1. Introduction

With the rapidly growth of air transport industry, the major airports in worldwide are getting more and more congested. According to the prediction in European Commission Flightpath 2050 report, the number of global air traffic passengers may increase from 2.5 billion in 2011 to 16 billion in 2050 [1]. Airports are very important hubs for air traffic network and remain a major bottleneck for the development of air transportation. To better handle the challenge between increasing air traffic demand and the limited airport capacity, it is of great significance to development more advanced and intelligent airport management techniques.

Aircraft in airports spend most of their time moving on the taxiway. In order to realize efficient and smooth airport ground traffic operation, precise aircraft taxi time prediction plays a vital role, which refers to the moving time period from the aircraft gate to the runway. Based on accurate taxi time, it is helpful to assist the airport controllers in estimating the airport operational status and establish better operation plans. Inputting accurate taxi time to the airport collaborative decision making (A-CDM) system, airport controllers could increase the utilization of airport gates, taxiway and runways, save the time of passengers, relieve congestion and also reduce the workload of themself[2].

Due to the increasingly importance of taxi time in advanced airport decision systems, taxi time prediction have attracted a large number of researches in the past decades and different prediction methods have been suggested, such as queuing model methods[3,4], regression methods[5,6], machine learning methods[7-10] and some other prediction methods[11,12]. Among these prediction methods, machine learning-based prediction methods are relatively popular and have demonstrated certain superiority in many researches, benefiting from the rapidly development of machine learning and big data technologies.

However, current research works mainly focus on developing new taxi time prediction approaches, the importance of feature selection receives relatively little research attention and only a few related works were published in recent years. As a data preprocessing technique, feature selection has shown to be very effective in improving the performance of machine learning algorithms in many fields [13,14,15]. Airport can be seen as a very complicated system, the taxi time of an aircraft might be affected by many factors, such as peak hours, weather, taxi path and the state of other aircraft, et al. Thus, it is necessary to identify the key factors affecting the taxi process of aircraft to support obtaining more accurate taxi time prediction results.

Motivate by the above idea and in order to enhance the taxi time prediction efficiency, this paper combines a popular feature selection strategy named RReliefF [16] with several machine learning algorithms to develop new taxi time prediction methods for airport. Different from most existing works that ignore the feature selection process or only simply pick some important features based on personal experience, the proposed methods first identify important features for taxi time prediction using the RReliefF feature selection method based on real airport data. After determining the most affecting features, machine learning algorithms are then employed to predict the taxi time based on these selected features rather than all features, which is helpful to reduce computational burden and has the potential to improve the prediction accuracy. In order to verify the effectiveness of the proposed methods, real flight data from an international airport are collected for taxi time prediction. Experiments showed that the proposed taxi time prediction methods could achieve similar or better performance than prediction methods without using feature selection.

The rest of this paper is organized as follows: Section 2 reviews the related works on taxi time prediction and introduces the existing prediction methods. Section 3describes the proposed taxi time prediction methods, including the feature selection strategy and the adopted machine learning algorithms. Then in Section 4, experiments are conducted to test and analyse the performance of the proposed taxi time prediction methods. Finally, conclusion and future work are given in Section 5.

2、Related works

The early researches on taxi time prediction can be traced back over 20 years ago, and queuing models are generally employed. For example, Idris et al.[3] studied the taxi out process of aircraft in Boston Logan International Airport and analyzed the main influence factors of taxi time. Then queuing model was built to simulate the taxi process and achieve more accurate taxi time prediction result. Simaiakis and Pyrgiotis [] divided the taxi out time into three parts, including the unimpeded taxi-out time, the time spent in the departure queue and the congestion delay due to ramp and taxiway interactions. After analyzing the dependence of taxi out time on many features, they developed approximation models of the taxiway and runway queues based on analytical queuing model to predict the taxi out time.

Regression methods were also widely used in taxi time prediction due to their simplicity and effectiveness. Jordan et al. [6]introduced the key characteristics for taxi time prediction of Dallas/Fort Worth Airport and attempted to approximate aircraft taxi time utilizing statistical regression models. They first suggested a simple linear regression model for taxi out processes based only on taxi distance, and further developed a relatively complex regression model by considering more influence factors. Test results on real data showed that their methods are relatively simple and could also obtain appealing prediction results for both taxi in time and taxi out time. Lordan et al.[] proposed a taxi time prediction method based on log-linear regression model for Barcelona-El Prat airport, and found out that their method is able to achieve precious taxi time prediction result if a large training sample covering sufficient airport operation data could be available. They also discussed the importance of different types of airport factors in the taxi time prediction process and gave some useful suggestions.

Machine learning algorithms are perhaps the most popular adopted methods for taxi time prediction, and a great deal of related works have been successfully proposed for different airports. By employing a stochastic dynamic programming framework, Balakrishna et al. [7] suggested a nonparametric reinforcement learning based taxi time prediction method. Experimental studies utilizing real data of Tampa International Airport indicated that the taxi time prediction result generated by their method could match the actual average taxi time well. Besides, the advantage of reinforcement learning based prediction method over traditional regression based method was also discussed. With the purpose of enhancing the prediction ability of common machine learning algorithms, Lian [8] et al. introduced two swarm intelligence optimization improved SVR (SVR) approaches for taxi time prediction, in which a particle swarm optimization and a firefly algorithm was used to improve the performance of SVR respectively. Comparison among linear regression, Softmax regression, artificial neural network and the above two improved SVR algorithms verified that the firefly algorithm improved SVR algorithm showed appealing improvement on prediction accuracy. Herrema [9] et al. studied the effectiveness of neural network, regression tree, reinforcement learning, and multilayer perceptron methods in forecasting taxi time of aircraft, and recommended the regression tree model as the most efficient prediction method through a case study in Charles de Gaulle Airport. Diana [10] investigated and compared the taxi time prediction capability of ten different algorithms including ensemble machine learning algorithms, ordinary least-squared algorithms, SVR algorithm, and regularized or penalized regression algorithms. Experiments using actual hourly data from Seattle/Tacoma International Airport showed that very complicated prediction models may perform poorly due to the lack of generalization ability, and more traditional models may also get a promising predictive performance.

In addition to the aforementioned prediction methods, some other techniques were also employed for taxi time prediction. For instance, Chen et al. [11] attempted to realize taxi time prediction utilizing fuzzy rule-based system (FRBS), and they pointed out that FRBS could supply better taxi time estimations than linear regression methods since FRBS possesses well mathematical properties and explanatory ability. Jeong et al. [12] suggested constructing a node-link model for an airport and predicting the unimpeded taxi time by computing the link travel times on the model.

It can be seen from the above works that existing researches mainly concentrate on developing different kinds of taxi time prediction methods and testing their performances. Feature selection, which is also a key process affecting the prediction efficiency [13-15], receives less research attention, and existing works generally pick some features for prediction based on personal experience or some simple data analysis. In recent years, with the development of automatic data acquisition and processing capabilities, more and more airport data could be used for taxi time prediction, the importance of feature selection is gradually increased. Wang et al. [19] provided a relatively complete taxi time prediction feature group with up to 33 features for the first time. Then they integrated feature selection procedures with several taxi time prediction methods to investigate the influence of different features and some important features are identified and discussed. Recently, Park and Kim [20] conducted a very specialized study focusing on the influential factors to taxi time, in which real data from two different airports are utilized to analyze the aircraft taxi process. They indicated that the taxi time could vary differently in different airports and the taxi time of aircraft might be affected by some common factors like wind speed and number of moving flights.

3、Proposed method

Considering the significance of feature selection in the taxi time prediction process, this study combines a famous feature selection technique, i.e. RReliefF [16], with several machine learning algorithms to develop novel and efficient aircraft taxi time prediction methods for airport as well as investigate the effectiveness of feature selection in improving the taxi time prediction performance.

The framework of the proposed method is illustrated in Fig. 1. Firstly, various types of airport data that may have relationship with aircraft taxi time are collected and preprocessed to provide the underlying basis for future feature selection and prediction operation. Secondly, the RReliefF feature selection algorithm is employed to estimate the significance of different features and identify the most influential feature subset for taxi time prediction. Based on selected feature subset, taxi time prediction model could be built and trained by taking advantage of existing machine learning algorithm, then taxi time prediction result is finally generated. Next, the feature selection algorithm RReliefF and the adopted machine learning algorithms will be described in detail.

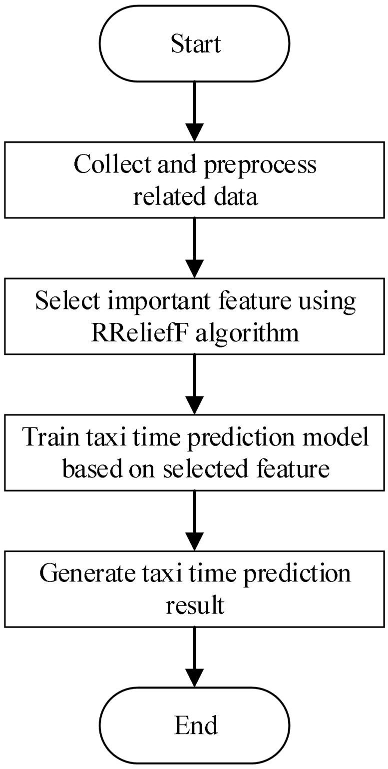

Figure 1. Framework of the proposed method.

3.1 RReliefF feature selection algorithm

RReliefF is an extension of traditional Relief feature selection algorithm for regression problems and has commonly been used to decide optimal feature subset before the regression model is learned [16]. RReliefF is a kind of feature weighting algorithm, its basic idea is to estimate the weight of features according to how well their values distinguish between samples that are close to each other. Features with higher weights are regarded as more important factors for the regression task and will be retained to form a feature subset. In this way, the dimension of data is reduced to speed up the model learning process and the model prediction performance might also be improved.

According to [16], let $W[A]$ denotes the weights of attribute A, $W[A]$ is defined as an approximation of a Bayes rule as follows:

$$W[A] = \frac{P_{diffC|diffA}\, P_{diffA}}{P_{diffC}} - \frac{(1 - P_{diffC|diffA})P_{diffA}}{1 - P_{diffC}}$$

in which P_{diffA} , $P_{diff\,C}$ and $P_{diff\,C|diffA}$ are defined by the following equations:

$$P_{diffA} = P(\text{different value of } A \mid \text{nearest instances})$$

$$P_{diffC} = P(\text{different prediction} \mid \text{nearest instances})$$

$$P_{\text{diff}C|diffA} = P(\text{different prediction} \mid \text{different value of } A \text{ and nearest instances})$$

Based on the computation results of P_{diffA}, $P_{diff\,C}$ and $P_{diff\,C|diffA}$, $W[A]$ could be directly calculated using the probability of the predicted values of two different instances. The detailed computation steps of RReliefF could be found in [16] and will not be repeated here.

The computation result of RReliefF gives the weight of each feature representing its importance. In order to further determine the number of features to form the feature subset for prediction, a feature selection ratio τ ($0<\tau<1$) is introduced. For a total number of n features with their wights obtained by RReliefF, then the top $\lfloor \tau \cdot n \rfloor$ features that have the biggest weights would be selected to constitute the feature subset, which would be finally used for taxi time prediction in the next step. The setting and influence of the feature selection ratio τ will be analyzed and discussed in the experimental study section.

3.2 Prediction models

In addition to feature selection, the actual employed prediction model would also have great impact on the final taxi time prediction performance. To investigate the prediction ability of different models as well as test the efficiency of feature selection procedure more comprehensively, this study adopts three different prediction models for taxi time prediction, including multiple linear regression (MLR), support vector regression (SVR) and artificial neural network (ANN).

(1) multiple linear regression

MLR is one of the most popular model for taxi time prediction and is used as the performance baseline for comparison. MLR is generally used to estimate the linear relationship between two or more explanatory (independent) variables and one response (dependent) variable.

The MLR can be generally formulated as follows:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n + \varepsilon$$

where y is the response variable, x is the explanatory variables, β is the slope coefficient for each explanatory variable and β_0 is generally a constant term, ε denotes the error term of MLR model.

(2) support vector regression

SVR is a generalization and adaptation of support vector machine for regression problems, it is a very popular prediction model with relatively good accuracy and is considered a nonparametric technique because it relies on kernel functions.

Given a set of l training samples $\{(x_1, y_1), (x_2, y_2), \ldots, (x_l, y_l)\}$, in which $x_i (i \in [1, l])$ is a input training data sample and y_i is its target value. The regression function to be derived is:

$$f(x) = \langle w, \phi(x) \rangle + b$$

Where x is an unseen data sample and $f(x)$ is its prediction value, $\phi(x)$ is the kernel feature map of x, and w is the weight in the kernel feature space $\phi(x)$, b denotes the constant offset.

Kernel in SVR model is a function which could map the samples into high dimension feature space to improve computation efficiency. Different kernel functions have different impacts on the prediction ability of SVR in handling different regression problems. Rather than using the traditional linear kernel function, this study suggests utilizing the Gaussian kernel function to exploit the non-linear prediction ability of SVR with the hope of achieving a higher accuracy in taxi time task. The formula of Gaussian kernel function is as follows:

$$K(x_i, x_j) = \exp\left(-\frac{\left\| x_i - x_j \right\|^2}{2\sigma^2}\right), \sigma > 0$$

Where x is a data sample and σ is covariance.

The training method of SVR has been introduced in many studies and thus would be omitted here to save space.

(3) artificial neural network

ANN, also known as neural network, is a kind of machine learning method and is also at the heart of deep learning methods. The name and model structure of ANN are inspired by the way that neurons signal to each other in real human brain.

An ANN is comprised of several node layers, generally including an input layer, one or more hidden layers and an output layer. Each node in a layer connects to node in another layer and has a weight, the connection between two nodes is called edge and the edge weight represents the strength of the connection between two nodes. In the model training process, as the training data transfers from one layer to another ,the ANN continuously adjusts the connections and weights to learn more about the data and improve its regression accuracy over time. There are different kinds of ANNs, the back-propagation neural network (BPNN) is one of the most mature and widely used ANN model and thus is employed in this study. The training process of BPNN could be obtained in many references.

For the convenience of research, we directly invoke the implementations of these prediction modes, i.e. MLR, SVR and ANN, from the library of Matlab, and adopt the recommended default parameter settings to obtain reliable prediction results. In the following content, the proposed RReliefF feature selection-based taxi time prediction methods are named as R-MLR, R-SVR and R-ANN, respectively.

4. Experiments and discussions

4.1 Experimental Setup

For sake of investigating the performance of the proposed three feature selection-based taxi time prediction methods, historical flight data from Lisbon Airport is used for experiments to ensure the reliability of the results. Lisbon Airport is the main international and domestic airport of Portugal and also one of the most congested airports of Europe, which is suitable for taxi time prediction research. After data preprocessing, we extract 5,000 data records concentrated on taxi out flights with 15 features, including flight number, aircraft model, scheduled departure date (month, day, hour, minute), number of other take off aircraft 15 minutes before the current aircraft takes off, number of other landing aircraft 15 minutes before the current aircraft takes off, air temperature, wind direction, wind speed, visibility, cloud coverage, cloud height, day or night. And the respond of each data record is the taxi out time of a flight. In should be noted that the values of these features are all converted to numerical variables and standardized to avoid algorithm failure.

To thoroughly verify the effectiveness of the proposed methods, 10-fold cross-validation is employed in all the experiments, and three different performance measures are used to judge the performance of different prediction methods, including root mean square error (RMSE), mean absolute percentage error (MAPE), and the taxi time prediction accuracy (PA). Specifically, PA is defined as the percentage of prediction accuracy within a specific-error absolute value, i.e. 3 minutes. The computation method of PA is as follows:

$$\mathrm{PA} = \frac{\text{number of } |y-\hat{y}| \leq 3}{N} \times 100\%$$

Where y is the actual taxi time, $\hat{y}$ is the predicted taxi time, N is the total number of samples used. As for RMSE and MAPE, they are very popular performance measures for machine learning methods.

4.2 Parameter analysis

In the feature selection process of the proposed method, RReliefF generates the weights of different features and then a feature selection ratio parameter τ ($0<\tau<1$) is used to decide how many features would be adopted for prediction. To measure the influence of τ on the prediction ability, R-SVR is taken as example and its prediction performance is tested with different values of τ. The value range of τ is set to within [0.2, 0.8], and considering there are totally 15 features in the dateset, the number of selected features is ranging from [3, 12]. Figure 2 presents the results of PA obtained by R-SVR.

It can be seen from Fig.2 that the number of selected features dose have certain influence on the performance of R-SVR, but the influence is not very great since the PA obtained by R-SVR is generally within 60%-62%. When the number of selected features is ranging from 3 to 5, R-SVR could always achieve appealing prediction performance with the PA greater than 62%. However, with the further growth of the number of selected features, R-SVR has experienced a performance degradation.

Furthermore, with more and more features are selected and used for prediction, the performance of R-SVR gets better again gradually. But in this case, RReliefF feature selection procedure almost lose its impact.

Overall, R-SVR performs better when the number of selected features ranges from 3 to 5, which verifies the influence and effectiveness of the feature selection method. To ensure a relatively robust prediction ability and also reduce the computational complexity, the recommended number of selected features is set to 5 in this study and the corresponding feature selection ratio is approximately 0.3.

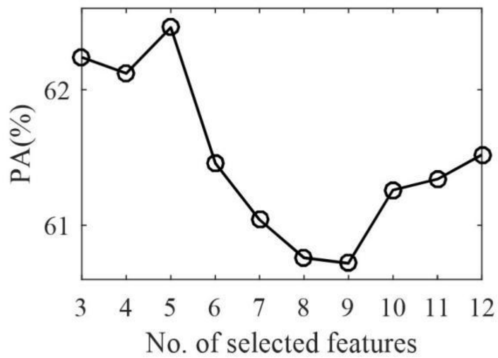

Figure 2. PA obtained by R-SVR with different feature selection pressures.

4.3 Comparisons and discussions

Table 1 Experimental results obtained by six prediction methods

Method	RMSE	MAPE	PA
MLR	3.65	22.18%	61.98%
R-MLR	3.71	22.62%	60.80%
SVR	3.73	21.94%	61.58%
R-SVR	3.75	21.80%	62.46%
ANN	3.67	22.18%	61.80%
R-ANN	3.69	22.50%	61.72%

5 Conclusion and future work