m/ for arrivals compared to the actual speed on the testing set. We also discuss the application of our model on predicting taxi-speed profiles, unimpeded taxi-time and potential conflict. The performance of our approach is evaluated and compared with baseline methods such as the median and 20th percentile method. For generating taxi-speed profile, the proposed method has the error $\approx 77\%$ of the baseline model ($baseline_1$). For unimpeded time results, our proposed method outperforms two baseline models with MAPE on the testing set is 11.03% for the arrival flights and 16.8% for departure flights. Precise unimpeded time prediction is an example that shows the benefit of a more realistic speed model in ground movement optimization problems. Another application, our method can be used to detect conflicts of aircraft during taxiing by providing better 4D trajectory prediction given the assigned taxi-route.

In some cases, stopping in taxiing is required. For example, in the airport has crossing runways, the aircraft has to stop before the active runway until it is allowed for clearance. In the future, the model will extend to accept those required stopping points in unimpeded taxiing. Furthermore, more features will be investigated to improve the predictability, including visibility, time pressure, the familiarity of the pilot to the airport layout, etc.

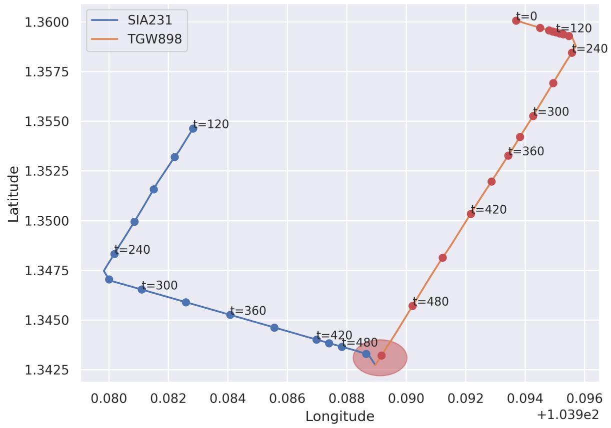

Fig. 14: The illustration of predicted positions of two aircraft through time using our method

VIII. Acknowledgement

This research / project* is supported by the Civil Aviation Authority of Singapore and Nanyang Technological University, Singapore under their collaboration in the Air Traffic Management Research Institute. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not reflect the views of the Civil Aviation Authority of Singapore.

We wish to thank Dr. Man for his insight into the feature engineering process.

References

[1] L. Adacher, M. Flamini, and E. Romano, "Airport ground movement problem: minimization of delay and pollution emission," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 12, pp. 3830–3839, 2018.

[2] R. Morris, C. S. Pasareanu, K. Luckow, W. Malik, H. Ma, T. S. Kumar, and S. Koenig, "Planning, scheduling and monitoring for airport surface operations," in *Workshops at the Thirtieth AAAI Conference on Artificial Intelligence*, 2016.

[3] S. Atkins and C. Brinton, "Concept description and development plan for the surface management system," *Journal of Air Traffic Control*, vol. 44, no. 1, pp. 1–8, 2002.

[4] S. Zelinski and R. Windhorst, "Modelling and simulating airport surface operations with gate conflicts," *The Aeronautical Journal*, vol. 123, no. 1259, pp. 1–19, 2019.

[5] H. Khadilkar and H. Balakrishnan, "Network congestion control of airport surface operations," *Journal of Guidance, Control, and Dynamics*, vol. 37, no. 3, pp. 933–940, 2014.

[6] M. E. Baltazar, T. Rosa, and J. Silva, "Global decision support for airport performance and efficiency assessment," *Journal of Air Transport Management*, vol. 71, pp. 220–242, 2018.

[7] S. Rathinam, J. Montoya, and Y. Jung, "An optimization model for reducing aircraft taxi times at the dallas fort worth international airport," in *26th International Congress of the Aeronautical Sciences (ICAS)*, 2008, pp. 14–19.

[8] H. Balakrishnan and Y. Jung, "A framework for coordinated surface operations planning at dallas-fort worth international airport," in *AIAA Guidance, Navigation and Control Conference and Exhibit*, 2007, p. 6553.

[9] J.-B. Gotteland, N. Durand, J.-M. Alliot, and E. Page, "Aircraft ground traffic optimization," 2001.

[10] J. V. Montoya, R. D. Windhorst, S. Stroiney, K. Griffin, A. Saraf, Z. Zhu, and S. Gridnev, "Analysis of airport surface schedulers using fast-time simulation," in *2013 Aviation Technology, Integration, and Operations Conference*, 2013, p. 4275.

[11] C. Gong, "Kinematic airport surface trajectory model development," in *9th AIAA Aviation Technology, Integration, and Operations Conference (ATIO) and Aircraft Noise and Emissions Reduction Symposium (ANERS)*, 2009, p. 7076.

[12] H. Lee, W. Malik, and Y. C. Jung, "Taxi-out time prediction for departures at charlotte airport using machine learning techniques," in *16th AIAA Aviation Technology, Integration, and Operations Conference*, 2016, p. 3910.

[13] P. Balakrishna, R. Ganesan, and L. Sherry, "Accuracy of reinforcement learning algorithms for predicting aircraft taxi-out times: A case-study of tampa bay departures," *Transportation Research Part C: Emerging Technologies*, vol. 18, no. 6, pp. 950–962, 2010.

[14] S. Ravizza, J. Chen, J. A. Atkin, P. Stewart, and E. K. Burke, "Aircraft taxi time prediction: comparisons and insights," *Applied Soft Computing*, vol. 14, pp. 397–406, 2014.

[15] T. Nam, P. Duc-Thinh, and A. Sameer, "A map-matching algorithm for ground movement trajectory representation using a-smgcs data," in *International Conference on Artificial Intelligence and Data Analytics in Air Transportation (AIDA-AT) 2020*, 2020.

[16] JEPPESEN. (2008) Airport information: Wsss (changi). [Online]. Available: http://www.fly-sea.com/charts/WSSS.pdf

[17] A. Federal Aviation Administration. (2019) Aspm taxi times: Definitions of variables. [Online]. Available: https://aspmhelp.faa.gov/index.php/ASPM_Taxi_Times:_Definitions_of_Variables

[18] Y. Dovoedo and S. Chakraborti, "Boxplot-based outlier detection for the location-scale family," *Communications in statistics-simulation and computation*, vol. 44, no. 6, pp. 1492–1513, 2015.

[19] M.-H. Roy and D. Larocque, "Robustness of random forests for regression," *Journal of Nonparametric Statistics*, vol. 24, no. 4, pp. 993–1006, 2012.

[20] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine Learning in Python ," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[21] G. Louppe, L. Wehenkel, A. Sutera, and P. Geurts, "Understanding variable importances in forests of randomized trees," in *Advances in neural information processing systems*, 2013, pp. 431–439.

[22] Y. Zhang and Q. Wang, "Methods for determining unimpeded aircraft taxiing time and evaluating airport taxiing performance," *Chinese Journal of Aeronautics*, vol. 30, no. 2, pp. 523–537, 2017.