

Detecting Flight Trajectory Anomalies and Predicting Diversions in Freight Transportation (Extended Abstract)

Claudio Di Ciccio,¹ Han van der Aa,² Cristina Cabanillas,² Jan Mendling,²
Johannes Prescher²

Abstract: Timely identifying flight diversions is a crucial aspect of efficient multi-modal transportation. When an airplane diverts, logistics providers must promptly adapt their transportation plans in order to ensure proper delivery despite such an unexpected event. In practice, the different parties in a logistics chain do not exchange real-time information related to flights. This calls for a means to detect diversions that just requires publicly available data, thus being independent of the communication between different parties. The dependence on public data results in a challenge to detect anomalous behavior without knowing the planned flight trajectory. Our work addresses this challenge by introducing a prediction model that just requires information on an airplane's position, velocity, and intended destination. This information is used to distinguish between regular and anomalous behavior. When an airplane displays anomalous behavior for an extended period of time, the model predicts a diversion. A quantitative evaluation shows that this approach is able to detect diverting airplanes with excellent precision and recall even without knowing planned trajectories as required by related research. By utilizing the proposed prediction model, logistics companies gain a significant amount of response time for these cases. The work summarized in this extended abstract has been published in [Di16].

Keywords: Air transportation, Airplane trajectory, Logistics, Machine learning, Prediction methods

1 Introduction

The growth of inter-continental trade has led to a notable increase in multi-modal transport. Multi-modal transport involves at least two modes of transportation on two consecutive transportation legs, which have to be synchronized. This is, for instance, the case when air freight cargo is unloaded at airports in order to be distributed into the hinterland by trucks, or sea ship cargo being redistributed at sea ports. Because multi-modal transport faces increasing challenges in terms of efficiency, describing and planning such sequential dependencies is a common concern. A crucial problem in this context is that different parties involved in a transportation chain hardly exchange real-time information related to individual deliveries. This makes it difficult for a receiving party to respond in a timely way to unexpected events that occur earlier on in the transportation.

The impact of such unexpected events is especially prominent in supply chains that involve cargo airplanes. In case an airplane has to land in an airport that is not the intended destination (i.e. the flight is diverted), re-planning and adaptation mechanisms must be triggered

¹ Institute for Information Business, Vienna University of Economics and Business,
{claudio.di.ciccio, cristina.cabanillas, jan.mendling, johannes.prescher}@wu.ac.at

² Department of Computer Sciences, Vrije Universiteit Amsterdam, Faculty of Sciences,
j.h.vander.aa@vu.nl

so that other parties involved in the chain can continue with the delivery of the cargo. Although diversions are relatively rare, their impact on the logistic chain is significant. To recognize the impact of a diversion on costs and CO₂ emissions, it must be considered that the freight of an airplane is, on average, loaded onto 30 trucks.³ If the airplane diverts to a different destination airport, the logistics service provider has to cancel (or reroute) the trucks that have been sent to the expected destination, and in parallel arrange for new transportation means to pick up the cargo in the actual landing airport. Therefore, this requires a rerouting of up to 60 trucks for a single airplane. Optimization of scheduling around such unexpected events is therefore recognized as one of the most important fleet management decisions. In order for these corrective actions to be effective, it is crucial that the logistics service provider becomes aware of the airplane diversion as soon as possible [Bu13]. Unfortunately, the communication between logistics service providers and cargo airlines is in practice not as prompt as required. In fact, logistics service providers do not receive real-time information and are generally even only notified once an airplane has already landed at another airport. Therefore, it is desirable to identify an anomalous flight behavior without depending on such information completeness.

In this paper, we address the problem of alerting receiving parties, e.g. trucking companies, in case a delivering airplane is diverted. Based on real scenarios, we make use of event data that is semi-publicly available. More specifically, our contribution is a prediction model that detects flight trajectory anomalies based on minimal input data. We implemented the model as a prototype and tested it on a sample of flights yielding a high predictive accuracy. The prediction model provides considerable gains in response time.

To the best of our knowledge, our research work is the first that addresses the issue of predicting the diversions of flights. We also remark here that our approach operates under the requirements that trajectories are not known a priori, and that there is no limited geographical area that is specifically meant to be put under analysis. Previous techniques have challenged related issues in the area of monitoring aircraft routes based on flight data [Kr02, GSF11, Gu14]. Nevertheless, not only they pursue different goals, but also their operating conditions change in terms of information they require: The need of planned flight trajectories as input, the circumscription of the geographical area in scope, or the higher number of factors used to detect anomalous behavior.

2 Automated Flight Diversion Detection

This section describes the proposed prediction model for the automated detection of diverting airplanes. During a flight, an airplane transmits updates on its *(i)* position, *(ii)* velocity, *(iii)* and altitude. We refer to these updates as *flight track events*. Whenever our model receives a flight track event, it predicts whether the airplane is diverting or whether it is still heading towards its intended destination. To make this prediction, the model performs three subsequent steps. (1) Given the receipt of a flight track event, the prediction model combines the received information with the previous gathered events within a given time

³ According to a major logistics service provider that we have collaborated with in this research project.

interval. The length of the time interval L is customizable. Such a combination leads to the generation of five *features*, based upon the framework of [Ca14]: (i) *Distance completed* from the origin airport, (ii) *distance gained* towards the destination, (iii) *velocity deviation*, (iv) *altitude deviation*, and (v) *phase*, namely the covered fraction of the distance from the origin to the destination. (2) The second step uses a one-class classifier to process the treated data. Given the aforementioned features as input, it determines whether the behavior in the time interval should be considered as normal or anomalous. In our implementation, we resort on one-class Support Vector Machines (SVMs) with Gaussian kernel as the classifiers [CV95]. We opted for one-class SVMs because they can be trained on behavior observed in regular flights. It is a valuable advantage in our setting, since the diversions are rare. (3) Finally, we augment the classification of the behavior with the classifications of the airplane's prior intervals. If the number of consecutive anomalous intervals in the flight history surpasses a certain threshold t , our model predicts a diversion.

3 Evaluation

We have evaluated the prediction accuracy and response time gains achieved by using our model on real flight data, acquired from FlightRadar24 and FlightStats for a period ranging from 10/07/2013 to 16/07/2013, and from 14/07/2013 to 11/08/2013, respectively. The data sets contained labeled data both of regular and diverted data, mainly related to routes over the United States and Europe. We have first conducted a grid search in the training and validation phase to tune the best configuration of the parameters (the SVM-specific parameters, the length of the time intervals L , and the number of consecutive anomalies that lead to a diversion alert t). Our final objective was to maximize the F-score, recall and precision of predictions [Mi97], keeping the time needed for a diversion prediction $L \cdot t$ as low as possible. To this end, we adopted a K-fold cross-validation approach with $K = 5$ [Ko95] on approximately the 80% of the flight data. Once the best performing parameters combinations were collected, we proceeded with the test phase on the remaining 20% of data. Gathered test results turned out to be in line with the validation phase, thus showing that the classifiers do not suffer from overfitting with respect to the training data. The minimum and maximum F-score respectively amounted to 0.76 and 0.82. The corresponding values for precision and recall floated in the range 0.78–0.96, and 0.68–0.79. The best classifiers raised a diversion alert after processing consecutive events for 12 to 16 minutes.

To assess the response time gain, we have considered two separate metrics: (i) the time-gain w.r.t. planned arrival time, namely the response time gained to cancel or redirect road transportation assigned to pick up cargo at the original arrival airport, and (ii) the time-gain w.r.t. actual arrival time, i.e., the response time gained to arrange road transportation to pick up cargo at the new arrival airport. Using the configuration that performed best in terms of highest F-score, our approach is on average able to predict a diversion 120 minutes before the originally scheduled landing time, and 62 minutes before the actual landing occurs. This gives logistic service providers more than one hour to react to a probable diversion. This is a significant gain in comparison to the case where logistic service providers have to wait for a notification of the diversion, which often occurs up to two hours *past* the actual landing time.

4 Conclusion

In this paper we tackle synchronization problems in multi-modal transport and the challenge to timely react to unexpected behavior. Our contribution is a model for the prediction of flight diversions based on the automated detection of anomalous behavior. In contrast to prior research, our technique does not require information on planned flight paths. We model the flight trajectory as a sequence of positional updates that describe flight's location, altitude and velocity. Such data is transformed into relevant features that characterize the behavior of an airplane during a time interval, which are processed by a one-class classifier. We evaluated our technique on an extensive set of real-world data demonstrating its accuracy in terms of the F-measure and a substantial time-to-prediction gain.

We plan to extend our work in several ways in the future. Firstly, we intend to expand the approach such that it not only predicts the occurrence of diversions, but also computes to which airport the airplane will most likely divert. Also, knowing how diversions can be predicted for airplanes, we intend to investigate the prediction of breakdowns and diversions in other transportation contexts, e.g. road, inland waterway or railway transportation, such that the model can be used in any multi-modal logistics scenario.

Acknowledgement. The research work has received funding from the European Union's Seventh Framework Programme under grant agreement 318275 (GET Service).

References

- [Bu13] Burgholzer, Wolfgang; Bauer, Gerhard; Posset, Martin; Jammerneegg, Werner: Analysing the impact of disruptions in intermodal transport networks: A micro simulation-based model. *Decision Support Systems*, 54(4):1580–1586, 2013. Rapid Modeling for Sustainability.
- [Ca14] Cabanillas, Cristina; Di Ciccio, Claudio; Mendling, Jan; Baumgrass, Anne: Predictive Task Monitoring for Business Processes. In: *BPM*. Springer, pp. 424–432, 2014.
- [CV95] Cortes, Corinna; Vapnik, Vladimir: Support-Vector Networks. *Machine Learning*, 20(3):273–297, 1995.
- [Di16] Di Ciccio, Claudio; van der Aa, Han; Cabanillas, Cristina; Mendling, Jan; Prescher, Johannes: Detecting flight trajectory anomalies and predicting diversions in freight transportation. *Decision Support Systems*, 88:1 – 17, August 2016.
- [GSF11] Gariel, Maxime; Srivastava, Ashok N.; Feron, Eric: Trajectory Clustering and an Application to Airspace Monitoring. *Trans. Intell. Transport. Sys.*, 12(4):1511–1524, 2011.
- [Gu14] Guo, Yuejun; Xu, Qing; Yang, Yu; Liang, Sheng; Liu, Yu; Sbert, Mateu: Anomaly detection based on trajectory analysis using kernel density estimation and information bottleneck techniques. Technical report, University of Girona, 2014.
- [Ko95] Kohavi, Ron: A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. In: *IJCAI*. pp. 1137–1145, 1995.
- [Kr02] Krozel, Jimmy: Intelligent Tracking of Aircraft in the National Airspace System. In: *American Institute of Aeronautics and Astronautics*. 2002.
- [Mi97] Mitchell, Thomas M.: *Machine Learning*. McGraw Hill series in computer science. McGraw-Hill, Inc., New York, NY, USA, 1 edition, 1997.