MATEC Web of Conferences **105**, 00005 (2017) DOI: 10.1051/matecconf/201710500005
IWTSCE'16

Multi-objective optimization approach for air traffic flow management

Rabie Fadil[1], *Badr* Abou El Majd[1], *Hicham* Rahil[2], *Hassan* El Ghazi[3], and *Naima* Kaabouch[4]

[1] *LIMSAD Lab, Faculty of Science Ain Chock - Casablanca, Morocco*
[2] *Air Navigation Department, Mohammed V International Airport,Casablanca, Morocco*
[3] *STRS Lab, National Institute of Posts and Telecommunication, Rabat, Morocco*
[4] *Electrical Engineering Department, University of North Dakota, Grand Forks, United States*

Abstract. Air traffic has long been a generally high growth sector and all forecasts indicate that this trend will continue at a similar pace for the next twenty years. The regular traffic demand growth has led to congestion at airports and in space.
In this paper, we will create in first a probabilistic model which describes the uncertainty of the aircraft's trajectory, and its presence in a sector during a time interval. We define as result, a multi-objective optimization problem whose objective functions are the expected cost of delay and the expected cost of congestion.
Then we use the Non-dominated Sorting Genetic Algorithm (NSGA-II) to solve an instance included 21 flights and 1 sector, and is able to provide a good approximation of the Pareto front.
The decision-making stage was then performed with the aid of data clustering techniques to reduce the size of the Pareto-optimal set and obtain a smaller representation of the multi-objective design space, thereby making it easier for the decision-maker to find satisfactory and meaningful trade-offs, and to select a preferred final design solution.

1 INTRODUCTION

The Operational Research community has studied many variants of the air traffic flow management problem since the beginning of the 90s. Along the years, the models have been refined in order to take into account new operational constraints.

One of the first formulations of air traffic flow management is the ground holding problem, which minimizes the sum of airborne and ground delay costs when the demand for the runways exceeds the allowed capacities. The first work under this formulation has been the Single Airport Ground Holding Problem [1]. After that, [2] has proposed a stochastic and dynamic version of the formulation.

Then, the Multi-Airport Ground Holding Problem was addressed by [3]. This does not take into account the sector capacities,rerouting and speed changes.

The first two limitations were overcome with the model of [4]. Also, this work has the merit to use realistic instances with several thousand flights. To our knowledge, the most comprehensive formulation is the Air Traffic Flow Management Rerouting Problem [5] which integrates all phases of a flight, different costs for ground and air delays, rerouting, continued flights and cancellations. Also, with the same mathematical framework, [6], [7] have formulated the problem in terms of routes instead of nodes.In the same manner, there is also the work of [8] which describes an optimization problem to minimize directly the probability of congestion of the sectors with the concept of chance constraint.

Besides, other techniques were used to solve similar problems. [9] uses stochastic optimization methods for handling sector congestion with take-off delays and rerouting. Constraint programming was also used by [10] and [11]. The former solves the slot allocation problem with sector capacity constraints and the former minimizes an air traffic complexity metric for multiple sectors.

More recently, a multi-objective optimization approach has been used in air traffic control by [11] to minimize an aggregated complexity metric, designed and validated by *Eurocontrol* [1], over sectors. Also, [12] uses the multi-objective to model the trade-off between sector congestion and delays.

2 THE MATHEMATICAL MODEL

In this section, we briefly present the mathematical model wich is based on [13] and [14]. It consists of two submodels, the flight model and the sector model. The former is used to compute the expected cost of delay and the probability of presence in the sectors. The latter takes the probability of presence for each flight

[1] European Organisation for the Safety of Air Navigation, it coordinates and plans air traffic control for all of Europe. This involves working with national authorities, air navigation service providers, civil and military airspace users, airports, and other organisations. Its activities involve all gate-to-gate air navigation service operations.

and compute the expected cost of congestion. First, the flight model is defined as follows:

$$\begin{aligned} p_N(t_N) &= \int_\Omega \dots \int_\Omega p_{1:N}(t_{1:N}) dt_{1:N-1} \\ &= \int_\Omega \dots \int_\Omega \prod_{i=2}^{n} p_{i|i-1}(t_i|t_{i-1}).p_1(t_1) dt_1 \dots dt_{N-1} \\ &= \int_\Omega \dots \underbrace{\left[\int_\Omega p_{3|2}(t_3|t_2) \left[\underbrace{\int_\Omega p_{2|1}(t_2|t_1).p_1(t_1) dt_1}_{p_2(t_2)} \right] dt_2 \right]}_{p_3(t_3)} \dots dt_{N-1} \end{aligned}$$

Where Ω is the time line, $p_1(t_1)$ is the marginal density function for the flight f to enter the airspace at time t_1 and $p_{i+1|i}(t_{i+1}|t_i)$ is the conditional density function to be on X_{i+1} at time t_{i+1} given it was on X_i at time t_i. Then the probability to be in sector during the time interval $\Delta t = [t_{min}, t_{max}]$ is :

$$P\left(S_{s,f}^{\Delta t}\right) = F_i\left(t_{max}\right) - F_j\left(t_{min}\right) \tag{1}$$

Where $F_i(t)$ is the cumulative density function denoting the probability that the flight f has flown over X_i by time t.

Thereafter, the probability to have n flights in the sector s during time Δt can be expressed by :

$$\begin{aligned} P\left(K_s^{\Delta t} = n\right) &= \frac{1}{N_s^{\Delta t} + 1} \sum_{l=0}^{N_s^{\Delta t}} exp(-iwln) \,. \\ & \prod_{f=1}^{N_s^{\Delta t}} \left[1 - P\left(S_{s,f}^{\Delta t}\right) + P\left(S_{s,f}^{\Delta t}\right) exp(iwl) \right] \end{aligned} \tag{2}$$

Where $i = \sqrt{-1}$ and $w = \frac{2\pi}{N_s^{\Delta t}+1}$.

Now, we have all the elements to define our multi-objective optimization problem. Because of the stochastic context, one way to define the cost functions is to use their expected values. So the total expected cost of the delays is defined as follows:

$$\begin{aligned} C_1(\gamma) &= \sum_{f \in \mathscr{F}} E_{\phi_f}\left(T_{n_f}^f ; \gamma_{|f}\right) \\ &= \sum_{f \in \mathscr{F}} \left[\int_\Omega (t - A_f)^2 . p_{n_f}^f (t; \gamma_{|f}) dt \right] \end{aligned} \tag{3}$$

Where $\mathscr{F}$ is the set of flights, A_f is the scheduled time of arrival of flight f, n_f the number of waypoints in the flight plan f and p_{n_f} refers to the marginal density function associated to the arrival point $X_{n_f}^f$. $\gamma_{|f}$ denotes the decision vector, the target times of arrival on the different waypoints of the flight f. Indeed, in the optimization context, we are trying to find the vector γ that minimizes the cost functions. For the expected cost of congestion, the function is:

$$C_2(\gamma) = \int_{\overline{\Omega}} \sum_{n=C_s+1}^{N_s^{\Delta t}} (n - C_s)^2 . P(K_s^{\Delta t} = n; \gamma) dt \tag{4}$$

Where C_s is the capacity of the sector s and $N_s^{\Delta t}$ is the maximum number of flights that can be inside the sector at the same time with probability higher than 0 and $\overline{\Omega}$ a temporal horizon sufficiently large to encompass the supports of the marginal distributions.

3 EXPERIMENTS

In this section, we describe the experiment on minimizing both the expected cost of delays and the expected cost of congestion with a multiobjective optimization algorithm and a probabilistic model. Then the results on a simple instance are detailed.

The chosen variation operators are shown in the table below:

NSGA-II operators

Encoding: Real
Population size : 1000
Maximal number of evaluations : 30000
Selection : Binary tournament
Crossover : Simulated Binary Crossover , $P_c = 0.9$
Mutation : Polynomial, $P_m = 1/NumberOfVariables$

3.1 Experimental Setting

The chosen instance for this study, implies 21 flights, including 15 departures and 6 arrivals in the Terminal Manoeuvring Area (TMA) surrounding Mohammed V International Airport.

All the results were computed on an Intel® Core™ i5-4200M CPU with 4× 2.5 GHz and 4Go of memory. The model was coded in Java, for NSGA-II and SMPSO we have used the implementation provided by the jMetal framework. We have also used the Apache Commons Mathematics Library for the probabilistic distributions and the computation of the integrals. This library is freely downloaded at : http://commons.apache.org/

3.2 Analysis

Now, we will analyze each step of the methodology. The first step consists in computing the marginal probabilities over the waypoints, as depicted on figure 1, We can see that the value of the modes decreases and the supports of the distributions increase with time. This simply translates the fact that uncertainty on the target time of arrival increases with time. The next step is the computation of the probability for a flight to be in a sector along the time. This is done with (1) and figure 2 shows the result.

The probability increases when the aircraft approaches the sector and once the flight is inside the sector, the probability should not decrease until the exit.

Thereafter, we need to compute the probability that there will be n flights inside the sector at a given time with (2) and we can remark in figure 3 that the probability of congestion decreases with time since the variance of the marginals increases with time.

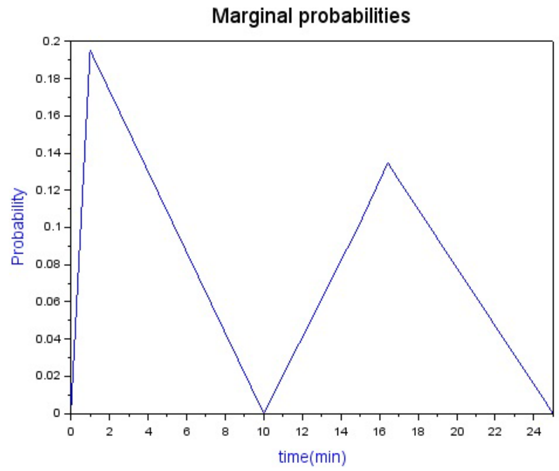

Figure 1: Marginals probability over two waypoints

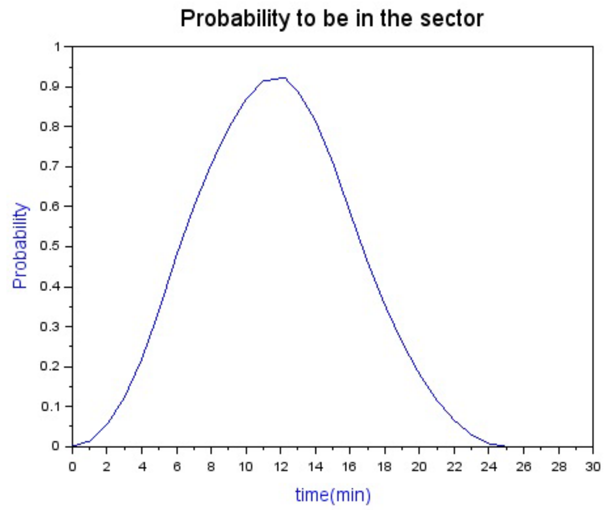

Figure 2: Probability to be in the sector

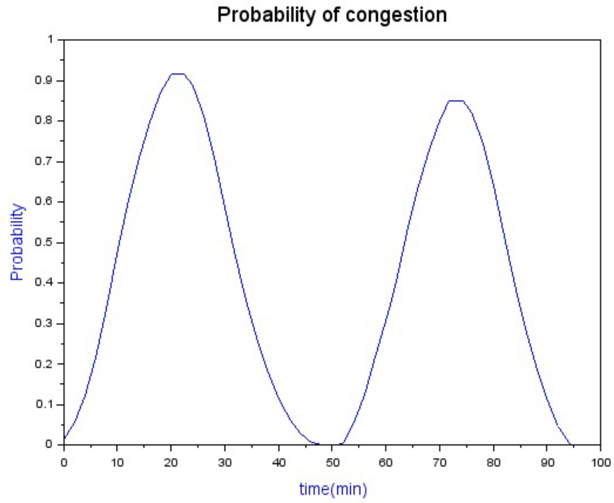

Figure 3: Probability of congestion

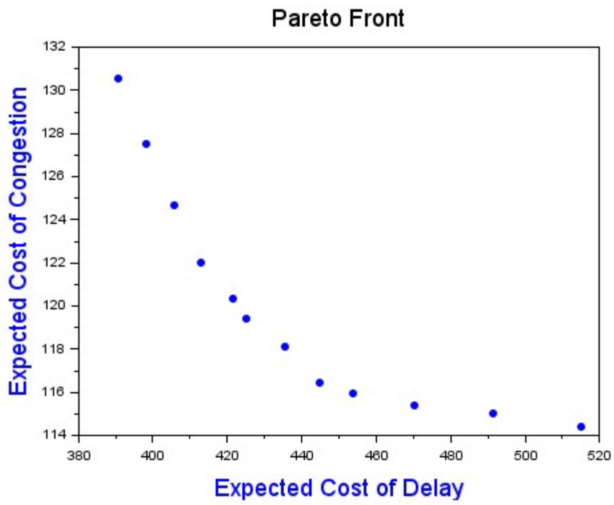

Figure 4: Pareto front obtained with NSGA-II

When all these distributions are known, the probabilistic model can compute the expected cost of delays and expected cost of congestion with (3) and (4) respectively. We can note that minimizing the probability of congestion for every timestamp will effectively minimize the expected cost of congestion.

Finally, when the cost functions are known, one can optimize by given different objectives to the flights. Figure 4 shows the Pareto front for 12 different solutions. Every solution is a complete schedule in the decision space and so, the Pareto front shows the trade-off between minimizing both the delays and the congestion. If the decision maker wants to minimize the chance that delaying the flights congests the sector, he shall select a solution toward the lower right corner of the graph. Otherwise, if she/he believes that the controllers can manage more flights, he can choose a schedule in the top left corner of the graph, resulting in much less delays, at the price of a higher congestion in some sectors.

4 Post-Pareto Analysis

In this section, we attempt to reduce the number of trade-off solutions while considering the stochastic nature of the objective functions.

We will use a standard agglomerative hierarchical clustering technique. We have chosen a hierarchical clustering technique (over a partitioning one, such as k-means) because it allows us to easily adapt the level of granularity of the clusters, and it does not require the number of clusters to be specified in advance.

When carrying out a hierarchical cluster analysis, the process can be represented on a diagram known as a *dendrogram*. This diagram illustrates which clusters have been joined at each stage of the analysis and the distance between clusters at the time of joining.

In our approach, we have generated a dendrogram from the solutions enclosed in the Pareto front as depicted in figure 4. Figure 5 shows the dendrogram obtained using the *Euclidean distance* as a measure of similarity and the *Ward's method* as a hierarchical clustering procedure.

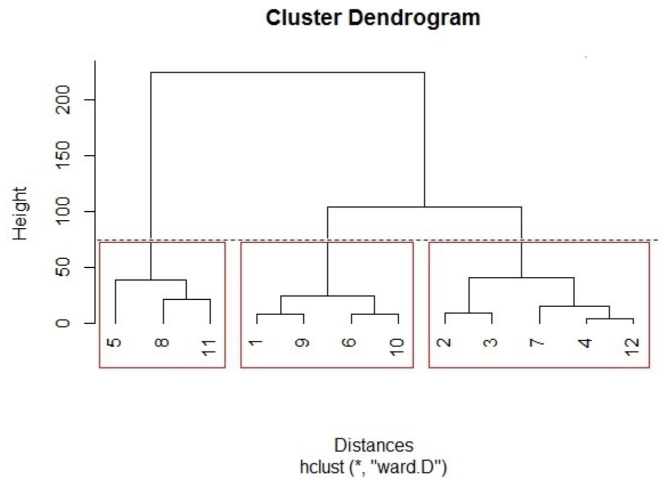

Figure 5: Dendogram

Visual inspection of a dendrogram can give indications about the spread of solutions in the solution. The height of a node represents the distance between its two sub-clusters. Therefore, nodes that are low on the dendrogram form clusters whose elements are closer to each other than nodes that are placed higher up. For example, the cluster 1, 9 forms a cluster whose elements are closer to each other than the cluster 5, 8, 11. The visual inspection of a dendogram can also reveal long branches leading to isolated solutions, which may indicate the presence of "outliers". An outlier in our context is a solution that is radically different from other solutions in the set.

Once the dendrogram has been generated, we need to decide a cut-off at a certain level in order to obtain a set of clusters. For example, with a cut-off value of 60, shown with dotted line on figure 5, we have generated 3 clusters formed by the set of elements under the three branches cut by that line and framed by rectangles.

A recurrent problem that many clustering algorithms encounter is the choice of the number of clusters. Thus, different cluster validity indices have been suggested to address this problem.

The Silhouette [15] and The Elbow method [16] are two of these cluster validity techniques. They are helpful metrics to guide the choice of the best number of clusters.

Figure 6 shows the Scree plot which represents the number of clusters against the total within sum of squares, we can see that the error decreases as the number of clusters gets larger, this is because when the number of clusters increases, they should be smaller, so distortion is also smaller. The idea of the elbow method is to choose the number of clusters at which the SSE decreases abruptly. This produces an "elbow effect" in the graph. In this case, the Elbow method has selected 3 clusters.

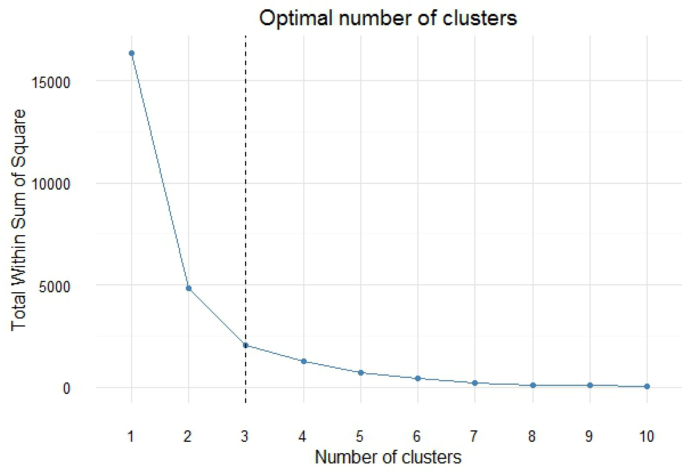

Figure 6: Scree plot

Figure 6 shows the clusters silhouette plot with an average silhouette width of 0.47, which indicates a satisfactory cluster quality. After we have validated the number of clusters, we need to select a representative solution for each cluster. To do this, the solution that is closest to its respective cluster centroid is chosen as a good representative solution. The obtained results of the clustering analysis are depicted in table 1 and figure 8 represents the final solutions.

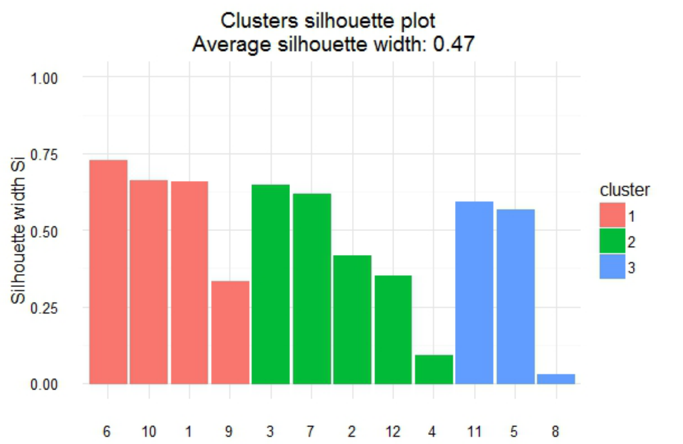

Figure 7: Silhouette plot

5 Conclusion

This paper has presented a probabilistic model to handle the propagation of the uncertainty from the trajectories to the sectors. Then the probabilistic model was used within an optimization algorithm for scheduling all flights at boundaries of the sectors in order to minimize the expected cumulated delays and the expected sector congestion.

Table 1: Final compromise solutions

Cluster	Solutions	Cluster centroid		Representative solutions	
		Expected Cost of Delay	Expected Cost of Congestion	Expected Cost of Delay	Expected Cost of Congestion
cluster1	1,6,9,10	401.97	126.15	398.24	127.48
cluster2	2,3,4,7,12	436.29	118.02	435.51	118.11
cluster3	5,8,11	492.33	114.90	491.5	114.99

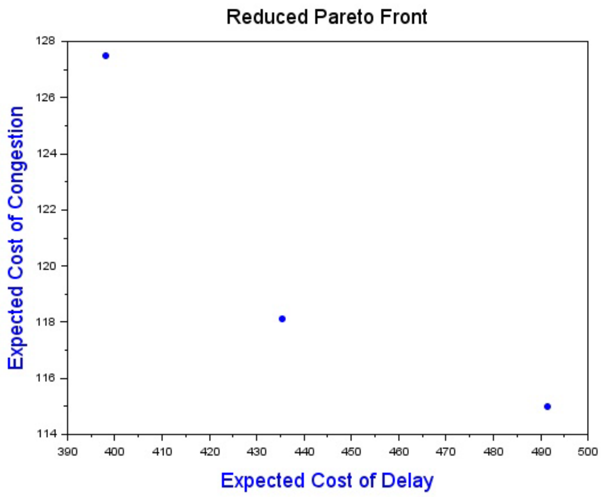

Figure 8: Reduced Pareto Front

NSGA-II was proposed to solve the multi-objective problem, which consists of minimizing both the expected cost of delay and the expected cost of congestion.

Furthermore, in order to illustrate how the theoretical model can be useful in practice, we presented some results on an instance with 21 flights, including 15 departures and 6 arrivals in the Terminal Manoeuvring Area (TMA) surrounding Mohammed V International Airport.

The decision-making stage was then performed with the aid of data clustering techniques to reduce the size of the Pareto-optimal set in order to help the decision-maker to select a preferred final design solution.

References

[1] A. Odoni, " The Management Problem in Air Traffic Control ", *Flow Control of Congested Networks*, pp. 269-288, 1987.

[2] O. Richetta and A. R. Odoni, " Dynamic Solution to the Ground-Holding Problem in Air Traffic Control ", *Transportation Research Part A: Policy and Practice*, vol. 28, no. 3, pp. 167-185, May 1994.

[3] P. Vranas, D. Bertsimas, and A. Odoni, " The Multi-Airport Ground- Holding Problem in Air Traffic Control " ,*Operations Research*, vol. 42, no. 2, pp. 249-261, 1994.

[4] D. Bertsimas and S. S. Patterson, " The Air Traffic Flow Management Problem with Enroute Capacities ", *Operations Research*, vol. 46, no. 3, pp. 406-422, 1998.

[5] D. Bertsimas, G. Lulli, and A. R. Odoni, " An Integer Optimization Approach to Large-Scale Air Traffic Flow Management ", *Operations Research*, vol. 59, no. 1, pp. 211-227, 2011.

[6] A. Agustin, A. Alonso-Ayuso, L. Escudero, and C. Pizarro, " On Air Traffic Flow Management with Rerouting. Part I: Deterministic Case ", *European Journal of Operational Research*, vol. 219, no. 1, pp. 156-166, May 2012.

[7] A. Agustin, A. Alonso-Ayuso, L. Escudero, and C. Pizarro," On Air Traffic Flow Management with Rerouting. Part II: Stochastic Case ", *European Journal of Operational Research*, vol. 219, no. 1, pp. 167-177, May 2012.

[8] G. Clare and A. Richards, " Air Traffic Flow Management under Uncertainty: Application of Chance Constraints ", in *Proceedings of the* 2^{nd} *International Conference on Application and Theory of Automation in Command and Control Systems*, ser. ATACCS '12. Toulouse, France, France: IRIT Press, 2012, pp.20-26.

[9] S. Oussedik, D. Delahaye, and M. Schoenauer, " Air Traffic Management by Stochastic Optimization ", in 2^{nd} *USA/Europe RD Seminar on AirTraffic Management*, 1998, pp. 1-10.

[10] N. Barnier, P. Brisset, and T. Rivière, " Slot Allocation with Constraint Programming: Models and Results ", in 4^{th} *USA/Europe RD Seminar on Air Traffic Management*, 2001.

[11] P. Flener, J. Pearson, M. Ågren, C. Garcia Avello, M. Celiktin, and S. Dissing, " Air-Traffic Complexity Resolution in Multi-Sector Planning using Constraint Programming ", in 7^{th} *USA/Europe RD Seminar on Air Traffic Management*, 2007.

[12] D. Delahaye, S. Oussedik, and S. Puechmorel, "Airspace Congestion Smoothing by Multi-Objective Genetic Algorithm ", in *Proceedings of the 2005 ACM symposium on Applied computing*, ser. SAC '05. New York,ACM, 2005, pp. 907-912.

[13] G. Marceau Caron, P. Savéant and M. Schoenauer, "Multiobjective tactical planning under uncertainty for air traffic flow and capacity management ",*IEEE Congress On Evolutionary Computation. IEEE Computational Intelligence Society*,2013.

[14] A. Hadjaz, G. Marceau Caron, P. Savéant and M. Schoenauer, "Increasing Air Trafc: What is the Problem ? ",*The Second SESAR Innovation Days* ,2012.

[15] P. J. Rousseeuw,(1987). " Silhouettes: a graphical aid to the interpretation and validation of cluster analysis ". *Journal of Computational and Applied Mathematics*. Vol.20, pp.53-65.

[16] Robert L. Thorndike (December 1953). " Who Belongs in the Family ? ". Psychometrika .18 (4): 267-276. doi:10.1007/BF02289263.