ENRI Int. Workshop on ATM/CNS. Tokyo, Japan (EIWAC2010)

[EN-024]Clustering radar tracks to evaluate efficiency indicators

+ R. Winkler * A. Temme * C. Bösel * R. Kruse **

* German Aerospace Center (DLR) Braunschweig
Institute of Flight Guidance
Department of Air Transportation
Brausnchweig, Germany
[roland.winkler | anette.temme | christoph.boesel]@dlr.de

** Otto-von-Guericke University Magdeburg
Institute of Computer Science
Department of Computational Intelligence
Magdeburg, Germany
kruse@iws.cs.uni-magdeburg.de

Abstract It is expected that air traffic increases considerably in the next years whereas the possibility to extend the available infrastructure is limited. Therefore, the infrastructure will have to be used more efficient than today. The airspace surrounding an airport is one of the resources that can not be extended like the number of runways (which is hard enough). On the contrary, in highly populated areas like London, airspace is a valuable resource that is needed to separate approaching aircraft. In this paper, we analyse approaching aircraft within the last 100NM of an airport to compare different approach strategies.

The efficiency of separation strategies for approaching aircraft is measured by calculating performance indicators which consist basically of three values for the last 100NM: the flight time median, flight time variance, and an idealized flight time. The last indicator takes the complexity of the airspace into account, like Mountains, populated areas, or flight restricted areas (i.e. military bases). The ideal flight time represents the time an aircraft typically needs to cross the last 100NM if it is not disturbed by other aircraft.

To obtain useful values of the three presented indicators, approaching aircraft are clustered into groups. In other words, the (unsorted and unstructured) radar tracks must be sorted into groups in which they are comparable: it makes no sense to analyse the flight time for aircraft that approach the airport from different directions. The clustering can be done by two general algorithms which are the main focus of this paper: clustering by origin and runway configuration respectively and clustering by radar track similarity. Both clustering approaches can be further subdivided by using different clustering methods. In this paper, we present the clustering methods as well as compare their results. The analysis is done solely on radar tracks, no information on airport structure, airspace structure, or runway configuration is used. Also the clustering technique is designed to hold minimal human interaction which allows automated analysis of many airports.

Keywords air traffic management, clustering, radar tracks, efficiency indicators

1 INTRODUCTION

Measuring the performance of airports is the basis to increase its efficiency. By observing the efficiency, it is possible to evaluate strategy changes and to decide if they have positive or negative influence on the air-traffic system. This study does not try to improve the efficiency of an airport, nor does it make suggestions how that can be done. While there are many aspects of air traffic efficiency, this paper deals with the efficiency of the air traffic management near an airport for approaching aircraft. The ATM Airport Performance (ATMAP) framework [8] defined a number of key performance indicators (KPI) like the flight time median, flight time variance, and idealized flight time. The median and the variance can be computed quite easily. The idealized flight time however requires a model of the airspace because it represents the time an aircraft would typically require to reach the airport given the airspace structure of the airport. One major influence in the flight time of approaching aircraft is the separation strategy. The efficiency of different separation strategies might be measurable using the above mentioned KPIs.

This study is done with several constraints. The analysis should be designed in such a way, that human interaction is limited, i.e. the evaluation does not need hand picked parameters. New airports might

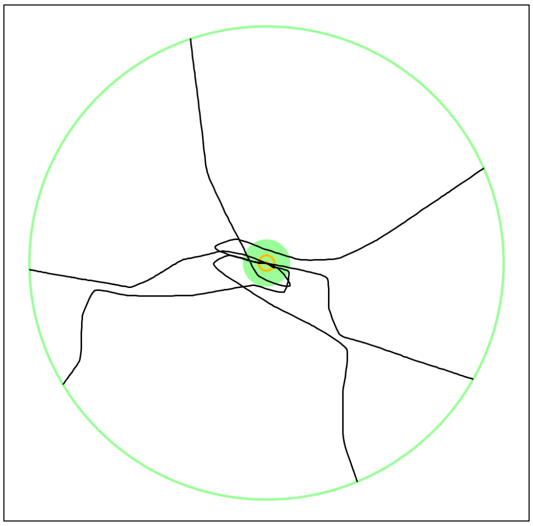

Figure 1 A collection of typical flight tracks.

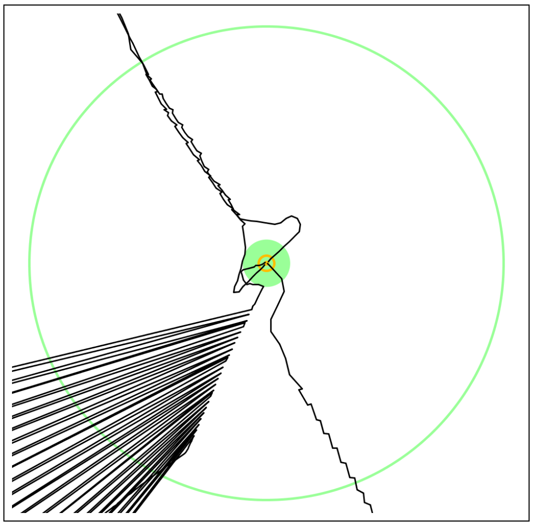

Figure 2 Some flight tracks containing errors.

be added without major work on the models or algorithms. These constraints are important because there are over 100 airports considered with each between 200 and 8000 approaching flight tracks. Therefore, the KPI calculation should work on all of these airports without human interaction due to the enormous workload hand picked parameter would require.

1.1 Near airport airspace: TMA and ASMA

The airspace around an airport, terminal maneuvering area (TMA), is usually organized in a complex, 3-dimensional structure. The structure of the TMA defines the influence area and influence possibilities of the airport. For this analysis, the exact structure of the TMA is not important. Therefore, the airport sequencing and metering area (ASMA) is defined as the cylindrical airspace, surrounding the airport. The cylindrical airspace has the airport reference point (ARP) as centre and an undefined height. Following the suggestion of the PRU [8], the radius of this cylinder is defined to 100 NM because it is estimated that all flight operations related to landing operations such as descending and sequencing are performed in this area. Also it is estimated that the TMA of an airport is included in the ASMA. The analysis in this paper is performed solely on flight track data inside the ASMA and a list of airport locations. To ensure that all flights are comparable, flights starting inside the ASMA are excluded from the analysis.

1.2 Available Information

The data that is available for this analysis consists of radar tracks and the coordinates of airports. The radar data consists of a collection of civil aircraft movements within one week, collected from ground radar and transponder information. Each data entry consists of a collection of elements such as flight track id, time stamp and position. The flight track id is unique for each flight and can be used to identify transits. Therefore, a sequence of radar trajectories with an identical flight track id defines one flight track. In Figure 1 some typical examples for flight tracks are presented. The green outer circle represents the ASMA, the green, filled circle the ILS area and the yellow inner circle represents the approximate area of the airport.

However, there are several errors that can occur. For example if two radar stations have overlapping areas and not perfectly calibrated positioning or not perfectly synchronised time measurements, the flight tracks appear to 'jump' as shown in figure 2. Sometimes, two aircraft fly exactly behind one another from the viewpoint of a radar station with the result that the radar signals can not be identified by the radar station and the recorded route jumps back and forth from one aircraft position to the other. These corrupted flight tracks have to be dealt with in some way during the data analysis phase.

For the analysis, only the flight track parts inside the ASMA are of interest. For this reason, they are cut of by extrapolating the exact space/time location at which the aircraft entered the ASMA. As there are radar records outside the ASMA, there are also radar recordings after the aircraft already has landed. To ensure that the flight track analysis is done solely in the air, the flight tracks are cut of at 6 NM distance to the ARP because it is safe to assume that the aircraft is about to land any moment. Again, this final inner point is extrapolated using the radar recording before and after the 6 NM border. An alternative to this simple model would be to include the Airport layout for

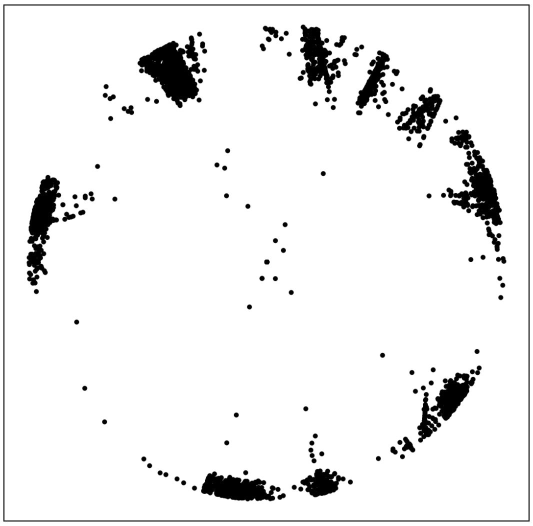

Figure 3 A data set of entrance points of one airport

every analysed airport but this would require an enormous workload with almost no benefit.

1.3 Reasons for clustering

Calculating the flight time median or flight time variance for all approaching flights of one airport at once does not make any sense. The airport layout and weather conditions define the landing direction of approaching aircraft. For example, let the landing direction be exactly west at one period of time. Than aircraft approaching from east might be able to fly a straight line while aircraft approaching from west (and though having an initial approach direction east) would have to surround the airport before being able to land. Therefore, these aircraft have very different distances to cover inside the ASMA and calculating the flight time median and variance for these flights would be biased. For this reason, flights have to be grouped in such a way, that similar flights are in one group. The groups should be large enough so that a statistical value such as the median still contains some meaning, but at the same time, flights of one group should be somehow similar.

For example it makes sense to put flights in different groups that enter the ASMA from very different directions. For the same reason, flights that land on different runways or runway configurations should be put in different groups. The process putting previously unorganised data in groups is called 'clustering'. To cluster the flight tracks, several different approaches are possible which are presented in Section 2.

1.4 Cluster dependent KPI

After the flight tracks are clustered, cluster depending KPIs (C-KPI) can be calculated. The calculation of the flight time median as well as the flight time variance are pretty much straight forward. However, the idealized flight time is more complex. In the ATMAP project [8], the idealized flight time is defined as the time an aircraft requires for passing the ASMA if no other flight interferes, but this approach might not be suited fr all cases. For example, there might not be such a flight track for each cluster that is not interfering with other flight tracks. Furthermore, such a flight can not be regarded to be typical for an airport because a situation of no interference is not typical for major airports. Therefore, a more sophisticated method is presented in Section 3.

Once all C-KPIs are calculated, the final KPIs can be defined. Several ways are possible, using the mean or median of C-KPIs is possible as well as for example always using the worst C-KPI value for the final KPI.

2 CLUSTERING FLIGHT TRACKS

There are many different clustering approaches which all have different qualities. A clustering algorithm is a software tool that uses unorganised data to produce grouped data and the process itself is called clustering. One group of data objects is called cluster. In this paper, we consider several different kinds of clustering algorithms.

Independent of the algorithm, there is a general problem that is called noise. The term noise refers to data objects that seem to not belong to any cluster because they are just too far away from all clusters. Noise is a serious problem in clustering because with noise, it is not clear if a data object actually belongs to a group or if it is noise. For example the odd shaped flight tracks in Figure 2 can be considered noise. Or an aircraft entering the ASMA from a quite unusual direction can also be considered noise as for some entrance locations in Figure 3.

2.1 Clustering ASMA entrance location and landing points

As stated in the last section, flight tracks have to be clustered according to their ASMA entrance and landing point. Since both processes are quite similar, we present only the clustering of the ASMA entrance points. The ASMA entrance point of a flight is defined as the first recorded radar observation of that flight inside the ASMA. In Figure 3, such a data set of ASMA entrance points is presented. In the middle of the picture, the airport is located. The points near the middle are the first, observed locations of an aircraft. Since it is desired to analyse the way the aircraft took from outside the ASMA to the airport in the middle, these flight tracks are considered to be noise and have to be removed from further calculations. Also some

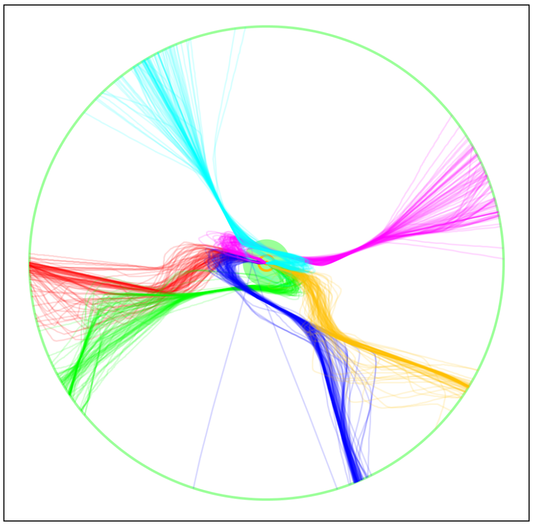

Figure 5 A sample of 6 flight track clusters defined by 3 entrance clusters an 2 landing clusters.

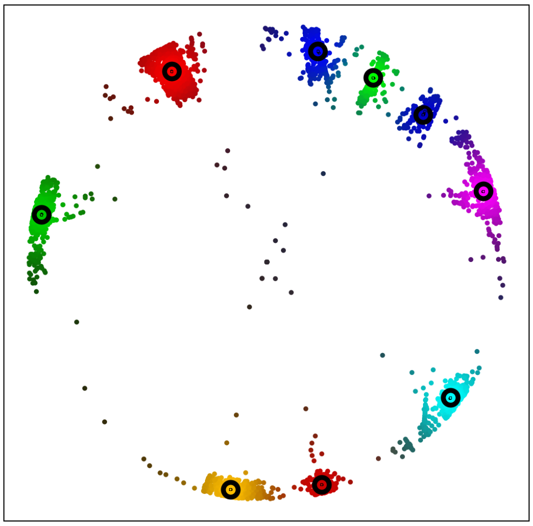

Figure 4 NFCM performed on a data set of ASMA entrance locations of an airport.

spots are very far away from more populated areas in the data set which are also regarded as noise because they do not hold statistically relevant value.

Several different clustering algorithms can be used to cluster such data sets. Here, we present a prototype based clustering algorithm called noise clustering (NFCM) [2, 7, 9] which is a version of the Fuzzy c-Means [4, 1, 6] clustering algorithm that is able to detect noise. However, NFCM needs several input parameter to work properly. NFCM needs the number of clusters in advance, a roughly expected cluster size to identify noise data objects and some other minor parameters. Since this are a lot of parameters that would have to be hand picked (which is not desired), a different clustering algorithm is used to calculate the necessary preconditions for NFCM. We used a density based clustering algorithm called DBScan [5] to identify the necessary parameters. The result of the clustering algorithm is presented in Figure 4. Note that the color of the data objects define their cluster association. Black is considered to be the noise cluster. However, NFCM is a fuzzy clustering algorithm that means, each data object is not assigned solely to one cluster, but gets a grade of membership. This is much more stable than a unique assignment and it provides insight into data objects that can not be uniquely assigned because they are in between two clusters. This also holds for noise data objects and so, the darker a data object is presented, the more it can considered to be noise.

The same process as for ASMA entrance positions is done for landing positions. With this, each flight track is assigned to two cluster identifications that in combination define the approach cluster. In Figure 5, a sample of flight tracks of 6 approach clusters is presented which are formed by 2 landing clusters and 3 entrance clusters.

2.2 Clustering flight tracks by flight track similarity

An alternative to clustering flight tracks using the ASMA entrance point and landing point is clustering flight tracks by flight track similarity. For this, task NFCM can not be used because flight tracks are (mathematically) more complex objects than just points. However, DBScan can be used if there is an appropriate distance or dissimilarity measure for flight tracks. Several distance measures can be used for this task, like measuring the difference of flight tracks by the area in between them or by the supremum of the pointwise minimal distance. In our tests, the first measure is too complex and therefore needs too much time to calculate. The second one yields reasonable results but it is hard to find the correct DBScan parameters that fit a data set. However, if taken the average of point wise minimal distances, the algorithm works quite well. All similarity calculations have the problem that they can not differentiate well between flight tracks that are different at the entrance of the ASMA and those that differ near the airport. Because flight track difference near the airport is more important than at the border of the ASMA, the point wise distances is weighted according to their distance to the airport. With this approach, similar flight tracks are put into the same cluster quite effectively. In Figure 6, some results are presented.

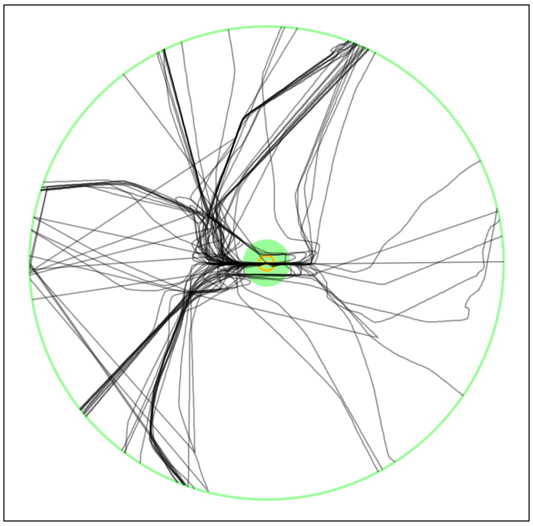

Figure 7 The noise cluster, generated with clustering by entrance location

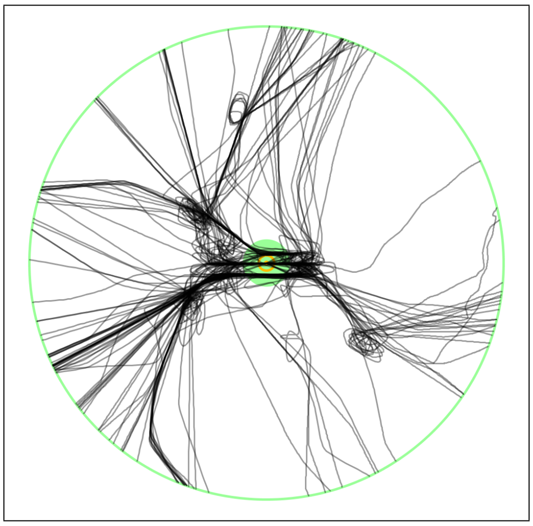

Figure 8 The noise cluster, generated with clustering flight track similarity

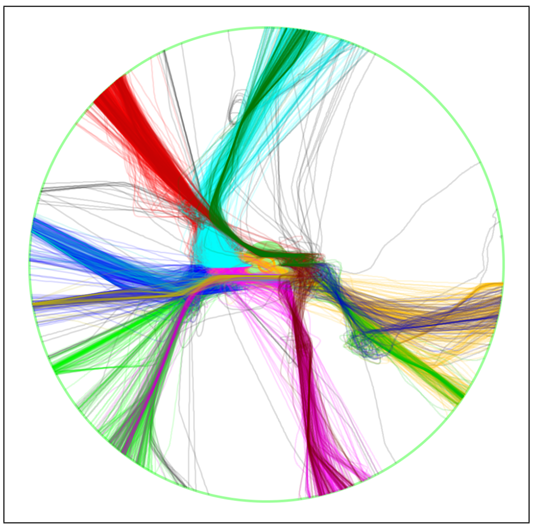

Figure 6 The result of clustering by flight track similarity of one airport

2.3 Comparing clustering methods

Both clustering approaches find clusters in the data that are to some extend similar. However, the second approach is capable of detecting unusual flight patterns which is not possible with the first method. Observe Figure 7 and 8 which shows the flights detected to be untypical by the two discussed methods. In the first approach, that are all flights that do not belong to an approach cluster. In the second example that are all flights that are too different from the others to be part of a cluster. These two noise definitions are quite different. For example, clustering by flight track similarity provides the possibility to find flights that used holdings which is uncommon in the presented example. On the other hand, at the day of operation, it is possible to determine the approach cluster in advance. Therefore, it is possible to use the statistical information provided by this study to determine the time the aircraft might need inside the ASMA. This could increase the effectiveness of flight time prediction.

3 CALCULATING IDEALIZED FLIGHT TIME

An idealized flight time is calculated for each approach cluster separately. The calculation is done by calculating an idealized approach route for each approach cluster and using the average flight time inside the ASMA to calculate the time to cover this approach route. The idealized approach route should be typical for an approach cluster, it should be as short as possible and a civil aircraft should be able to fly this route. This is not a trivial problem, which is why we use an indirect approach. We calculated an overflight frequency map (OFM) of the ASMA for each cluster. An example of an OFM is presented in Figure 9. The darker the colour at one point, the more aircraft flew over that point.

From this OFM, the idealized approach route is generated. This task might be easy for humans, but it is not trivial for a computer. To solve this problem, we used so called ant systems [3]. Ant systems are designed after their biological archetype, who use swarm intelligence to find an optimal route from their layer (the ARP) to some food resource (the ASMA border). The digital version is quite similar. The ant

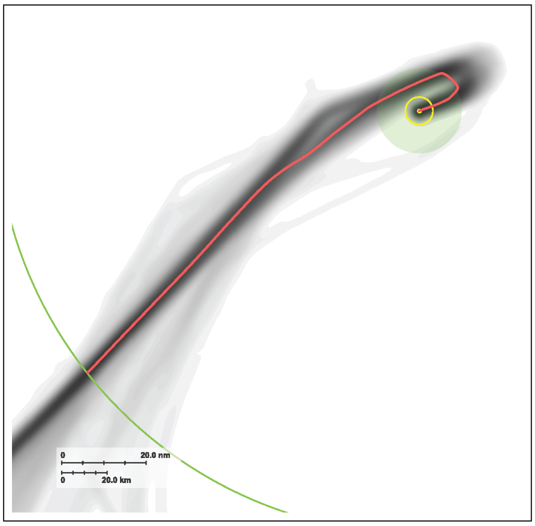

Figure 9 The OFM of one approach cluster with its idealized approach route.

system finds an optimal route under the above mentioned optimization goals (short, typical, and flyable) from the ARP to the ASMA border.

The algorithm generates randomly a set of example routes that are evaluated. The best generated example routes are than used to construct a new generation of example routes. This is iterated until no further optimization occurs. The final best example route is taken as idealized approach route. Observe Figure 9, the idealized approach route is presented as red line inside the approach cluster. Using the OFM for calculating the idealized flight time is very robust for corrupted flight tracks i.e. those presented in Figure 2. Small errors are almost not noticeable and even large errors occur very rarely and do not influence the approach route calculation.

The idealized approach route can be used in simulation studies if an exact airport and airspace model is not required. Here, it is used to calculate the idealized flight time. The idealized flight time was calculated by using the average flight time of the approach cluster. It is also possible to map a distance to airport dependent flight speed profile on the idealized approach route if that is available for the specific situation. It might be possible to construct one from the given data and is topic for future development.

4 CONCLUSIONS

We have presented in this paper two very different possibilities of clustering radar tracks of aircraft approaching an airport using radar tracks. The clusters were used to evaluate KPIs for performance measures of air traffic management inside the ASMA. For the KPI idealized flight time, an idealized approach route was generated that can be used in other studies than performance evaluations. For our study, no complex airport or airspace models are needed, furthermore, only very limited human interaction is needed for the calculation process.

Copyright Statement

References

[1] James C. Bezdek. *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press, New York, 1981.

[2] Rajesh N. Dave. Characterization and detection of noise in clustering. *Pattern Recogn. Lett.*, 12(11):657–664, 1991.

[3] M. Dorigo and Th. Stuetzle. *Ant Colony Optimization*. MIT Press, 2004.

[4] J.C. Dunn. A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters. *Cybernetics and Systems: An International Journal*, 3(3):32–57, 1973.

[5] Martin Ester, Hans-Peter Kriegel, S. Jörg, and Xiaowei Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. In *2nd International Conference on Knowledge Discovery and Data Mining*, pages 226–231. AAAI Press, 1996.

[6] F. Höppner, F. Klawonn, R. Kruse, and T. Runkler. *Fuzzy Cluster Analysis*. John Wiley & Sons, Chichester, England, 1999.

[7] Frank Klawonn. Noise clustering with a fixed fraction of noise. 2004.

[8] PRU. Atm airport performance (atmap) framework - measuring airport airside and nearby airspace performance. Technical report, Eurocontrol, 2009.

[9] Frank Rehm, Frank Klawonn, and Rudolf Kruse. A novel approach to noise clustering for outlier detection. *Soft Comput.*, 11(5):489–494, 2006.